Multiscale modeling of charge transfer and release in doped organic semiconductors

Massimiliano Comin

THÈSE

Pour obtenir le grade de
Docteur de l'Université Grenoble Alpes
Spécialité : Physique Théorique
Arrêté ministériel : 25 mai 2016

Présentée par Massimiliano Comin.

Thèse dirigée par Xavier Blase (Institut Néel, CNRS & UGA)
et co-dirigée par Gabriele D'Avino (Institut Néel, CNRS & UGA)

Préparée au sein de l'Institut Néel, CNRS
Avec l'École doctorale de physique de Grenoble.

Multiscale Modeling of Charge Transfer and Release in doped Organic Semiconductors

Modélisation Multi-échelle du Transfer et de la Séparation de Charge dans les Semiconducteur Organiques Dopés

Thèse soutenue publiquement le 29 mars 2022,
Devant le jury composé de :

Frédéric Castet Professeur des universités, Université de Bordeaux.	Rapporteur
Claudio Attaccalite Chargé de recherche, CNRS.	Rapporteur
Thierry Klein Professeur des universités, Université Grenoble-Alpes.	Président
Roberta Poloni Chargée de recherche, CNRS.	Examinatrice

Acknowledgements

This Thesis would not have been possible without the constant assistance and encouragement of many people.

This research has received funding from the *Grenoble Quantum Engineering* and the *Fondation Nanosciences* doctoral programs as well as the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie grant agreement No. 754303.

I would like to thank Frédéric Castet, Claudio Attaccalite, Roberta Poloni and Thierry Klein who kindly accepted to be part of my Ph.D. jury, as well as Denis Basko and Roberta Poloni for their participation to my *Comité de Suivi Individuel*.

I would like to express my gratitude to my supervisors Xavier Blase and Gabriele D'Avino for their outstanding guidance and their continuous support during the last three years. Xavier always provided insightful directions with an exceptional blend of pedagogy and patience. I am indebted to Gabriele for his invaluable scientific teachings, his well-calibrated incentives, his approach to problem solving (inherited from Mr. Wolf), and for the countless hours he has spent guiding me and helping me with *teh* codes I wrote. I would also like to thank Simone Fratini for the shared parenthood and his involvement throughout this Thesis via the many thrilling discussions and research ideas.

During my short stay in Karlsruhe, I was warmly welcomed by the Nanomatch team. My sincere thanks to Tobias Neumann, Timo Strunk and Franz Symalla for having integrated me in their group and for making me work on exciting subjects in a very pleasant working atmosphere, as well as for taking the time to introduce me to German cuisine and beer before I went back to Grenoble.

I would like to thank all the staff of the Néel Institute, and in particular Élodie Ferrara and Patrick Belmain for their kindness and helpfulness. My time at the Institute would not have been so pleasant without all the discussions, coffee breaks, foosball games, skiing sessions or after work beers. My sincere thanks to my fellow theoreticians Adolfo G. Grushin, Maxime Debertolis, Quentin Marsal, Richard East, and Sergueï Tchoumakov (who has this strange habit of always finding himself in the same lab as me).

I have been very lucky to share my office and my life as a doctoral student with three exceptional friends. Thanks to Theo Sepulcre, our local LaTeX supreme guide, for all the stimulating discussions and random anecdotes. I am grateful to Miguel Ángel Sánchez Martínez (Rodríguez González López), who blessed us with his extraordinarily expressive way of laughing. I would like to thank Katee Driscoll for her irradiating good mood and cheerfulness, and for all the good times spent together.

This paragraph is dedicated to another kind of mammal: a cat named Kum Sujuk, whose purrs and cuddles strongly contributed to my survival while I was struggling to write this Thesis.

Je suis infiniment reconnaissant envers mes amis, mes compagnons de mesure et de démesure lors de nos soirées de jeux de société ou de débauche. Je tiens à remercier Hugo pour son amitié inconditionnelle et intemporelle. Je remercie pour leur sincère amitié ceux avec qui j'ai grandi, Bastien, Franco, Juliette et Thibault, mais aussi Claire, Hugo, Jean, Manon, Marine, Marine, Myriam et Victor pour tous les moments passés ensemble. Merci aussi à Anas, Charlotte et Etienne pour leur amitié transatlantique. Cette Thèse, et bien d'autres choses, n'auraient jamais vu le jour sans vous. Merci bande de narvalos.

Vorrei dedicare questa Tesi ai miei genitori, Fabio e Letizia, e ai miei fratelli, Andrea e Niccolò, a cui devo tutta la mia gratitudine per l'aiuto, l'incoraggiamento e l'amore avuto da parte loro.

Pour finir, mes pensées vont de tout mon coeur à Elise, qui m'a supporté durant cette période avec une patience exemplaire.

Massimiliano

Contents

List of Figures

List of Tables

List of Acronyms

BCC	Body-Centered Cubic
BSE	Bethe-Salpeter Equation
BZ	Brillouin Zone
CBS	Complete Basis Set
CME	Clausius-Mossotti Equation
CP-SCF	Coupled-Perturbed Self-Consistent Field
CR	Charge Response
CT	Charge Transfer
CTCs	Charge-Transfer Complexes
DFT	Density-Functional Theory
DoS	Density of States
EA	Electron Affinity
EMBPT	Embedded Many-Body Perturbation Theory
FCC	Face-Centered Cubic
FCT	Fractional Charge Transfer
FT	Fourier Transform
GGA	Generalized Gradient Approximations
GTO	Gaussian-type orbital
HEG	Homogeneous Electron Gas
HF	Hartree-Fock
HOMO	Highest Occupied Molecular Orbital
ICT	Integer Charge Transfer
IP	Ionization Potential
IPES	Inverse Photoemission Spectroscopy
KMC	Kinetic Monte Carlo
KS	Kohn-Sham
LDA	Local Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MBPT	Many-Body Perturbation Theory
ME	Micro-Electrostatics

MM . . . . . Molecular Mechanics

OSC . . . . . Organic Semiconductor

PCF . . . . . Pair Correlation Function
PCM . . . . . Polarizable Continuum Model
PES . . . . . Photoemission Spectroscopy
PR . . . . . Participation Ratio

QM . . . . . Quantum Mechanics
QM/MM . . . . . Quantum/Classical
QMC . . . . . Quantum Monte Carlo

RDF . . . . . Radial Distribution Function
RKS . . . . . Restricted Kohn-Sham
ROKS . . . . . Restricted-Open Kohn-Sham

SC . . . . . Simple Cubic
SCF . . . . . Self-Consistent Field
SOS . . . . . Sum-Over-States
STO . . . . . Slater-Type Orbital

UKS . . . . . Unrestricted Kohn-Sham

List of Publications

Published

M. COMIN AND L. J. LEWIS. "Deep-Learning Approach to the Structure of Amorphous Silicon". *Physical Review B* 100, (2020).

J. W. R. MACDONALD, G. PIANA, **M. COMIN**, E. VON HAUFF AND G. KOCIOK-KÖHN, C. BOWEN, P. LAGOUDAKIS, G. D'AVINO AND E. DA COMO. "Charge Transfer Excitons in a Donor–Acceptor Amphidynamic Crystal: The Role of Dipole Orientational Order". *Materials Horizons* 7, (2020).

P.-F. LOOS, **M. COMIN**, X. BLASE AND D. JACQUEMIN. "Reference Energies for Intramolecular Charge-Transfer Excitations". *Journal of Chemical Theory and Computation* 17, (2021).

M. COMIN, S. FRATINI, X. BLASE AND G. D'AVINO. "Doping-Induced Dielectric Catastrophe Prompts Free-Carrier Release in Doped Organic Semiconductors". *Advanced Materials*, (2021).

M. COMIN, V. LEMAUR, A. GIUNCHI, D. BELJONNE X. BLASE AND G. D'AVINO. "Doping of semicrystalline conjugated polymers: dopants within alkyl chains do it better". *Journal of Materials Chemistry C*, (2022).

In Preparation

M. COMIN, S. FRATINI, X. BLASE AND G. D'AVINO. "Dielectric Catastrophe in Doped Organic Semiconductors: Study of Collective Screening Phenomena in Inhomogeneous Media". *In preparation, (2022)*.

A. GIUNCHI, **M. COMIN**, AND G. D'AVINO. "Comparing formalisms for polarization energies in organic semiconductors: from Micro-Electrostatics to Many-Body Perturbation Theory". *In preparation, (2022)*.

Abstract

The success of electronic and optoelectronic technologies relies on the possibility to control the charge carrier concentration in materials and modulate their electric properties by introduction of dopant impurities. The physics of doping is well understood in the context of inorganic semiconductors, in which the advent of shallow donor or acceptor impurity levels is correctly predicted within the Hydrogenic model. Conversely, the mechanism for molecular doping in organic semiconductors is believed to be qualitatively different. These excitonic semiconductors typically reach the degenerate limit only at impurity concentrations of 5-10%, which are orders of magnitude larger than those needed in their inorganic analogues. This has been related to the Coulomb binding between ionized dopants and charge carriers, which is particularly strong in organic materials featuring low dielectric constants. The mechanisms determining the charge release upon doping and the ensuing conductivity enhancements remain elusive so far.

Doping in organic semiconductors has been depicted as a two-step process, namely the ionization of dopant impurities and the subsequent release of free charges available for conduction. The present Thesis investigates these two aspects by means of a multiscale framework encompassing many-body *ab initio* electronic structure approaches, parameterized Hamiltonian models and classical polarizable models.

By taking the technologically relevant case of a doped polymer as a case study, our calculations target the ground and excited state properties of host-dopant complexes, drawing a coherent picture of the different factors at play in the ionization process such as the Coulomb electron-hole (excitonic) binding, environmental electrostatic interactions and the crucial role of the position of the dopant in the polymer structure. By combining many-body perturbation theory with the Micro-Electrostatic framework, our results explain the striking differences in conductivity arising from samples with different morphologies, and confirm the appearance of low-lying charge-transfer excitations from the dopant to the host semiconductor, as the first step to dopant ionization.

We have then focused on the release of carriers at finite doping loads, for which we propose a general mechanism in terms of collective screening phenomena. A multiscale model for the dielectric properties of doped organic semiconductors is set up by combining first principles and Micro-Electrostatic calculations. Our results predict a large nonlinear enhancement of the dielectric constant (tenfold at 8% load) at doping concentrations comparable to those determining orders-of-magnitude conductivity enhancements in experiments. The system approaches a dielectric catastrophe upon increasing doping, which is attributed to the presence of highly polarizable host–dopant complexes. The leading contribution, as compared to the Clausius-Mossotti relation applied on an effective homogeneous polarizability, arises from the formation of soft and eventually unstable polarization modes. The enhanced screening in the material drastically reduces the (free) energy barriers for electron–hole separation, rationalizing the possibility for thermal charge release. Our results suggest that such a doping-induced dielectric catastrophe represents a driving factor for the insulator-to-metal transition in doped organic semiconductors.

Résumé

Le succès des technologies électroniques et optoélectroniques repose sur la possibilité de contrôler la concentration des porteurs de charge dans les matériaux et de moduler leurs propriétés électriques par l'introduction d'impuretés dopantes. La physique du dopage est bien comprise dans le contexte des semi-conducteurs inorganiques, dans lesquels l'apparition de niveaux d'impuretés donneurs ou accepteurs peu profonds est correctement prédite dans le cadre du modèle Hydrogénoïde. En revanche, le mécanisme de dopage moléculaire dans les semi-conducteurs organiques est qualitativement différent. Ces semi-conducteurs excitoniques n'atteignent généralement la limite dégénérée qu'à des concentrations d'impuretés de 5 à 10%, qui sont des ordres de grandeur supérieures à celles requises dans leurs analogues inorganiques. Ce phénomène a été mis en relation avec l'interaction de Coulomb entre les dopants ionisés et les porteurs de charge, qui est particulièrement forte dans les matériaux organiques à faible constante diélectrique. Les mécanismes qui déterminent la libération de charge lors du dopage et l'augmentation de la conductivité qui en résulte restent à ce jour encore méconnus.

Le dopage des semi-conducteurs organiques peut être décrit comme un processus en deux étapes, à savoir l'ionisation des impuretés dopantes et la libération ultérieure de charges libres, disponibles pour la conduction. Cette Thèse se propose d'étudier ces deux aspects par le biais d'un formalisme multi-échelle englobant des approches de structure électronique *ab initio* à N corps, des modèles Hamiltoniens paramétrés ainsi que des modèles polarisables classiques.

En prenant comme cas d'étude un polymère dopé technologiquement important, nos calculs portant sur les propriétés de l'état fondamental et des états excités des complexes hôte-dopant dressent un tableau cohérent des différents facteurs en jeu dans le processus d'ionisation, tels que la liaison électron-trou (excitonique) de Coulomb, les interactions électrostatiques environnementales et le rôle crucial de la position du dopant dans la structure du polymère. En combinant la théorie des perturbations à N corps avec l'approche Micro-Électrostatique, nos résultats expliquent les remarquables différences de conductivité associées à des échantillons de morphologies différentes, et confirment l'apparition d'excitations optiques à transfert de charge de faible énergie résultant du dopage comme première étape du mécanisme d'ionisation des dopants.

Nous nous sommes ensuite intéressés à la séparation de charge à dopage fini, pour laquelle nous proposons un mécanisme général en termes de phénomènes d'écrantage collectif. Un modèle multi-échelle pour les propriétés diélectriques des semi-conducteurs organiques dopés est mis en place en combinant des approches *ab initio* avec des calculs Micro-Électrostatiques. Nos résultats prédisent une forte augmentation non linéaire de la constante diélectrique (décuplée à 8% de dopage) à des concentrations d'impuretés semblables à celles correspondant expérimentalement à des augmentations de conductivité de plusieurs ordres de grandeur. Le système tend vers une catastrophe diélectrique en fonction du dopage, ce qui est attribué à la présence de complexes hôte-dopant hautement polarisables. La contribution principale, en comparaison avec la relation de Clausius-Mossotti appliquée à une polarisabilité homogène effective, provient de la formation de modes collectifs de polarisation à très basse énergie et potentiellement instables. Le fort écrantage résultant dans le matériau réduit considérablement la barrière d'énergie (libre) pour la séparation électron-trou, permettant donc la libération thermique des porteurs de charge. Nos résultats suggèrent qu'une catastrophe diélectrique induite par le dopage représente un facteur moteur pour la transition isolant-métal dans les semi-conducteurs organiques dopés.

Chapter 1

Introduction

1 Introduction

"Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

Lewis Carroll

ALL KNOWN LIFEFORMS ARE MADE OF CARBON, the chemical element forming the backbone of every organic molecule. Its ability to form single, double or triple bonds allows to form long linear chains and aromatic rings, resulting in strong and stable compounds ranging from DNA and proteins to organic molecules and polymers, graphene sheets or diamond. Combined with other elements such as Hydrogen, Nitrogen and Oxygen, the conceivable structures pertain to a set of endless possibilities. These properties give organic materials some of their unique properties as compared to their inorganic counterparts, such as Silicon. Nonetheless, being the second most abundant element on Earth after Oxygen, Silicon has served from Antiquity to the Modern Era for countless purposes ranging from architecture to nanotechnology.

A key step made towards understanding the vibrant complexity of phenomena taking place in solids can be traced back to the beginning of the 20$^{\text{th}}$ century. In his *Annus Mirabilis*, Albert Einstein explained the photoelectric effect as the absorption of discrete light *quanta* (see Fig. 1.1) [1], an idea that later led to the development of Quantum Mechanics [1]. Throughout the rest of the century, deciphering the electronic structure of materials has ranked among the greatest challenges in theoretical physics. Phenomenal success has been achieved in that matter, eventually resulting with the semiconductor revolution that followed the invention of the transistor in 1947 [2], a small semiconductor device which has had a profound impact on information technology as well as far-reaching consequences on our societies.

Most of the research interests were then focused on the electronic properties of inorganic semiconductors such as Silicon. The enthusiasm for the study of organic materials originated from the discovery of the conducting properties of π-conjugated molecular crystals and polymers in the 70s [3], which gave birth to a dynamic and strongly multidisciplinary branch of materials science. In the following decades, the research in the field progressively invaded areas that were considered of exclusive domain of inorganic materials such as electronics, photonics, magnetism, with more recent new twists on thermoelectricity [2] and spintronics [3]. Because of their nature of synthetic plastic materials, organics present several advantages over inorganic semiconductors and functional oxides. They are intrinsically cheap, based on abundant and non-toxic elements, they are suitable for the easy production of bio-compatible, light-weighted, mechanically flexible and possibly optically transparent devices [4]. Furthermore, the virtually infinite number of compounds that can be synthesized offers an unprecedented opportunity for the tuning of the sought physical properties. Last but not least, the prospect of better understanding biological systems at the microscopic level lies ahead of us as a vast *terra incognita* to explore.

[1] Einstein (1905).

1. Albert Einstein was awarded the 1921 Physics Nobel Prize *"for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect"*.

2. Bell Labs scientists John Bardeen, Walter Brattain, and William Shockley won the 1956 Nobel Prize in Physics *"for their researches on semiconductors and their discovery of the transistor effect"*.

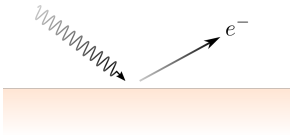

Figure 1.1: Illustration of the photoelectric effect: electron emission from a metal surface is induced by impinging photons.

3. The 2000 Nobel Prize in Chemistry was awarded jointly to Alan J. Heeger, Alan G. MacDiarmid and Hideki Shirakawa *"for the discovery and development of conductive polymers"*.

[2] Bubnova et al. (2011).

[3] Dediu et al. (2009).

[4] Forrest (2004).

A unifying characteristic of organic solids is their weak Van der Waals intermolecular interactions, leading to narrow bands, to a strong sensitivity to intramolecular and soft lattice vibrations, and to molecules that largely retain their identity in the crystal. At large (metallic) densities, electronic correlations and electron-phonon coupling compete and cooperate leading to complex and rich phase diagrams, including charge ordering, ferroelectricity [5, 6] and magnetoelectric phenomena [5]. On the other hand, in the low density limit (pristine semiconductor), charge transport is strongly hampered by the lattice motion, with only few molecular crystals having mobilities exceeding 1 cm s/V [7, 8, 9].

Doping of Organic Semiconductors (OSCs) by introduction in the host matrix of strong electron or hole donating molecules allows to control the carrier density, increase the electrical conductivity and align energy levels at interfaces [4, 10]. While recent years have witnessed remarkable advances concerning the chemistry and the structural control of doped organic semiconductors, the microscopic phenomena that govern molecular doping and the evolution of transport and optical properties with doping load remain to date unclear.

In particular, the electrical doping of semiconductors has attracted the interest of researchers since decades, in view of the fundamental physics underneath and its potential for applications. Doping indeed lies at the heart of countless developments in electronics, including organic light-emitting diodes and thermoelectricity. On the other hand, insulator-to-conductor transitions are arguably at the origin of two among the richest fields in condensed matter physics: the Mott transition in correlated electron systems and Anderson localization in disordered systems. While we now have a rather complete understanding of these phenomena in inorganic wide-band semiconductors, with basic ideas presented in textbooks, the understanding of the mechanism for molecular doping in organic semiconductors is still in its infancy and it is currently sparking an outstanding research effort [11, 12, 13, 14, 15, 16, 17].

The electrical doping of a semiconducting material consists of two elementary steps: (i) the charge transfer between the dopant impurity and the host semiconductor, leading to the formation of an electron-hole pair, and (ii) the charge on the host is released and becomes available for conduction. The standard picture of doping for inorganic semiconductors is the Hydrogenic model, developed by Kohn and Luttinger in 1955 [18]. In this model, shown in Fig. 1.2, the Charge Transfer (CT) between the impurity and the host semiconductor is always expected to happen spontaneously, due to the different number of valence electron between the host semiconductor (e.g. Silicon) and. donor and acceptor impurities (e.g. Phosphor, Boron). Besides, the strong covalent bonds and dispersive energy bands that are typical of inorganic semiconductors lead to an efficient delocalization of Bloch states and a large dielectric screening. The exceeding carrier introduced upon doping is then effectively delocalized, allowing the impurity to be modeled as a Hydrogenic system, accounting for the dielectric constant and effective mass determined by the material. The carrier's binding energy and effective radius are then given by (see, e.g. [19])

$$E_b = \frac{m^*}{\varepsilon^2}\,\mathrm{Ry} \tag{1.1a}$$

$$r = \frac{\varepsilon}{m^*}\,a_0, \tag{1.1b}$$

where ε is the static relative permittivity and m^* the effective mass of the material [4].

[5] Kagawa et al. (2010).

[6] D'avino et al. (2017).

[7] Podzorov et al. (2004).

[8] Fratini et al. (2016).

[9] Fratini et al. (2017).

[10] Salzmann and Heimel (2015).

[11] Walzer et al. (2007).

[12] Salzmann et al. (2012).

[13] Salzmann et al. (2016).

[14] Kang et al. (2016).

[15] Png et al. (2016).

[16] Gaul et al. (2018).

[17] Tietze et al. (2018).

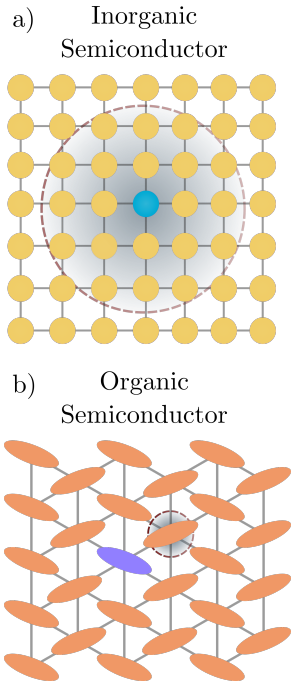

Figure 1.2: Illustration of the Hydrogenic model for a) inorganic semiconductors and b) organic semiconductors.

[18] Kohn and Luttinger (1955).

[19] Kittel (1976).

4. The values are given in Rydberg ($Ry = m_e e^4/8\varepsilon_0^2\hbar^2$) and Bohr radii ($a_0 = 4\pi\varepsilon_0\hbar^2/m_e e^2$).

Taking the values of Silicon, i.e. $\varepsilon \simeq 12$ and $m^* \simeq 0.2$, one finds an effective radius $r = 32$ Å, which is an order of magnitude greater than typical nearest-neighbor distances $r_0 \simeq 2$ Å, and a binding energy of $E_b = 19$ meV, which is comparable with the room temperature thermal energy $k_B T \simeq 26$ meV. Thus, dopant impurities in inorganic semiconductors feature shallow and thermally accessible impurity levels, resulting in the release of one free charge per dopant that is available for transport.

The situation is diametrically opposed in organic solids, featuring narrow bands and a low dielectric screening. Indeed, in such systems dopant ionization is not granted and the microscopic mechanisms controlling the charge transfer are still under debate. Moreover, considering the typical values of the effective mass $m^* = 2.5$ [20, 21, 22] and static dielectric constant of organic semiconductors $\varepsilon = 3.5$ [23, 24], the simple relations above yield a large binding energy $E_b = 2.78$ eV and a small effective radius $r = 0.74$ Å. Such an unphysical value of the effective radius is dramatically lower than the typical lattice constant of organic crystals $a \simeq 10$ Å and is at odds with the hypothesis that carriers are effectively delocalized. Moreover, according to the Hydrogenic picture, and in contrast with experiments, transport should not occur because of a too large binding energy. This signals the breakdown of the Hydrogenic model when applied naively to organic semiconductors, and calls for a more refined treatment taking into account the molecular nature of these materials.

[20] Troisi and Orlandi (2005).

[21] Doi et al. (2005).

[22] Hatch et al. (2009).

[23] Tavazzi et al. (2008).

[24] Dressel et al. (2008).

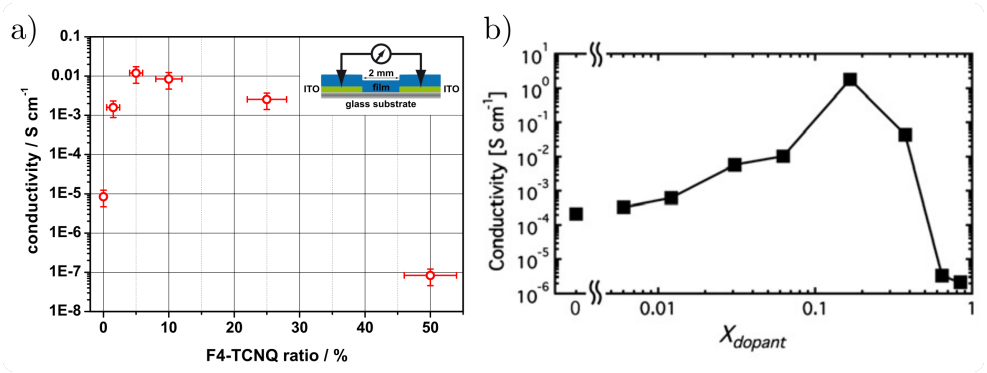

Figure 1.3: Conductivity of Pentacene (a)[12] and P3HT (b)[25] doped with F4TCNQ as a function of the doping load, expressed as a molar fraction, i.e. dopants per semiconductor molecule or monomer unit.

A central open question thus concerns the release of free doping-induced charge carriers in organic materials. Experimental data for small molecules and polymers show that a boost of the electrical conductivity by orders of magnitude is achieved at doping loads of about 5-10% [26, 11, 27, 28, 12, 29, 13, 30], as shown for a molecular and a polymer system in Fig. 1.3. These are orders-of-magnitude higher than those adopted for inorganic semiconductors ($\sim 10^{-5}$), and it is presently unclear why such large dopant concentrations are needed to boost their electrical conductivity.

[26] Gao and Kahn (2003).

[27] Reineke et al. (2009).

[28] Harada et al. (2010).

[29] Kleemann et al. (2012).

[30] Jacobs and Moulé (2017).

[31] Jacobs et al. (2018).

Clues can be found in that the doping efficiencies, as measured by the number of free charges per dopant unit, are often limited by the partial degree of charge transfer between the host OSC and molecular dopants [17, 31]. Even in the case of full charge transfer, experimental studies suggest that only a fraction of doping-induced charges contributes significantly to transport while the rest of the electron-hole pairs, located on the ionized dopant and host semiconductor molecules, remain Coulombically bound [32].

[32] Pingel and Neher (2013).

Moreover, several studies have shown that the conductivity follows a thermally activated behavior [33, 34, 35, 36], and the activation energy has been related to the Coulomb interaction between an ionized dopant and the charge injected in the semi-

[33] Maennig et al. (2001).

[34] Shen et al. (2003).

[35] Olthof et al. (2012).

[36] Schwarze et al. (2019).

conductor [36]. Photoemission measurements [16] and theoretical calculations [37, 38, 39] set this binding energy in the 400-700 meV range in the low-doping regime, a value that is too large to permit a significant release of free carriers at room temperature. Experimental data for both p-doped and n-doped OSCs have shown a universal tendency for a conspicuous reduction of the activation energy with dopant concentration [36, 34, 35, 40]. Besides, the conductivities of a set of polymers heavily-doped by ion-exchange have been shown to be independent on the ion size [41], which is at odds with the interpretation of transport limited by Coulomb interactions. This set of evidences, together with the high loads needed to boost conductivity, points to a collective depinning mechanism for charge carriers taking place upon increasing doping concentration, beyond trap filling effects at ultralow doping [35, 42].

[37] Li et al. (2017).

[38] Li et al. (2019).

[39] Privitera et al. (2020).

[40] Zhang and Kahn (2018).

[41] Jacobs et al. (2021).

[42] Tietze et al. (2015).

Theoretical studies were able to partly reconcile the apparently contrasting pictures for the mechanisms of dopant ionization in organic semiconductors by studying the paradigmatic case of Pentacene doped with F4TCNQ [43, 29, 12]. By employing state-of-the-art many-body techniques, Li et. al [37] demonstrated that full dopant ionization takes place at room temperature in spite of a large photoemission gap of ~700 meV because of a large excitonic electron-hole binding energy, stabilizing the charge-transfer excitation energy down to ~30 meV. While such a large exciton binding energy allows for spontaneous dopant ionization at room temperature, it also determines a charge carrier that remains strongly bound to its parent dopant. Thus, the excitonic electron-hole interaction can be considered both as a blessing and a curse, since it is necessary for the charge transfer process between the dopant and the host semiconductor but also strongly impedes the release of free charges available for conduction.

[43] Ha and Kahn (2009).

The above considerations bring us to the central fundamental question we address in the present Thesis:

How is it possible to reconcile the observed conductivity increase in organic semiconductors at large doping loads with the evidence for strongly bound carriers ?

The beginning of an answer might well lie in the low-energy excitations arising in host-dopant complexes, since these could strongly contribute to the polarizability. At doping loads relevant for organic semiconductors, highly polarizable complexes would eventually result in a large boost of the dielectric constant, thereby screening the Coulomb interactions keeping charge carriers strongly bound to dopant ions. Together with the experimental data mentioned above, this suggests a scenario in which the insulator-to-conductor transition takes place at increasing doping loads because of collective screening phenomena. Such a scenario, hinging essentially on the Clausius-Mossotti, or Lorentz-Lorenz, non-linear relation between molecular polarizability and the long-range macroscopic dielectric constant, has been discussed by Herzfeld [44] in an early attempt to relate atomic properties to potential metallic behavior. The transition from the insulator side is explained by the divergence of the dielectric constant, resulting in a material's response that perfectly counteracts any external field, a phenomenon dubbed *dielectric catastrophe*. In the context of inorganic semiconductors, such a transition was predicted by Mott [45] and first reported for doped Silicon [46]. So far, this phenomenon has not been discussed in the context of organic semiconductors, for which we propose an analogous mechanism supported by a rigorous theoretical analysis.

[44] Herzfeld (1927).

[45] Mott and Littleton (1938).

[46] Rosenbaum et al. (1983).

It is interesting to give away qualitatively the forthcoming developments for the impatient reader. Let us model a paradigmatic organic material as a face-centered cubic lattice, where each site represents a molecule with polarizability $\alpha_0 = 50\,\text{Å}^3$. A reasonable lattice constant would be A =12.26 Å, which is selected to give a dielectric constant representative of OSCs $\varepsilon = 3.5$. Consider now that doping this material results in a spontaneous host-dopant charge transfer, with a lowest optical excitation of ~30 meV, similar to what reported for F4TCNQ-doped Pentacene [37]. Then, the polarizability of the dopant site can be approximated as $\alpha_1 \simeq 2\,\langle\mu\rangle^2 / S_1$, where $\langle\mu\rangle$ is the transition dipole moment. A typical charge transfer between a donor and acceptor molecule results in a transition dipole moment [5] of ~3 e Å, giving a dopant polarizability α_1=600 Å^3. Consequently, according to the Clausius-Mossotti relation applied to the effective polarizability of the doped material $\alpha_{\text{eff}} = (1-\rho)\alpha_0 + \rho\alpha_1$, where ρ is the doping load, the dielectric constant is $\varepsilon = 64$ at 10% doping load. This considerable enhancement of the dielectric constant would then result in a screened Coulomb potential between two neighboring charges of $V_{\text{eh}} = 1/\varepsilon a \simeq 26$ meV, which is precisely of the order of the thermal energy $k_B T \simeq 26$ meV, at room temperature (T =300 K). Therefore, it may well be possible that the advent of a doping-induced dielectric catastrophe is the sought collective mechanism explaining the charge release in doped organic semiconductors. Of course, this intuition will be developed and demonstrated in the following.

5. The transition dipole moment can be calculated exactly in the Mulliken two-state Donor-Acceptor model as

$$\mu_{CT} = \mu_0 \sqrt{\rho(1-\rho)}$$

where ρ is the amount of transferred charge. With a typical distance between molecules of 6 Å, and assuming a partial charge transfer $\rho = 0.5$, the transition dipole moment is found to be 3 e Å.

The Thesis is organized as follows. After a brief discussion of the experimental probes of neutral and charged excitations in Chapter 2, we set the stage by presenting the theoretical methods that pertain to the calculation of the ground-state and excited properties of organic semiconductors in Chapter 3. We will discuss and compare Density-Functional Theory, Many-Body Perturbation Theory, and specifically the *GW* and Bethe-Salpeter formalisms, Micro-Electrostatics, and their combination resulting in a multiscale approach to electronic excitations.

Chapter 4 will be devoted to the study of ionization efficiency in a typical crystalline conjugated polymer (PBTTT). We will investigate the intimate relationships between the electronic properties and the structure of the doped polymer, with specific reference to the position occupied by the dopant in the polymer lamellae. This will reveal the major role of the electrostatic landscape of the host polymer, which affects the energetics of the charge transfer process, ultimately leading to qualitatively different ground states for dopants placed within the conjugated backbone or in the Alkyl side chains.

Finally, in Chapter 5 we unravel the collective screening phenomena that underlie charge release in doped organic semiconductors. Following a multiscale approach, the polarizability associated with low-energy charge-transfer degrees of freedom in host-dopant complexes will be assessed. We will discuss the advent of a doping-induced dielectric catastrophe, taking place at the doping loads corresponding to the experimental conductivity boosts (a few %), rationalizing the possibility of free carrier release in doped organic semiconductors.

The final chapter will summarize the main findings of the present Thesis, and discuss the future perspectives that branch out of this research work.

Chapter 2

Probing the Atomic and Electronic Structure

Probing the Atomic and Electronic Structure

2

He who can properly define and divide is to be considered a god.

Plato

Contents

THE ADVENT OF SPECTROSCOPIC TECHNIQUES for the characterization of solids revolutionized our ability to understand their nature by providing access to their structural, electronic and vibrational properties. The aim of this chapter is to establish in a non-exhaustive fashion the experimental context in which the structural and electronic properties of materials are probed. Particular attention will be given on the microscopic phenomena at play and the relationships with their theoretical modeling, since these will serve as a common thread for the discussion that will follow.

2.1 STRUCTURAL EXPERIMENTS

Discovered in 1895, X-rays are a type of high-energy electromagnetic radiation with wavelengths comparable to typical inter-atomic distances in solids. They have been widely used for imagery experiments and are still nowadays a tool of choice for determining the atomic structure of materials.

X-ray diffraction experiments are operated by shining radiation on the system under study and measuring the scattered intensity as a function of the deviation angle. The energy of X-rays being much greater than that of valence electrons, their absorption and refraction are negligible and the interaction may be modeled as the elastic scattering of an electromagnetic wave by free electrons, known as Thomson scattering [1]. Thus the scattered waves maintain the same amplitude, and each point in the sample at position $\mathbf{r}$ deviates the beam by a momentum $\mathbf{q}$. This results in a scattered intensity $I(\mathbf{q})$ being simply the Fourier transform of the electronic density $n(\mathbf{r})$

$$I(\mathbf{q}) = \left| \int d\mathbf{r}\, n(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \right|^2 . \tag{2.1}$$

1. X-rays have energies of 10^2-10^5 eV and thus wavelengths of $10^{-1} - 10^2$ Å, whereas typical energies of valence electrons are a few eV and typical inter-atomic distances are a few Å.

When the atomic positions in the sample feature a periodic structure, the electronic density has the same periodicity and its only non-vanishing components are the reciprocal lattice vectors $\mathbf{G}$. Thus the condition for scattered photons to interfere constructively is that $\mathbf{q} = \mathbf{G}$: this the famous Bragg's law depicted in Fig. 2.1.

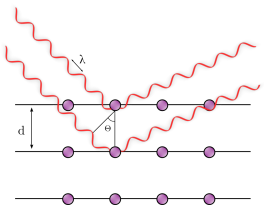

Figure 2.1: Illustration of Bragg's law.

In contrast with perfect crystals, organic materials (and especially polymers) often feature structural disorder due to the weakness of electrostatic and dispersion interactions that are responsible for their bonding. A moderate disorder translates into

a broadening of the diffraction peaks. Many organic compounds and polymers are amorphous, meaning they lack any long-range order. In this case the static structure factor, defined as $S(\mathbf{q}) = I(\mathbf{q})/N \langle f(\mathbf{q}) \rangle^2$ where $f(\mathbf{q})$ is the atomic structure factor, can be related to the real space structure as

$$S(\mathbf{q}) = 1 + \bar{n} \int d\mathbf{r}\, g(r) e^{i\mathbf{q}\cdot\mathbf{r}}, \tag{2.2}$$

where $g(\mathbf{r})$ is called the Radial Distribution Function (RDF) (or Pair Correlation Function): it designates the average density in a spherical shell at distance $\mathbf{r}$ from any particle. Being a one-dimensional information, it is impossible to recover the full atomic structure of amorphous materials from X-ray diffraction experiments. Thus one must rely on atomic-scale simulations validated against the RDF to obtain the atomic positions of such systems [2].

2. For example, my M.Sc. work [47] treated the application of Deep Learning algorithms to the modeling of the structure of amorphous silicon.

2.2 Photoemission Experiments

The discovery of the photoelectric effect by Hertz and Hallwachs in 1887 and its subsequent formalization by A. Einstein in 1905 led to the development of photoemission spectroscopy experiments, in which the electron binding energy is assessed by shining polarized light sources [3] on a material and measuring the resulting photoelectric current [48]. Photoemission spectroscopy techniques allow not only to assess the electronic band structure of the system under study, but also many other properties such as the Fermi surface and the electron-phonon coupling. When low-energy sources are used, the penetration depth is only of a few atomic layers, making photoemission spectroscopy a particularly surface-sensitive technique.

3. The photon energies used in photoemission experiments range from Ultraviolet (Ultraviolet Photoemission Spectroscopy, UPS) to X-ray (X-ray Photoelectron Spectroscopy, XPS) i.e. from 10 to 10^3 eV.

4. The Auger effect designates the electron emission upon filling of an inner shell vacancy by another electron. Auger electron spectroscopy is particularly used for probing the chemical and compositional surface environments of materials.

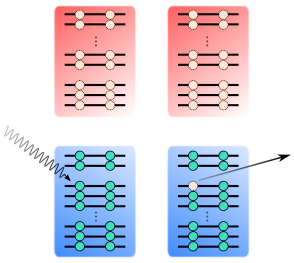

Figure 2.2: Single-particle picture of the photoemission process.

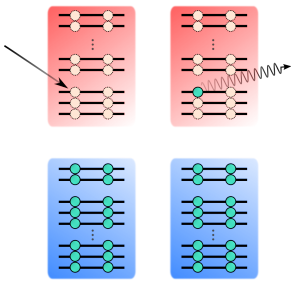

Figure 2.3: Single-particle picture of the inverse photoemission process.

In Direct Photoemission Spectroscopy (PES) experiments a photon impinges the sample and if its energy $\hbar\omega$ is larger than the binding energy of an electron, the latter is eventually extracted. This inherently many-body event is usually understood phenomenologically as a three-step process [49]. Firstly, the incoming photon triggers a direct optical transition between an occupied and an unoccupied electronic state, following the allowed dipole transition selection rules, and possibly exciting phonons. The generated hole may decay radiatively or non-radiatively through the Auger effect [4]. Secondly, the promoted electron propagates towards the surface of the sample and scatters elastically and inelastically with other electrons on its way, thus altering its energy and momentum, and eventually extracting other electrons. Thirdly, the electron escapes the surface potential barrier into a free vacuum state, losing an energy that is equal to the difference between the electrostatic energy of an electron outside the sample and the Fermi level. The latter is known as the work function ϕ, and is a fixed characteristic of the material's surface.

Although this description in terms of particles and quasiparticles represents already a simplification with respect to the full many-body problem, further approximations are commonly done. The hole creation is assumed to be instantaneous and to have no effect on the electronic structure under the frozen-orbital approximation. Moreover, the extrinsic losses during the propagation of the photoelectron are neglected in the so-called sudden approximation, and the phonon coupling and the decay of the hole are also neglected. These features typically translate into satellite peaks in the measured spectrum. Finally, the energy dependence of the electron-photon scattering cross-section is also not considered.

These considerations allow us to draw a single-particle picture of the photoemission

process, as depicted in Fig. 2.2. A photon of energy $\hbar\omega$ impinges the system which is in its N-electron ground-state of energy $E_0[N]$. The photoelectron is sent to the vacuum and the system remains in an excited $(N-1)$-electron state of energy $E_i[N-1]$, where i labels the i^{th} excited state of the system. The conservation of energy yields

$$\hbar\omega + E_0[N] = E_K + E_i[N-1], \tag{2.3}$$

where E_K is the kinetic energy acquired by the electron in the vacuum. This expression allows to define the photoemission electronic levels corresponding to the energy required to expel an electron from the system, namely

$$\epsilon_i = E_0[N] - E_i[N-1],. \tag{2.4}$$

In the particular case of an electron expelled from the Highest Occupied Molecular Orbital (HOMO), the associated Ionization Energy (or Ionization Potential, IP) corresponds to the minimum energy required to extract an electron from the system, i.e. $IP = E_0[N-1] - E_0[N]$, where the IP is taken by convention to be a positive quantity, as depicted in Fig. 2.4. Since the IP is formally defined as a difference of ground-state total energies between the systems with N and $N-1$ electrons, it is not in general related to ground-state properties of the neutral N-electron system but rather to charged excitations corresponding to electron removal energies.

On the other hand, Inverse Photoemission Spectroscopy (IPES) allows to probe the unoccupied states of a material by impinging an electron beam (whose kinetic energy is ~ 20 eV) on the sample and measuring the energy of the emitted radiation.

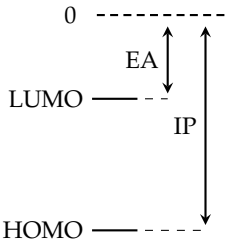

Figure 2.4: Energy level diagram depicting the convention for the signs of the IP and EA.

Assuming the same approximations as in the case of direct photoemission, the inverse photoemission process can be described phenomenologically as the instantaneous capture of an electron in an unoccupied state and its subsequent radiative decay, as depicted in Fig. 2.3. The conservation of energy reads

$$E_k + E_0[N] = \hbar\omega + E_i[N+1]. \tag{2.5}$$

It is then possible to define the inverse photoemission electronic level, i.e. the binding energy of the captured electron, as

$$\epsilon_i = E_i[N+1] - E_0[N]. \tag{2.6}$$

In the particular case of an electron captured in the Lowest Unoccupied Molecular Orbital (LUMO) the associated Electron Affinity (EA) corresponds to the minimum energy released upon electron capture, i.e. $EA = E_0[N] - E_0[N+1]$, where the EA is taken by convention to be a positive quantity for bound states, as depicted in Fig. 2.4. Again, this is a difference of total energies and is not in general a ground-state property but rather a charged excitation corresponding to the electron addition energy.

Upon performing PES and IPES on the same systems it is possible to obtain the entire photoemission spectrum, i.e. a map of addition and removal electron energies. We will be concerned with finding a theoretical formalism properly accounting for the (direct or inverse) photoemission process so as to yield accurate electronic energy levels associated with charged excitations of the system. We will see in particular that ground-state formalisms show a moderate success in the prediction of the photoemission spectrum because relaxation effects upon electron addition or removal play an important role in the description of the electronic structure (see

Sec. 3.2.5). Indeed, one needs a quasiparticle formalism such as the *GW* approximation in order to capture these effects since the latter captures perturbatively the screening of the additional charges by the other electrons in the system.

2.3 Optical Absorption Experiments

Absorption spectroscopy experiments proceed by shining a light source on a thin slab of the material under study and measuring the transmitted intensity with a photodetector for different incoming light frequencies. This results in a direct measure of the intensity loss and thus the macroscopic optical constants; the latter can further be related to the optical transitions in the material [5]. The infrared, visible and ultraviolet spectral regions span a range of energies that are close to the typical gaps of semiconductors, and optical absorption experiments were among the first methods developed to characterize them. One should note that contrary to photoemission experiments that scan all electronic states, optical absorption spectroscopy can only probe dipole-allowed transitions. As another crucial difference, such excitations do not change the particle number, making them neutral excitations where both electrons and holes may interact.

5. Each spectral region accounts for different degrees of freedom; for instance rotational lines are often found in the microwave region, vibrational lines in the infrared, electronic lines in the visible and ultraviolet and inner electronic shell lines in the X-ray region.

The optical absorption is related to the electronic structure of the system by considering the propagation of the impinging electromagnetic radiation through the sample. The solution of Maxwell's equations in a medium without source are plane waves with a complex wavevector $\tilde{k} = \omega/c\sqrt{\varepsilon\mu} + i\sigma\omega\mu/c^2$, where ω is the angular frequency, ε the relative permittivity, μ the relative permeability and σ the electric conductivity. The complex wavevector is expressed in terms of the complex refraction $\tilde{n}$ index or the complex dielectric function $\tilde{\varepsilon}$ as $\tilde{k} = \tilde{n}\,\omega/c$, where $\tilde{n} = \sqrt{\tilde{\varepsilon}\mu} = n_1 + in_2$ and $\tilde{\varepsilon} = \varepsilon_1 + i\varepsilon_2$ [6]. Assuming the propagation is along the z direction, the electric field has a the form

$$\mathbf{E}(z,t) = \mathbf{E}_0 e^{i(\tilde{k}z - \omega t)} \tag{2.7}$$

where $\mathbf{E}_0$ is the electric field at $z = 0$. This form shows that the real part of the wavevector is associated to the propagation of light while its imaginary part is associated to the losses. In particular, it implies that the optical intensity falls off exponentially in a medium with a decay constant

$$\alpha(\omega) = 2n_2(\omega)\frac{\omega}{c} \tag{2.8}$$

which is called the absorption coefficient, and is thus directly related to the optical constants $n_1(\omega), n_2(\omega)$ of the material. The absorption coefficient can also be calculated from the ratio of the power removed from the incident beam per unit volume and time divided by the incident energy flux

$$\alpha(\omega) = \frac{\hbar\omega\Gamma(\omega)}{\Phi(\omega)} \tag{2.9}$$

where $\Gamma(\omega)$ is the number of optical transitions per unit volume and time and $\Phi(\omega)$ the incident energy flux. The latter can easily be computed from the Poynting vector as $\Phi(\omega) = n_1\omega^2/(8\pi)|A|^2$ where $\mathbf{A}$ is the vector potential in the Coulomb gauge.

6. All these complex quantities are in general frequency dependent. Moreover, the real and imaginary parts of $\tilde{n}$ and $\tilde{\varepsilon}$ are not independent as they are related by the Kramers-Kronig relations.

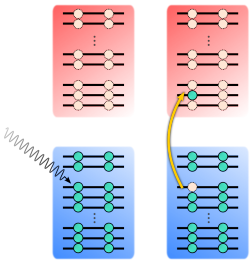

Figure 2.5: Single-particle picture of the optical absorption process.

The transition probability is a purely quantum-mechanical quantity that is related to the electronic structure of the material. Assuming the frozen-orbital approximation, the optical transition process can be described from an independent-particle picture as the promotion of an electron from an occupied to an unoccupied state upon absorption of an impinging photon, as depicted in Fig. 2.5. The system will stay for

a finite time in this excited state and eventually decays back into its ground state. This phenomenon can be described by considering that the incoming electromagnetic field is considered small enough as to be a perturbation of the Hamiltonian of the electronic system [7], the transition probability is then obtained from Fermi's golden rule

$$\Gamma(\omega) = \frac{2\pi e}{\hbar m_e} \left| \langle v|\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}|c\rangle \right|^2 \rho_{cv}(\omega), \tag{2.10}$$

where $\rho_{cv}(\omega)$ is the joint (conduction-valence) density of states, $\hat{\mathbf{A}}$ is the vector potential and $\hat{\mathbf{p}}$ is the momentum operator. The matrix elements involved in Eq. (2.10) are mostly dictated by the symmetries of the system's quantum state. In particular, they lead to the selection rules that determine the allowed transitions, and the symmetry-forbidden (allowed) transition states are called *dark (bright)* states. The joint density of states $\rho_{cv}(\omega)$ yields the number of states which feature energy difference between the conduction and valence bands equal to the photon energy, per unit volume and energy. Its critical points, called the Van Hove singularities, correspond to the electronic transitions that contribute the most to the optical properties of the system, e.g. the dielectric function.

7. The interaction of an electronic system with Hamiltonian $\hat{H}_0$ with an electromagnetic field can be accounted for by changing the momentum operator $\hat{\mathbf{p}}$ to $\hat{\mathbf{p}} + e\hat{\mathbf{A}}$, where $\mathbf{A}$ is the vector potential in the Coulomb gauge. Ignoring quadratic terms in $\mathbf{A}$, this results in a Hamiltonian
$$\hat{H} = \hat{H}_0 + \frac{e}{m_e}\hat{\mathbf{A}} \cdot \hat{\mathbf{p}}.$$

Despite the effectiveness of the aforementioned approximations in describing qualitatively the optical absorption process, the transition energies computed from an independent-particle formalism are often overestimated. An optical transition generates in fact an interacting electron-hole pair that may behave like a localized quasiparticle, known as an exciton. The exciton binding energy stabilizes the optical transition, and thus bound excitons are usually found at lower energies than the photoemission gap. Furthermore, the other electrons in the system react to the formation of the exciton and generally screen the electron-hole interaction; which has the effect of reducing the binding energy. We will see that these many-body effects can be correctly captured by an adapted two-body quasiparticle picture, such as the Bethe-Salpeter formalism, and we will specifically discuss the formation of charge-transfer excitons in doped materials in Chapter 4.

2.4 Dielectric Spectroscopy Experiments

The dielectric properties of semiconductors play a fundamental role for electronic and transport phenomena in that they influence the quasiparticle energies. As we mentioned above, the propagation of electromagnetic waves within a material is governed by the complex dielectric function $\tilde{\varepsilon}(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$, which translates the ability of the material to polarize in response to an electric field oscillating at a given frequency ω. Its real and imaginary parts quantify the propensity of a dielectric material to respectively store and dissipate the energy of a field. The dielectric function is thus responsible for the screening of the electric fields in the material, such as those generated by charged quasiparticle excitations or neutral excitons, notably reducing their binding energies.

There are several experimental techniques for determining the optical constants of a semiconductor as a function of the photon energy, mostly differing in their accuracy and frequency range. Parallel-plate experiments are arguably the simplest way to measure the dielectric function $\tilde{\varepsilon}$ in the static limit. In these experiments, the sample is inserted between two parallel electrodes to form a capacitor and an AC source is applied. The complex impedance of the system is $\widetilde{Z}(\omega) = -i\omega C_0 \tilde{\varepsilon}(\omega)$, where $C_0 = \varepsilon_0 A/d$ is the capacitance of the system without the sample, A the area of

overlap of the two electrodes and d their separation. Then, the dielectric function is obtained by measuring the impedance (in Impedance Spectroscopy Experiments) or the capacitance (in Capacitive Experiments) with an impedance analyzer or an LCR meter. This type of experiments allows for very precise measurements at low frequencies, typically below 1 GHz, and have the benefit of being relatively simple to set up.

On the other hand, optical spectroscopy experiments allow to measure the dielectric function of semiconductors in a broad energy range by irradiating the sample under study with a light source of the desired frequency. The incident photons are scattered by the system, which results from their absorption and re-emission, altering their phase, polarization and propagation direction. The scattered photons interfere and give rise to macroscopic transmitted and reflected fields, which can then be measured. The appropriate method depends on the chosen energy range of the incident light $\hbar\omega$: transmission experiments are well suited when $\hbar\omega < \Delta$, where Δ is the optical gap, while reflection experiments are better suited when $\hbar\omega > \Delta$ [8].

8. Indeed, when $\hbar\omega < \Delta$ the absorption coefficient $\alpha(\omega)$ is negligible, whereas when $\hbar\omega > \Delta$ the absorption coefficient is finite and the transmitted intensity decreases exponentially with the sample's thickness, thus making it difficult to detect.

A very popular dielectric spectroscopy technique is Spectroscopic Ellipsometry, where s- and p-polarized light waves [9] of varying frequency are irradiated onto a sample at oblique incidence [50]. The name derives from the fact that when the incident beam is linearly polarized, the reflected beam is elliptically polarized. Upon light reflection or transmission, s and p polarizations undergo different changes in amplitude and phase. Ellipsometry then measures simultaneously the amplitude ratio and phase difference between the two polarizations. In the case of reflective Ellispometric experiments, depicted in Fig. 2.6, the angle of incidence is typically set to the Brewster angle, as it maximizes the difference between s and p reflectivities, and thus the sensitivity of the measurement.

9. s- and p-polarized beams refer to light polarizations respectively perpendicular and parallel to the plane of incidence.

[50] Fujiwara (2007).

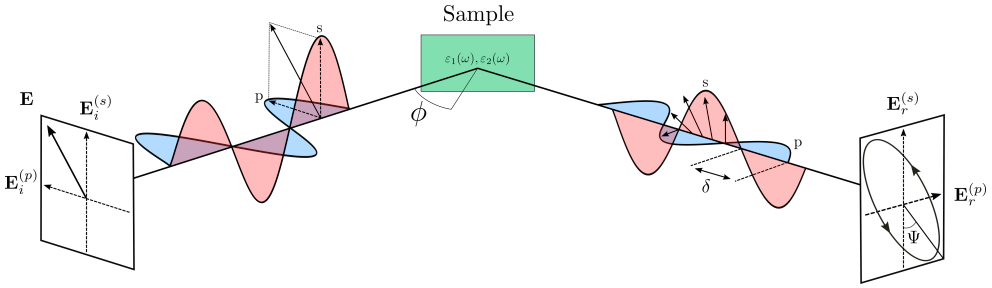

Figure 2.6: Measurement principle of reflection Ellipsometry (adapted from [50]). In this example, the incident light beam is linearly polarized with $\Psi = 45\,\text{deg}$ and $\delta = 180\,\text{deg}$.

The complex reflectivities are then related to the complex refractive index $\tilde{n}$ and the angle of incidence ϕ by the Fresnel formulae

$$r_s = \frac{\mathbf{E}_r^{(s)}}{\mathbf{E}_i^{(s)}} = \frac{\cos\phi - \sqrt{\tilde{n}^2 - \sin^2\phi}}{\cos\phi + \sqrt{\tilde{n}^2 - \sin^2\phi}} \tag{2.11}$$

$$r_p = \frac{\mathbf{E}_r^{(p)}}{\mathbf{E}_i^{(p)}} = \frac{\tilde{n}^2\cos\phi - \sqrt{\tilde{n}^2 - \sin^2\phi}}{\tilde{n}^2\cos\phi + \sqrt{\tilde{n}^2 - \sin^2\phi}} \,. \tag{2.12}$$

The complex ratio of reflectivities $\tilde{\rho}$ is expressed in terms of its amplitude $\tan\Psi$ and phase δ as

$$\tilde{\rho} = \frac{r_p}{r_s} = \tan\Psi e^{i\delta} \,. \tag{2.13}$$

The angles Ψ and δ are the quantities that are actually measured in Ellipsometry experiments. The complex dielectric function can then be determined from $\tilde{\rho}$ and ϕ

using the expression

$$\tilde{\varepsilon}(\omega) = \sin^2\phi \left[1 + \tan^2\phi \left(\frac{1-\tilde{\rho}}{1+\tilde{\rho}}\right)^2\right] . \tag{2.14}$$

Typical intensity measurements require the measurement of the real part of the dielectric function $\varepsilon_1(\omega)$ over a wide range of frequencies in order to recover its imaginary part $\varepsilon_2(\omega)$ from the Kramers-Kronig relations. In contrast, Spectroscopic Ellipsometry experiments are much more sensitive and accurate since they measure simultaneously only the change in modulus and phase of the polarization components of the light at each wavelength, allowing to withdraw the full complex dielectric function $\tilde{\varepsilon}(\omega)$ [10]. Spectroscopic Ellipsometry is typically applied in the infrared, visible and ultraviolet spectral region, allowing to track the evolution of the dielectric function over a broad energy range characterized by the occurrence of electronic excitations of different nature.

10. Moreover, Spectroscopic Ellipsometry can also be used to investigate many other quantities such as film thickness or doping concentration, making it a highly accurate and versatile tool for the optical characterization of thin films.

We mentioned earlier the facilitating role of the dielectric constant for the release of free carriers upon doping in the context of the Hydrogenic model. Thus, characterizing the dielectric properties of doped semiconductors will thus be key in the forthcoming study. We will discuss the relationship between doping and the dielectric function in Organic Semiconductors in Chapter 5, considering in particular the low-frequency limit of $\varepsilon_1(\omega)$, namely in the plateau value attained at frequencies much smaller than the lowest-energy transition.

Chapter 3

Theoretical Background

3 Theoretical Background: From Mean-Field to Embedded Many-Body Perturbation Theory

La Loi éternelle qui règle la chute des pommes a supplanté la Loi de l'Éternel, qui, pour une pomme, fit chuter Adam.

Edgar Morin

Contents

SINCE THE DISCOVERY OF THE ELECTRON in 1897 and the subsequent formalization of modern physics, the theoretical modeling of materials has never stopped intriguing physicists. The fundamental basis for understanding the microscopic phenomena at play in solids ultimately relies upon capturing their electronic structure, which means that one must deal with the interacting many-electron problem in diverse, realistic situations. The lack of exact solutions even for the few-body problem [1] is reminiscent of the fact that approximated theories are needed in order to properly describe large interacting systems.

1. As R. D. Mattuck beautifully pointed out [51], in Newtonian Mechanics the three-planets problem lacked an exact solution. With the advent of General Relativity and Quantum Electrodynamics, the two-body and one-body problems became insoluble. And within modern Quantum Field Theory, the zero-body problem (i.e. the vacuum) is irresolvable!

However, it does not suffice to devise any sort of tractable approximation of a theory to succeed in this task. The role of the theoretical physicist is to understand the limits of such approximations and to design appropriate formalisms to the description of the problem at hand. The requirement of adaptability of a given framework to a given problem is not decidable *a priori*, one must ultimately confront its theory to the experimental observations and their interpretation in order to conclude on its predictive value.

This chapter aims at giving a wide introduction to the theoretical frameworks that will be considered throughout this Thesis. The Many-Body problem will be introduced in the context of Solid State Physics, showing the necessity of approximate formalisms. Following is a formal introduction to Density-Functional Theory (DFT) as a mean-field theory for the ground state, after which the more advanced Many-Body Perturbation Theory (MBPT) frameworks of the *GW* approximation and the Bethe-Salpeter Equation (BSE) are discussed to tackle charged and neutral excited states. A change in direction is taken in order to address the classical framework of Micro-Electrostatics (ME), needed to describe many-body systems at a larger, coarser scale. The final sections are devoted to the description of the hybrid Quantum/Classical (QM/MM) frameworks combining these different levels of theory for the accurate and computationally efficient characterization of the energetics of solid-state systems via Embedded Many-Body Perturbation Theory (EMBPT).

3.1 THE MANY-BODY PROBLEM

3.1.1 Electronic structure: introduction and notation

Within Quantum Mechanics, the state of an isolated N-atom system at some fixed time can be described by the wavefunction $\Psi(\underline{\mathbf{r}}, \underline{\sigma})$ that depends on the positions $\underline{\mathbf{r}} = \mathbf{r}_1, \ldots, \mathbf{r}_N$ and spins $\underline{\sigma} = \sigma_1, \ldots, \sigma_N$ of its constituents. The wavefunction is interpreted as the complex-valued probability amplitude of finding the N particles with spins $\underline{\sigma}$ and positions $\underline{\mathbf{r}}$.

Every measurable physical quantity is represented by an observable $\hat{O}$, a linear self-adjoint operator acting on the Hilbert space, the eigenvectors of which form a basis of the Hilbert space. The result of measuring such physical quantity is governed by Born's rule: the outcome λ must pertain to the spectrum of $\hat{O}$ and it arises with probability $\langle\Psi|\hat{P}_\lambda \hat{O} \hat{P}_\lambda|\Psi\rangle$, where $\hat{P}_\lambda$ is the projection operator onto the eigensubspace associated to λ.

The simplest example is certainly the total energy E, which is represented by the Hamiltonian operator $\hat{H}$. The time evolution of the wavefunction is related to the Hamiltonian by the Schrödinger equation

$$\hat{H}(t)\,|\psi(t)\rangle = i\hbar \frac{\partial}{\partial t}\,|\psi(t)\rangle \qquad \text{(Time-dependent)} \tag{3.1a}$$

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle \qquad \text{(Stationary).} \tag{3.1b}$$

Eqs. (3.1) are 3N-dimensional linear differential equations that must be solved for $|\Psi\rangle$. The stationary equation reduces to an eigenvalue problem whereby the wavefunction is an eigenstate of the Hamiltonian operator, the eigenvalues of which are the allowed total energies.

If one is only interested in the ground-state of the system at hand, the variational principle states that the total energy can be obtained by minimizing the expectation value of the Hamiltonian over all acceptable (i.e. normalized and properly symmetrized) many-body wavefunctions, namely

$$E_0 = \min_{\Psi} \langle\Psi|\hat{H}|\Psi\rangle\,. \tag{3.2}$$

This expectation value is a correlated 3N-dimensional integral and is prohibitively expensive to directly evaluate numerically [2]. Of course, a better sampling of the very high-dimensional space of wavefunctions would drastically reduce the computation times, but we still need to develop approximations such as mean-field or many-body perturbation theory to treat this problem for large systems.

From now on, we shall use the so-called Hartree natural units whereby $e, m_e, \hbar, 4\pi\varepsilon_0 = 1$. In these units, energies are measured in Hartrees ($\mathrm{Ha} = e^2/a_0$), distances in Bohrs $a_0 = 4\pi\varepsilon_0\hbar^2/(m_e e^2)$, masses in electron masses and charges in electron charges. For the sake of simplicity, we will also disregard spin variables $\underline{\sigma}$ unless explicitly mentioned; this amounts to imply a spin summation along with any spatial integral thereafter.

2. As a simple example, consider a H_2O molecule which has $N = 10$ electrons. If we decide to compute the above expectation value on a very coarse $10 \times 10 \times 10$ real-space grid, we end up having to sum 10^{3N} terms. Modern computers perform up to 10^{15} floating point operations per second (1 petaflops), meaning one such calculation would require more than 30 million years!

3.1.2 The Many-Body Hamiltonian

Let us now consider the problem of characterizing the properties of a system composed of N electrons and nuclei. The non-relativistic, stationary Hamiltonian of such a system is composed by the kinetic energy of electrons (nuclei), the mutual Coulomb interactions between electrons (nuclei), and the Coulomb attraction between electrons and nuclei. It can be expressed in the position operators basis of the Hilbert space, whereby the Hamiltonian depends explicitly on the atomic coordinates:

$$\hat{H} = -\sum_I \frac{\nabla_I^2}{2m_I} - \sum_i \frac{\nabla_i^2}{2} + \sum_{I<J} \frac{z_I z_J}{|\mathbf{r}_I - \mathbf{r}_J|} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{iI} \frac{z_I}{|\mathbf{r}_i - \mathbf{r}_I|}\,, \tag{3.3}$$

where indices i, j (I, J) run over electrons (nuclei), m_I is the nuclei mass and z_I the atomic number of nucleus I. The attractive Coulomb interaction between electrons and nuclei is responsible for the cohesion of the system, while electron-electron and nuclei-nuclei interactions are responsible for all the correlations in the many-body state of the system.

In principle, the Schrödinger equation for this Hamiltonian should be solved in order to obtain the wavefunction $|\psi\rangle$ from which all the properties of interest can be computed. This would amount to diagonalizing the Hamiltonian operator $\hat{H}$, thus obtaining a wavefunction that cannot be split into an electronic and nuclear part.

3.1.3 Born-Oppenheimer approximation

[52] Born and Oppenheimer (1927).

Because the electron mass $m_e \simeq 0.511$ MeV/c^2 is about 1836 times lower than the mass of nuclei $m_p \simeq m_n \simeq 938.3$ MeV/c^2, electrons will in general have faster dynamics than nuclei. In 1927, Born and Oppenheimer [52] formulated the eponymous approximation that allowed to decouple the electrons and nuclei equations of motion.

To a first approximation, the small value of the ratio $\frac{m_e}{m_I}$ allows to neglect the nuclear kinetic energy T_n in the Hamiltonian of Eq. (3.3), an assumption that is generally confirmed by experimental studies. In this case the Hamiltonian no longer depends on derivatives of the nuclei positions. The nuclei Coulomb repulsion energy V_{nn} is then independent of electronic coordinates and becomes a classical variable on which the electronic states depend parametrically. Within the Born-Oppenheimer approximation, electronic transitions due to a variation in the nuclei potential are also neglected. This means that electrons are assumed to respond to atomic vibrations adiabatically: they react almost instantaneously while remaining at equilibrium in their initial state. Conversely, nuclei are assumed to feel only a time-averaged electronic density, and crystallographic experiments show that indeed nuclei make up a static spatial arrangement with only small oscillations about their equilibrium positions at room temperature.

The electronic Born-Oppenheimer Hamiltonian thus reads

$$\boxed{\hat{H}_e = \sum_i -\frac{\nabla_i^2}{2} + \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}.} \tag{3.4}$$

where the nuclei positions are now classical parameters that are often taken from X-ray diffraction experiments or computed self-consistently upon structural optimization, where the total electronic and nuclear energy is minimized with respect to the atomic

coordinates. We will consider this Hamiltonian as the starting point of all subsequent derivations.

To illustrate the validity of the Born-Oppenheimer approximation, we note that the typical frequencies of Carbon-Carbon stretching modes in organic molecules are of the order of 1500 $cm^{-1} \sim 5.10^{13}$ Hz. On the other hand, the typical band gap of organic semiconductors is 4 eV, which translates to a frequency associated with electronic motion of about 10^{15} Hz [53] [3]. While the mass criterion is always fulfilled for sufficiently heavy elements, the former frequency criterion seems however to be violated for metals. Nevertheless, experimental investigations showed that the adiabatic approximation is still mostly valid in such systems. This is because the non-adiabaticity of metals is governed by the ratio of the characteristic phonon frequency and the Fermi energy. For a wide range of temperatures, the thermal energy $k_B T$ is lower than the Fermi energy E_F, and therefore electronic excitations are confined to a narrow region around the Fermi surface, where most of their important physical properties come from, and thus non-adiabatic contributions to the motions of electrons in metals are often negligible.

[53] Coropceanu et al. (2007).

3. However the vibration modes frequency, the electron-vibration coupling strength and the electronic band dispersion are close in energy in organic systems. The Born-Oppenheimer approximation should thus be taken with care for the evaluation of transport properties, however the latter will not be discussed in this Thesis.

3.2 Density-Functional Theory

All DFT calculations performed in this Thesis use the ORCA quantum chemistry package [54]. It is a general-purpose *ab initio* program suite that uses Gaussian basis sets and features the possibility to be interfaced with external programs e.g. for hybrid quantum-classical calculations.

The solution of the many-body problem in condensed matter physics requires the development of approximated theories that are able to predict and interpret the results of observed phenomena. The key idea of Density-Functional Theory (DFT) is that the knowledge of the electron density, a much simpler object than the wavefunction, is sufficient to quantify any ground-state observable. DFT is a formally exact theory for the ground-state energy and density of a many-body system, although approximations must be performed when expressing the total energy as a functional of the electron density, an expression that lacks an analytic form. Despite that, DFT has known an impressive success in predicting material properties at a moderate computational cost, and still ranks nowadays among the favorite and most powerful theories for the description of many-body systems [55, 56].

[55] Bechstedt (2016).

[56] Martin (2020).

3.2.1 Density matrices, Exchange & Correlation

We begin our discussion by outlining some important properties of the electron density, and in particular the sum-rules it must obey.

A. Electronic density and pair density

The electron density is defined as the integral of the wavefunction over $(N-1)$ spatial coordinates

$$n(\mathbf{r}) = N \int d\mathbf{r}_2 \dots d\mathbf{r}_N \, |\Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2. \tag{3.5}$$

It follows that the electron density obeys the sum rule

$$\int d\mathbf{r}\, n(\mathbf{r}) = N. \tag{3.6}$$

Hence, when normalized, the electron density is to be interpreted as the probability of finding an electron at $\mathbf{r}$. The corresponding density operator $\hat{n}$ is an example of a one-body operator: it can be expressed as the single sum of operators acting on one particle at a time e.g. $\hat{n}(\mathbf{r}) = \sum_{i=0}^{N} \delta(\mathbf{r} - \mathbf{r}_i)$. For any such operator $\hat{O}_1 = \sum_{i=0}^{N} \hat{O}(\mathbf{r}_i)$, the expectation value is

$$\langle \Psi | \hat{O}_1 | \Psi \rangle = \sum_{i=0}^{N} \int d\underline{\mathbf{r}}\, \hat{O}(\mathbf{r}_i) |\Psi(\underline{\mathbf{r}})|^2 = \int d\mathbf{r}\, n(\mathbf{r}) O_1(\mathbf{r}), \tag{3.7}$$

where the variables $\mathbf{r}_i$ have been renamed thanks to the particles' indistinguishability. Thus, any one-body operator, such as the ionic potential or the kinetic energy, acts solely on the electron density and the action of any one-body Hamiltonian is fully characterized by it.

The pair electron density is defined analogously to the electron density as the integral of the wavefunction over $(N-2)$ spatial coordinates

$$m(\mathbf{r}, \mathbf{r}') = \frac{N(N-1)}{2} \int d\mathbf{r}_3 \dots d\mathbf{r}_N \, \left|\Psi(\mathbf{r}, \mathbf{r}', \mathbf{r}_3, \dots, \mathbf{r}_N)\right|^2. \tag{3.8}$$

It follows that the pair density obeys the sum rule

$$\int d\mathbf{r}\, d\mathbf{r}'\, m(\mathbf{r}, \mathbf{r}') = \frac{N(N-1)}{2}. \tag{3.9}$$

Hence, when normalized, the pair electron density represents the probability of simultaneously finding an electron at $\mathbf{r}$ and another at $\mathbf{r}'$. The corresponding pair

density operator $\hat{m}(\mathbf{r},\mathbf{r}')$ is an example of a two-body operator: it can be expressed as the double sum of operators acting on two particles at a time e.g. $\hat{m}(\mathbf{r},\mathbf{r}') = \sum_{i<j}\delta(\mathbf{r}-\mathbf{r}_i)\delta(\mathbf{r}'-\mathbf{r}_j)$. For any such operator $\hat{O}_2 = \sum_{i<j} O(\mathbf{r}_i,\mathbf{r}_j)$, the expectation value is

$$\langle\Psi|\hat{O}_2|\Psi\rangle = \sum_{i<j}\int d\underline{\mathbf{r}}\,\hat{O}(\mathbf{r}_i,\mathbf{r}_j)|\Psi(\underline{\mathbf{r}})|^2 = \int d\mathbf{r}\,d\mathbf{r}'\,O(\mathbf{r},\mathbf{r}')m(\mathbf{r},\mathbf{r}'). \tag{3.10}$$

Hence, any two-body operator acts solely on the pair electron density. It is worth noting that these results have a strong significance for our purpose: because the electronic Hamiltonian is formed by one-body and two-body operators, there is in principle no need to know the full many-body wavefunction $|\Psi\rangle$: the electron density and pair density are sufficient to compute the expectation values of any observable of the system.

B. *The exchange-correlation hole*

An obvious example of a two-body operator is the electron-electron interaction energy E_{ee}, which we can readily express as

$$E_{ee} = \int d\mathbf{r}\,d\mathbf{r}'\,\frac{m(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.11}$$

The pair electron density can be partitioned into a purely classical Hartree term [4] $m_H(\mathbf{r},\mathbf{r}') = n(\mathbf{r})n(\mathbf{r}')$ and exchange-correlation term

$$m(\mathbf{r},\mathbf{r}') = n(\mathbf{r})n(\mathbf{r}') + n(\mathbf{r})n_{xc}(\mathbf{r},\mathbf{r}')\ , \tag{3.12}$$

where the exchange-correlation density $n_{xc}(\mathbf{r},\mathbf{r}')$ is defined as all the information that is missing in the Hartree density for the quantum mechanical description of electrons [5]. The propagation of this decomposition to the electron-electron interaction gives

$$E_{ee} = E_H + E_{xc} \tag{3.13}$$

where the Hartree and the exchange-correlation energies are defined as

$$E_H = \frac{1}{2}\int d\mathbf{r}\,d\mathbf{r}'\,\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{3.14a}$$

$$E_{xc} = \frac{1}{2}\int d\mathbf{r}\,d\mathbf{r}'\,\frac{n(\mathbf{r})n_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.14b}$$

It is possible to obtain a sum-rule for the exchange-correlation density by straightforwardly combining the sum rules for the density in Eq. (3.6) and the pair density in Eq. (3.9), resulting in

$$\int d\mathbf{r}'\,n_{xc}(\mathbf{r},\mathbf{r}') = -1. \tag{3.15}$$

This means that because of exchange and correlation effects arising from the Pauli principle and the Coulomb correlations, the electron density at $\mathbf{r}$ dynamically generates a depletion hole around an individual electron: this is called the exchange-correlation (xc) hole. The exchange-correlation energy is then simply the Coulomb energy between the electron and its associated exchange-correlation hole within a factor $\frac{1}{2}$ which translates the fact that the hole grows with the charge. The electron thus dressed thereby forms a neutral quasiparticle that is weakly interacting with its surroundings - a description which may be of use in the construction of effective independent particle theories.

4. In a purely classical description, electrons are distinguishable and thus not subject to the Spin-Statistics theorem nor quantum mechanical correlations. This results in a factored pair density matrix in terms of individual uncorrelated densities that enters in the expression of the Hartree energy to form the so-called *direct* term of the Coulomb interaction.

5. Another way to look at Eq. (3.12) is to factorize the Hartree pair density and obtain

$$m(\mathbf{r},\mathbf{r}') = m_H(\mathbf{r},\mathbf{r}')(1+g(\mathbf{r},\mathbf{r}'))$$

where $g(\mathbf{r},\mathbf{r}')$ is the pair correlation function, which measures the amount of quantum correlations (including exchange).

The exchange and correlation contributions [6] can be further separated as

$$n_{xc}(\mathbf{r},\mathbf{r}') = n_x(\mathbf{r},\mathbf{r}') + n_c(\mathbf{r},\mathbf{r}'). \tag{3.16}$$

with the corresponding energies

$$\begin{aligned} E_x &= \frac{1}{2}\int d\mathbf{r}\, d\mathbf{r}'\, \frac{n(\mathbf{r})n_x(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ E_c &= \frac{1}{2}\int d\mathbf{r}\, d\mathbf{r}'\, \frac{n(\mathbf{r})n_c(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \end{aligned} \tag{3.17}$$

The exchange hole is rigorously defined from its Hartree-Fock (HF) expression [7], which swiftly provides us with the individual sum rules of the exchange and correlation densities

$$\int d\mathbf{r}'\, n_x(\mathbf{r},\mathbf{r}') = -1 \tag{3.18a}$$

$$\int d\mathbf{r}'\, n_c(\mathbf{r},\mathbf{r}') = \quad 0. \tag{3.18b}$$

Consequently, the interaction of the electron density with the xc-hole gives rise to a negative contribution to the total energy. The sum rules that we derived above will constitute important guidelines for the development of approximated functional forms of the exchange-correlation energy, as their violation resulted historically into poorly accurate results.

7. The exchange hole in Hartree-Fock is given by $n_x(\mathbf{r},\mathbf{r}') = -\frac{1}{n(\mathbf{r})}\left(\sum_{i=1}^{N} |\phi_i^*(\mathbf{r})\phi_i(\mathbf{r}')|\right)^2$, where ϕ_i are the HF one-body spin-orbitals.

3.2.2 Hohenberg and Kohn theorems

In 1964, Hohenbeg and Kohn [57] derived an exact theory for the ground state of many-body systems that went beyond the attempts to improve approximate theories such as the Thomas-Fermi-Dirac one, and paved the way to the development of modern Density-Functional Theory.

[57] Hohenberg and Kohn (1964).

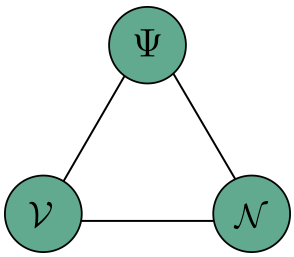

Figure 3.1: Illustration of the Hohenberg and Kohn theorem as an isomorphism between groups.

Let us first observe that the electronic Hamiltonian of Eq. (3.4) consists of two parts: a universal part composed by the kinetic energy $\hat{T}$ and the electron-electron interaction $\hat{V}_{ee}$, and a system-specific part formed by the ionic potential $\hat{V}$ which varies greatly with different atomic configurations. But the set of all physical (non-diverging) ionic potentials is defined modulo constant energy shifts representing arbitrary choices of the zero of energy, which do not change the electronic spectrum. We can therefore think of this set as a quotient group $\mathcal{V}$. This group is somehow connected to the quotient group $\mathcal{P}$ of ground-state wavefunctions modulo their phase, and the latter is again connected to the set of ground-state densities $\mathcal{N}$. We already know that the ionic potential acts solely on the charge density. The essence of the Hohenberg and Kohn theorem is to show that the universal kinetic energy and electron-electron interaction operators are also functionals of the charge density for the system in its ground-state.

Theorem 1: The external potential is a unique functional of the density.

Proof: Assume that there exists two potentials, V_1 and V_2 differing by more than a constant and giving rise to the same ground state density $n(\mathbf{r})$. They belong to distinct Hamiltonians $\hat{H}_1$ and $\hat{H}_2$ associated with different ground state wavefunctions Ψ_1 and Ψ_2 that are assumed here to be non-degenerate. Since the ground state densities

are the same for $\hat{H}_1$ and $\hat{H}_2$, we have

$$\langle\Psi_1|\hat{H}_1|\Psi_1\rangle < \langle\Psi_2|\hat{H}_1|\Psi_2\rangle = \langle\Psi_2|\hat{H}_2|\Psi_2\rangle + \int d\mathbf{r}\, n(\mathbf{r})\Big(V_1(\mathbf{r}) - V_2(\mathbf{r})\Big) \tag{3.19a}$$

$$\langle\Psi_2|\hat{H}_2|\Psi_2\rangle < \langle\Psi_1|\hat{H}_2|\Psi_1\rangle = \langle\Psi_1|\hat{H}_1|\Psi_1\rangle + \int d\mathbf{r}\, n(\mathbf{r})\Big(V_2(\mathbf{r}) - V_1(\mathbf{r})\Big) \tag{3.19b}$$

Adding these two inequalities results in a contradiction $E_1 + E_2 < E_2 + E_1$, which proves the theorem [8].

8. Note that the strict inequality holds only for non-degenerate ground states, but the demonstration can be generalized to any ground state.

Theorem 2: A universal functional for the energy $E[n]$ can be defined in terms of the density. The exact ground state is the global minimum value of this functional.

Proof: Since the ground state wavefunction is a unique functional of the density $\Psi = \Psi[n]$, if follows that any observable can be expressed as a functional of the density $O[n] = \langle\Psi[n]|\hat{O}|\Psi[n]\rangle$. In particular, the total energy is

$$E = E[n] = \langle\Psi[n]|\hat{H}|\Psi[n]\rangle = \langle\Psi[n]|\hat{T} + \hat{V}_{ee} + \hat{V}_{ion}|\Psi[n]\rangle = \hat{T}[n] + \hat{V}_{ee}[n] + \hat{V}_{ion}[n]. \tag{3.20}$$

Because the ground state density is unique, the problem of finding the ground state of a many-body system may be cast in a variational form

$$E_0 = \inf_n E[n] = \inf_n \left(F[n] + \int d\mathbf{r}\, V_{ion}(\mathbf{r})n(\mathbf{r})\right) \tag{3.21}$$

where $F[n] = T[n] + V_{ee}[n]$ is the universal Hohenberg and Kohn functional. This completes the proof.

3.2.3 Kohn-Sham DFT

A. *Kohn-Sham equations*

Thanks to the Hohenberg and Kohn theorems, the variational principle for the true interacting problem can be cast in terms of a minimization of $E[n]$ with the minimal constraint that the charge density sums up to the total number of electrons, i.e. the sum rule of Eq. (3.6) is satisfied. The resulting Lagrangian is

$$\Omega[n, \mu] = E[n] + \mu\left(N - \int d\mathbf{r}\, n(\mathbf{r})\right) \tag{3.22}$$

where the chemical potential μ plays the role of a Lagrange multiplier. The stationary condition $\frac{\delta\Omega[n,\mu]}{\delta n(\mathbf{r})} = 0$ yields

$$\frac{\delta F[n]}{\delta n(\mathbf{r})} + V_{ion}(\mathbf{r}) = \mu\,, \tag{3.23}$$

which is a 3D differential equation that is indeed much simpler than the full many-body Hamiltonian of Eq. (3.4). However, in order to solve this equation one needs to know the universal functional $F[n]$, for which an exact formulation is lacking.

[58] Kohn and Sham (1965).

In 1965, Kohn and Sham [58] introduced their celebrated approach to address the unknown functional form of $F[n]$ by mapping the true many-body problem into an effective non-interacting one which has the same ground state energy and electron density.

Indeed, a non-interacting system is fully determined by a set of one-body orbitals $|\phi_i\rangle$, for which the density and the kinetic energy T_0 take simple expressions [9]. Because the true many-body system and the fictitious non-interacting system must have the same ground-state density, they should obey to the same variational Euler-Lagrange equation and have the same chemical potential, namely

$$\frac{\delta T_0}{\delta n(\mathbf{r})} + V_{\text{eff}}(\mathbf{r}) = \frac{\delta F[n]}{\delta n(\mathbf{r})} + V_{\text{ion}}(\mathbf{r}) \,, \tag{3.24}$$

which serves as the definition of the effective potential V_{eff}. Because the functional form of the Hartree energy is known from Eq. (3.14a), we introduce the Hartree interaction potential

$$V_H(\mathbf{r}) = \frac{\delta E_H[n]}{\delta n(\mathbf{r})} = \int d\mathbf{r}' \, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.25}$$

It is then possible to write the effective potential as

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ion}}(\mathbf{r}) + V_H(\mathbf{r}) + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \tag{3.26}$$

where all the remaining functionals are pushed into the so-called *exchange-correlation* functional

$$E_{xc}[n] = T[n] - T_0[n] + E_{ee}[n] - E_H[n] \tag{3.27}$$

which is now the functional that has still to be determined [10]. The well-known Kohn-Sham equations (Eq. (3.30)) are then obtained from the Schrödinger equation of non-interacting electrons under the effective potential V_{eff}.

10. We observe that within DFT, the exchange-correlation energy contains not only the deviation of the electron-electron interaction from the Hartree potential $E_{ee}[n] - E_H[n]$ but also corrections to the kinetic energy $T[n] - T_0[n]$.

Another way of finding the Kohn-Sham equations is by minimizing the total energy under the strengthened constraint that all orbitals $\underline{\Phi} = \{\phi_i(\mathbf{r})\}$ are kept orthonormalized throughout the minimization process. This results in

$$\Omega(\underline{\Phi}, \underline{\Lambda}) = E[\underline{\Phi}] + \sum_{i \leqslant j} \Lambda_{ij} \Big(\delta_{ij} - \langle \phi_i | \phi_j \rangle \Big). \tag{3.28}$$

The Kohn-Sham equations are then obtained by minimizing the Lagrangian $\Omega(\underline{\Phi}, \underline{\Lambda})$ with respect to the orbitals ϕ_i^* in order to obtain the orbitals that correspond to the ground-state energy and density [11].

11. One can prove that differentiating with respect to ϕ_i yields the same result.

$$\frac{\delta \Omega(\underline{\Phi}, \underline{\Lambda})}{\delta \phi_i^*(\mathbf{r})} = \frac{\delta T[n]}{\delta \phi_i^*(\mathbf{r})} + \left[\frac{\delta E_{\text{ion}}[n]}{\delta n(\mathbf{r})} + \frac{\delta E_H[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \right] \frac{\delta n(\mathbf{r})}{\delta \phi_i^*(\mathbf{r})} - \sum_j \Lambda_{ij} \phi_j(\mathbf{r}) = 0. \tag{3.29}$$

A unitary rotation leaves the total energy unchanged and diagonalizes $\underline{\Lambda}$, whose eigenvalues are the Kohn-Sham energies ϵ_i, and one is left with the Kohn-Sham equations for the effective non-interacting system

$$\boxed{\left[-\frac{\nabla^2}{2} + V_{\text{eff}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}).} \tag{3.30}$$

Eq. (3.30) is a system of coupled equations that allows reconstructing the exact charge density and total energy from the single-particle orbitals ϕ_i and energies ϵ_i. As such, they achieve a major simplification in that they feature a much more attainable scaling with system size than the factorial scaling of the many-body problem, the only approximations made being those concerning the functional expression of E_{xc}. However, the interpretation of the Kohn-Sham orbitals and eigenvalues is not straightforward, and one must be careful in deriving excited properties from this formalism.

The peculiar nature of the Kohn-Sham equations relies on the fact that V_{eff} depends explicitly on the density and therefore implicitly on the Kohn-Sham orbitals ϕ_i. As a result, any change in the orbitals also affects the effective potential on which they depend: the Kohn-Sham equations thus need to be solved simultaneously by a Self-Consistent Field (SCF) procedure. The latter typically consists in forming an initial guess for the orbitals or the density, e.g. a superposition of atomic densities, and then propagating it through Eq. (3.30) until the density and the total energy are converged.

B. Hellmann-Feynmann theorem

The knowledge of the ground-state energy and density further allows the calculation of energy gradients with respect to the ionic positions by means of the Hellmann-Feynman theorem. The latter relates the derivative of the total energy with respect to any external and continuous parameter λ to the derivative of the Hamiltonian with respect to this same parameter

$$\frac{\partial E}{\partial \lambda} = \left\langle \psi_\lambda \middle| \frac{\partial \hat{H}_\lambda}{\partial \lambda} \middle| \psi_\lambda \right\rangle, \tag{3.31}$$

where the derivative of the wavefunction cancels by normalization since it involves the expression $d/d\lambda \, \langle \psi | \psi \rangle$. This allows for a very efficient calculation of the forces acting on the nuclei and thus leads to the possibility of minimizing the energy of the system with respect to ionic positions, namely relaxing the structure before studying its properties. Further quantities, such as vibrational modes or phonons, involving second-order derivatives (the dynamical matrix) are also dramatically simplified, but we will not address them here.

[59] Thomas (1927).

[60] Fermi (1927).

[61] Dirac (1930).

3.2.4 Exchange-Correlation Functionals

A. Local Density Approximation

The search for approximated expressions of the exchange-correlation functional is a central ongoing challenge for DFT and this is what makes it no longer an exact theory. Proposed in 1964 by Hohenberg and Kohn [57], the Local Density Approximation (LDA) is the first successful step taken towards this aim: the exchange-correlation functional is assumed to depend solely upon the value of the electronic density at each point i.e.

$$E_{xc}[n] = \int d\mathbf{r}\, n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})). \tag{3.32}$$

In other words the average exchange-correlation energy per electron ϵ_{xc} is assumed to be a function of $n(\mathbf{r})$ instead of a functional. While there are many approaches that can yield local approximations to the exchange-correlation energy, the most successful ones are those that are reminiscent of the Thomas-Fermi-Dirac theory [59, 60, 61] of the Homogeneous Electron Gas (HEG), also known as the *Jellium* [12], whereby ϵ_{xc} is just the exchange-correlation energy density of a homogeneous electron gas at the same density $n(\mathbf{r})$. In practice, LDA is nowadays synonymous with exchange-correlation functionals based on the HEG.

A decisive step was made by Ceperley and Alder [62] who performed Quantum Monte Carlo (QMC) calculations of the total energy of the interacting HEG at various densities. Because in the HEG the Hartree energy is exactly canceled with the ionic background, the exact correlation energy could be calculated by subtracting to the QMC calculations the non-interacting kinetic energy T_0 and the known exchange

12. The *Jellium* considers N interacting electrons in a solid of finite volume Ω compensated by a homogeneous positively charged background such that the system is globally neutral. Thomas and Fermi independently solved in 1927 the interaction-free HEG, obtaining the expression of the kinetic energy as a functional of the density,

$$T_0 = \frac{3}{10}\left(\frac{9\pi}{4}\right)^{\frac{2}{3}} \frac{N}{r_s^2}.$$

where $\frac{4\pi}{3} r_s^3 = \frac{1}{n}$. In 1930, Dirac obtained the exchange contribution to the total energy of the *Jellium* in the Hartree-Fock approximation, namely

$$E_x = -\frac{3}{4\pi}\left(\frac{9\pi}{4}\right)^{\frac{1}{3}} \frac{N}{r_s}.$$

The exchange energy is always negative, therefore stabilizing the electron gas at finite densities, because the antisymmetry requirement of the wavefunction allows electron of the same spin to avoid each other, minimizing the Coulomb repulsion energy.

[62] Ceperley and Alder (1980).

energy E_x. Various fits of the QMC data points are used when performing LDA calculations nowadays [63, 64, 65], and generalizations to spin-dependent functionals have been constructed, with separate densities for the spin-up and spin-down orbitals.

[63] Zunger et al. (1980).

[64] Vosko et al. (1980).

[65] Perdew and Wang (1992).

In principle, this approximation is strictly valid only in the limit of a homogeneous electron gas, which may seem a very crude way of describing solids. Despite this, LDA faced a large success in predicting structural properties such as bond lengths (see Table 3.1) and vibrational properties (see Fig. 3.2) even in solids featuring very inhomogeneous density profiles, hence widely contributing to the success of DFT.

The somehow unexpected success of LDA for inhomogeneous systems can be understood by the fact that the exchange-correlation energy of Eq. (3.14b) depends only on the spherical average of the exchange-correlation hole

$$\begin{aligned} E_{xc}[n(\mathbf{r})] &= \frac{1}{2}\int d\mathbf{r}\, d\mathbf{r}'\, n(\mathbf{r})\frac{n_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{2}\int d\mathbf{r}\, n(\mathbf{r}) \int_0^\infty dR\, R \int_{|\mathbf{r}-\mathbf{r}'|=R} \frac{n_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &= 2\pi \int d\mathbf{r}\, n(\mathbf{r}) \int_0^\infty dR\, R\, \bar{n}_{xc}(\mathbf{r},R)\,, \end{aligned} \tag{3.33}$$

where we defined $\bar{n}_{xc}(\mathbf{r},R)$ as the spherically averaged the exchange-correlation hole at a distance R around $\mathbf{r}$ with the Coulomb norm. Calculations show that LDA describes correctly the spherical average of the exchange-correlation hole, an important property that may be related to the fact that within LDA the exchange-correlation sum rule of Eq. (3.15) is satisfied i.e.

$$4\pi \int d\mathbf{r}\, n(\mathbf{r}) \int_0^\infty dR\, R^2\, \bar{n}_{xc}(\mathbf{r},R) = -1. \tag{3.34}$$

These are important physical conditions that are fulfilled by LDA and they explain its wide success particularly in the description of structural properties and charge density distributions of covalent, metallic and ionic materials, despite a more mitigated accuracy on the description of Hydrogen bonds or molecular crystals that are bonded by weak Van der Waals forces, namely non-local interactions that cannot be captured by local functionals. Likewise, vibrational modes or phonon energies are very often accurately described within LDA since relying on total energy variations (see Fig. 3.2).

	Metals	Non-metals
LDA	-0.136	-0.042
PBEsol	-0.039	0.026
PBE	0.046	0.085
TPSS	0.039	0.066

Table 3.1: Mean prediction error (Å) of lattice constants of DFT calculations using different functionals for a selection of 14 metals and 10 non-metals with respect to experimental data corrected for the zero-point anharmonic expansion contribution. Adapted from [66].

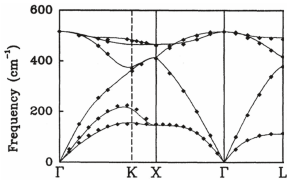

Figure 3.2: Experimental phonon dispersion in Silicon (solid line) and LDA predicted values (dots), taken from [67].

B. *Jacob's ladder to chemical heaven*

The aforementioned success of LDA is further mitigated by the fact that it tends to overbind atoms and molecules: bond lengths are slightly too small and vibrational modes energy too large. While LDA correctly describes the spherical average of the exchange-correlation hole, its value in a given direction can greatly differ from the true hole. Moreover, it does not present the correct asymptotic behavior of the exchange-correlation potential that an electron feels when it is pulled out from the system, which in finite systems must decay asymptotically like the Coulomb potential as $-1/r$. This scaling actually comes from the action of the non-local Fock operator $\hat{K}$ on the Kohn-Sham orbitals

$$\begin{aligned} \langle \mathbf{r}\,|\hat{K}|\phi_i\rangle &= -\sum_{n=1}^{N} \int d\mathbf{r}'\, \frac{\phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\phi_i(\mathbf{r}') \\ &\approx -\sum_{n=1}^{N} \frac{\phi_n(\mathbf{r})}{r} \int d\mathbf{r}'\, \phi_n^*(\mathbf{r}')\phi_i(\mathbf{r}') = -\frac{\phi_i(\mathbf{r})}{r}\,, \end{aligned} \tag{3.35}$$

where we made use of the orthogonality of the one-body orbitals and the limit $|\mathbf{r}-\mathbf{r}'| \approx r$. Anticipating on our discussion of the electronic properties, this comes as an important property advocating the use of some amount of exact exchange in the exchange-correlation functionals.

Jacob's ladder to chemical heaven represents the efforts to improve DFT beyond LDA by addressing the inhomogeneity of the system with a non-local dependence of the exchange-correlation functional on the density. Its rungs correspond to expansions of exchange-correlation energy as a functional of increasing derivatives of the density

$$E_{xc}[n] = \int d\mathbf{r}\, n(\mathbf{r})\, \varepsilon_{xc}[n(\mathbf{r}),\, \nabla n(\mathbf{r}), \nabla^2 n(\mathbf{r}), \ldots]. \tag{3.36}$$

Generalized Gradient Approximations (GGA) [68] correspond to a first correction to LDA with the integration of a dependence over the gradient of the density $\nabla n(\mathbf{r})$. These approximations do not naturally feature the correct asymptotic scaling, but current GGA functionals like PBE [68] and BLYP [69] tackle this by introducing a crucial cutoff procedure that interpolates smoothly between the desired short and long range functional forms [13]. This procedure is repeated to match other known exact relations, such as the high and low density limits or sum rules, but has the drawback of defining empirical parameters that must be fitted to experimental data.

[68] Perdew et al. (1996).

[69] Lee et al. (1988).

13. Acronyms of functionals often link to their authors e.g. P for Perdew, B for Becke, E for Ernzherof, L for Lee, Y for Yang and P for Parr.

The next rung comprises the so-called meta-GGA functionals which also include a dependence on the Laplacian of the density $\nabla^2 n(\mathbf{r})$, or the inclusion of the (non-interacting) kinetic energy density term $\tau(\mathbf{r}) = \frac{1}{2}\sum_i^{occ} |\nabla \phi_i(\mathbf{r})|^2$. The TPSS functional [70] or the Minnessota local functional M06-L [71] are common examples of meta-GGA functionals.

[70] Tao et al. (2003).

[71] Zhao and Truhlar (2006).

Hybrid functionals explicitly address the problem of recovering the asymptotic behavior of $V_{xc}(\mathbf{r})$ by incorporating a portion of exact exchange in the exchange-correlation functional [14]. Using exact exchange has the additional benefit of partially curing the self-interaction problem whereby the terms of the Hartree and exchange energies do not cancel exactly [15]. The general form is composed of a local correlation contribution, a local exchange contribution and the exact exchange potential, which depends on the set of one-body eigenstates :

$$E_{xc}^{Hybrid}[n] = \alpha E_x^{HF}[\{\phi\}] + (1-\alpha)E_x^{GGA}[n, \nabla n] + E_c^{GGA}[n, \nabla n], \tag{3.37}$$

where the amount of exact exchange is controlled by an empirical parameter α which must be tuned to a specific problem. The PBE0 hybrid functional was developed from perturbation theory and features $\alpha = 0.25$, while the B3LYP has a more involved empirical functional form

$$E_{xc}^{B3LYP} = 0.08E_x^{LDA} + 0.2E_x^{HF} + 0.72E_x^{GGA} + 0.19E_c^{LDA} + 0.81E_c^{GGA}. \tag{3.38}$$

PBE0 and B3LYP rate among the most common functionals used to describe finite systems, and we will mostly use the former throughout this Thesis, both for structural relaxations and as a starting point for subsequent *GW* calculations.

14. Thus, introducing an amount α of exact exchange results in an asymptotic behavior of the potential α/r, which can mimic the screened Coulomb potential in solids of dielectric constant $\varepsilon \sim \alpha^{-1}$.

15. An example of this problem is the Hydrogen atom, where the electron interacts with itself through the exchange-correlation functional.

Another strategy for addressing the need for having 100% of exact exchange in the long-range (at least in finite size systems) while keeping some amount of mixing in the short range is to split the Coulomb interaction kernel appearing in the exchange integral into two contributions

$$\frac{1}{r} = \frac{\alpha + \beta(f(\gamma r))}{r} + \frac{1-\alpha-\beta(1-f(\gamma r))}{r}. \tag{3.39}$$

where $f(\gamma r)$ is a separating function ranging from 0 at the origin to 1 at infinity (such as the error function) and α, β, γ are empirical parameters [16]. The first term is the long-range part and the second term is the short-range part; the parameters can be tweaked to control the desired amount of exact or DFT exchange in each part. Range-separated functionals have also been developed for describing extended systems in which long-range Coulomb interactions are screened by the macroscopic dielectric constant ε. The amount of long-range exact exchange should be qualitatively proportional to $1/\varepsilon$, which is indeed the rationale behind the range-separated hybrid HSE functional [72].

We mention another correction to DFT functionals that accounts for dispersion interactions, which are especially important in non-covalent molecular systems. Indeed, by assuming only a local or a semi-local dependence of V_{xc} on the electron density, DFT functionals typically neglect long-range electron correlations that are responsible for the ubiquitous dispersion interactions such as the London forces. The latter are accounted for in DFT by adding a dispersion correction to the total energy that is modeled as a simple and cost-effective pairwise interaction. Its functional form is obtained through the asymptotic expansion

$$E_D = \sum_{a,b} \sum_{n=6,8,10} \frac{C_{a,b,n}}{r_{ab}^n} , \tag{3.40}$$

where r_{ab} is the distance between basis points a and b and the $C_{a,b,n}$ coefficients are fitted to experimental values [17]. The semi-classical D2 and D3 corrections are the most widely used in this context [73].

The introduction of empirical parameters in the exchange-correlation functionals, and their sometimes necessary tuning to experimental data, translates the difficulty of DFT to perform as a predictive theory. In general, it is difficult to improve LDA in a universal manner that performs similarly for various system types and tasks (e.g. total energies, structural properties, ionization energies, charge densities) and that satisfy desirable properties (e.g. exchange-correlation sum rule, asymptotic behavior), which is why the development of accurate functionals is still a very active field. An interesting - though severe - discussion on the world of functionals can be found in [74].

16. By setting $\beta = 0$ one recovers conventional hybrid functionals, while range-separated functionals with $\alpha = 0$ are usually denoted by LC and the name of the semi-local functional, e.g. LC-PBE. When both $\alpha \neq 0$ and $\beta \neq 0$, they are denote by CAM and the name of the semi-local functional, e.g. CAM-B3LYP. The separating function $f(\gamma r)$ can be either the error function $\mathrm{erf}(\gamma r)$ or a Slater function $1 - \exp(-\gamma r)$, in which case the functionals are labeled with an additional Y in their name, e.g. CAMY-B3LYP and LCY-PBE. Finally, the optimal value of screening factor γ that controls the inverse effective length of the partitioning is still under discussion, but it is possible to tune it for a specific problem.

[72] Heyd et al. (2003).

17. When adopting atom-centered basis sets, the basis set superposition error (BSSE) results in a artificial interaction between fragments. The resulting spurious dispersion can be addressed by the empirical geometrical counterpoise correction (gCP).

[73] Grimme (2011).

[74] Medvedev et al. (2017).

3.2.5 Kohn-Sham band structure & Janak theorem

Similarly to the Hartree-Fock treatment, Kohn-Sham theory provides us with the single-particle energies ϵ_i that arise as Lagrange multipliers, but their identification with addition or removal energies for the calculation of band gaps is unclear. DFT is a ground-state density formalism and as such it is not designed to provide electronic energy levels. As discussed in Sec. 2.2, the fundamental gap is determined experimentally from photoemission and inverse photoemission techniques as the difference between the first Ionization Potential (IP) and the first Electron Affinity (EA) i.e.

$$\Delta = EA - IP = \big(E_0[N+1] - E_0[N]\big) - \big(E_0[N] - E_0[N-1]\big) \tag{3.41}$$

where $E_0[N]$ denotes the ground-state total energy of the system with N electrons. A first insight on the difficulty of interpreting the Kohn-Sham eigenvalues consists in remarking that their sum does not equal the total energy of the system. Indeed, the sum of the Kohn-Sham eigenvalues reads

$$\sum_{i=1}^{occ} \varepsilon_n = \sum_{i=1}^{occ} \langle \phi_i | -\frac{\nabla^2}{2} + V^{\mathrm{eff}}(\mathbf{r}) | \phi_i \rangle = T_0 + \int d\mathbf{r}\, n(\mathbf{r}) \Big(V(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \Big). \tag{3.42}$$

where the exchange-correlation potential is $V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}$. On the other hand, the ground-state total energy is

$$E_0[n] = T_0 + E_{ion}[n] + E_H[n] + E_{xc}[n] = T_0 + \int d\mathbf{r}\, n(\mathbf{r}) V(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}\, n(\mathbf{r}) V_H(\mathbf{r}) + E_{xc}[n]. \tag{3.43}$$

The difference between the total energy and the sum of the Kohn-Sham eigenvalues stems from the double counting of the Hartree energy in Eq. (3.42) and the difference between the exchange-correlation functional and $\int d\mathbf{r}\, n(\mathbf{r}) V_{xc}(\mathbf{r})$. Indeed, the total energy can be written as

$$E[n] = \sum_{i=1}^{occ} \varepsilon_n - E_H[n] + E_{xc}[n] - \int d\mathbf{r}\, n(\mathbf{r}) V_{xc}(\mathbf{r}). \tag{3.44}$$

As such, individual Kohn-Sham eigenstates cannot seamlessly be identified to differences of total energies. And indeed, the drastic mismatch between band gap predictions of HF and DFT and experimental values arises precisely from the misleading identification of eigenvalues to actual photoemission levels.

While Hartree-Fock eigenvalues can be understood with the Koopmans' theorem as a proper difference of total energies between systems with different number of electrons within the (though severe) frozen-orbital approximation, Janak's theorem [75] is an important relation that may allow to progress in the use of DFT also for the description of electronic excitations. The key idea is to generalize the Kohn-Sham approach by considering fractional occupation numbers f_i, therefore turning them into continuous variables. [18] The kinetic energy then becomes a functional of both the Kohn-Sham orbitals and the occupation numbers, and so is the total energy. The latter must then be minimized with respect to the orbitals and the occupation numbers in order to obtain the ground state of the system

$$E_0 = \min_{\Phi, f} E[\underline{\Phi}, \underline{f}], \tag{3.45}$$

and the minimization over the occupation numbers results in Janak's formula

$$\epsilon_i = \frac{\partial E}{\partial f_i}. \tag{3.46}$$

Hence, the infinitesimal variation of the total energy with respect to the occupation numbers is equal to the single-particle DFT eigenvalues. The Kohn-Sham formalism can be recovered by restricting f_i to integer values, and thus the Kohn-Sham eigenvalues are obtained by evaluating Eq. (3.46) at these values. This infinitesimal variation must be confronted to the experimental definition of the electronic energy levels as the variation of the total energy with respect to integer, and not infinitesimal, change of population.

It has been shown that the total energy must be piecewise-linear with respect to continuous occupation numbers [76, 77]. The different slopes of the linear segments reflect discontinuities in the chemical potential, that is, the fact that the Ionization Potential and the Electron Affinity are not the same. It has also been proved that the HF energy curve is concave while LDA is mostly convex [78], as depicted schematically in Fig. 3.3. This provides yet another rationale for the use of hybrid functionals in order to restore the piecewise linearity of the total energy as a function of the number of electrons between two integer values.

Even though the Kohn-Sham eigenvalues do not formally correspond to the photoemission energies measured in experiments, they remain a very valuable information

[75] Janak (1978).

18. Our discussion on DFT considered the zero temperature case in which the occupation numbers are simply $f_i = \Theta(\epsilon_i - E_F)$.

[76] Perdew et al. (1982).

[77] Perdew and Levy (1983).

[78] Li and Yang (2017).

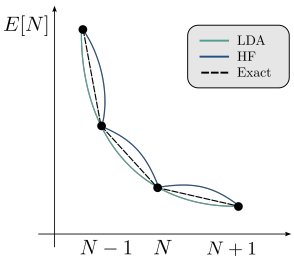

Figure 3.3: Typical shape of the exact, DFT and HF total energy curves: the exact total energy is piecewise-linear while LDA is usually convex and HF is concave. Mixing local functionals and exact exchange may restore linearity.

in that they provide an affordable starting point for the assessment of the electronic properties of many-body systems, and the search for exchange-correlation functionals providing accurate electronic energy levels is still an active research topic [79, 80]. Further, while band gaps in organic systems can be severely inaccurate (by a few eVs) if an optimally-tuned functional is not used, the dispersion of the bands and the spacing between occupied or unoccupied levels can be very satisfactory.

[79] Dabo et al. (2010).

[80] Kronik et al. (2012).

3.2.6 Spin Polarization

From the beginning of this chapter spin variables have been disregarded, meaning that the molecular orbitals and energies were assumed to be independent of the electron spin. This is known as the restricted Restricted Kohn-Sham (RKS), whereby each energy level is doubly occupied with a spin-up and spin-down electron sharing the same one-body molecular orbital. Whenever the system is closed shell and the Hamiltonian does not mix the two spin channels, the latter are decoupled and the RKS procedure can be justified.

However, there are many physical situations in which the electronic levels are not spin-degenerate, such as the Zeeman splitting in the presence of a magnetic field or when the spin-orbit coupling is not negligible e.g. in the presence of heavy atoms. Furthermore, a restricted treatment cannot describe molecules with unpaired electrons such as free radicals or anions and cations of closed-shell molecules. In such cases the orbitals and energies can be treated independently for each spin channel, resulting in the Unrestricted Kohn-Sham (UKS) approach. Unfortunately, the latter suffers from spin contamination: the ground state is not anymore an eigenfunction of the squared spin operator $\hat{\mathbf{S}}^2$. This is a desirable property since the non-relativistic Hamiltonian of DFT commutes with $\hat{\mathbf{S}}^2$. In particular, the difference between the expectation value of $\hat{\mathbf{S}}^2$ in the UKS approach and the exact one is [81]

[81] Szabo and Ostlund (2012).

$$\Delta \langle \hat{\mathbf{S}}^2 \rangle = N_\beta - \sum_{i,j}^{occ} \left| \left\langle \phi_i^\alpha \middle| \phi_i^\beta \right\rangle \right|^2. \tag{3.47}$$

This contamination is a result of an artificial mixing of different electronic spin states that would otherwise occupy the same molecular orbital. An intermediate solution is the Restricted-Open Kohn-Sham (ROKS) treatment, which considers spin-dependent occupations with spin-independent orbitals. In practice this results in having $2N_\alpha$ electrons (assuming $N_\beta > N_\alpha$) that are restricted in occupying the same molecular orbital and the remaining electrons that are unrestricted. This procedure has the advantage of providing a wavefunction that is always guaranteed to be an eigenfunction of the squared spin operator. We will be discussing such aspects in this Thesis as our study of doping will be associated with the possibility of charge-transfer states in the ground-state, leaving open-shell anions and cations difficult to treat within the RKS approach.

3.2.7 Gaussian basis sets

In practice, the equations of DFT must be solved numerically by mapping them into algebraic equations. This is implemented by expressing quantum states and operators in a finite vector space whose basis elements are called *basis functions*. For instance, the wavefunction is expressed as a vector, one-electron operators as matrices and two-electron operators as rank-four tensors. Given a finite set of basis functions $g_\mu(\mathbf{r})$, one can establish an approximate resolution of the identity by projecting a molecular

orbital $\phi_n(\mathbf{r})$ as

$$\phi_n(\mathbf{r}) = \sum_\mu C_{\mu n} g_\mu(\mathbf{r}) \tag{3.48}$$

where $C_{\mu n}$ are the expansion coefficients given by

$$C_{\mu n} = \sum_\nu \langle g_\mu | g_\nu \rangle^{-1} \langle g_\nu | \phi_n \rangle . \tag{3.49}$$

Therefore, numerical calculations only work with the basis coefficients $C_{\mu n}$, from which the molecular orbitals $\phi_n(\mathbf{r})$ can be retrieved. Because of the approximate nature of this projection, calculations must be converged with respect to the basis set size in order to approach the Complete Basis Set (CBS) limit. Such extrapolation schemes are done within DFT and other quantum chemistry methods.

The basis functions $g_\mu(\mathbf{r})$ can in principle be any function. However, the description of periodic and extended systems usually adopts plane waves $g_\mu(\mathbf{r}) = A_\mu e^{i\mathbf{k}_\mu \cdot \mathbf{r}}$, which are typically chosen within the solid state community. Conversely, the quantum chemistry community uses atomic orbitals, typically to describe finite and disordered systems. Atomic orbitals follow the usual radial-angular decomposition that is obtained from the solutions of the Hydrogen atom [19]

$$g_\mu(\mathbf{r}) = R_l(r) Y_{lm}(\theta, \phi), \tag{3.50}$$

where $Y_{lm}(\theta, \phi)$ are spherical harmonics. Slater-Type Orbitals (STOs) mimic the solutions of the Hydrogen atom, but with an additional adjustable localization parameter ζ. The radial part is a decaying exponential:

$$R_l(r) = A(l, \zeta) r^l e^{-\zeta r}. \tag{3.51}$$

While they are well motivated physically, STOs are not very popular in quantum chemistry codes because of the computational complexity of the integrals involving them. Gaussian-type orbitals (GTOs) cure this problem by approximating STOs with a linear combination of Gaussian functions, for which integrals can be expressed in closed form due to the Gaussian product theorem [20]. They feature a Gaussian radial part

$$R_l(r) = B(l, \zeta) r^l e^{-\zeta r^2}. \tag{3.52}$$

The definition of Gaussian basis sets proceeds as follows. For each element, the atomic orbitals (e.g. 1s, 2p, 2s ...) are described by z (zeta) atomic basis functions (STOs), and each basis function is decomposed into sum of g Gaussians (GTOs). Split-valence basis sets use more than one basis function for valence orbitals while keeping a single one for core orbitals, not involved in chemical bonding. The inclusion of unoccupied orbitals allows to account for the polarization of the electron density in atoms and molecules, like in the case of chemical bonding, and are called *polarization* functions. Further flexibility can be achieved by the addition of *diffuse* functions: these are extended Gaussians (with a very small exponent) that allow to describe the tail of the atomic orbitals far from nuclei. As a general principle, the needed coefficients are variationally optimized to minimize the total energy (as obtained from a chosen quantum chemistry method) of the corresponding isolated atom for a given basis size. Standard basis sets are compiled e.g. in the Basis Set Exchange website.

STO basis sets: The *STO-gG* minimal basis sets correspond to single-zeta ($z = 1$) basis function for every atomic orbital, each being composed by g Gaussian functions.

Pople basis sets [82]: These basis sets are denoted as $g_0 - g_1 \ldots g_z G$ where g_0 is

19. Some programs use instead a Cartesian decomposition

$$g(\mathbf{r}) = x^k y^n z^m R_l(\theta, \phi),$$

which is not equivalent to the spherical one, except for s- and p-type orbitals.

20. The product of two Gaussian functions

$$\sqrt{\frac{\zeta_1}{\pi}} \exp\left(-\zeta_1 (r - \mu_1)^2\right)$$

and

$$\sqrt{\frac{\zeta_2}{\pi}} \exp\left(-\zeta_2 (r - \mu_2)^2\right)$$

not centered on the same atom is again a Gaussian function

$$\sqrt{\frac{\zeta}{\pi}} \exp\left(-\zeta (r - \mu)^2\right)$$

with

$$\zeta = \zeta_1 + \zeta_2$$

and

$$\mu = \frac{\zeta_1 \mu_1 + \zeta_2 \mu_2}{\zeta_1 + \zeta_2}.$$

[82] Frisch et al. (1984).

the number of Gaussians for the single basis function describing core orbitals and each valence orbital is composed of z atomic basis functions, each composed of g_i Gaussians. The addition of polarization and diffuse functions on heavy atoms are denoted respectively by a $*$ or $+$ in the name, and a doubled symbol means that these functions are added also to Hydrogen. For example, the triple-zeta basis set with polarization functions is labeled *6-311G**, where the polarization function is a d-type orbital. We will be using such basis sets for large systems in this Thesis.

Karlsruhe basis sets: Another class of split-valence basis sets is the Karlsruhe family, denoted as $def2-xVP$, where $x = S, TZ, QZ$ stands for double, triple and quadruple zeta ($z = 2, 3, 4$). P stands for polarization functions and an added D stands for diffuse functions. For example, *def2-SVPD* is a double-zeta basis set with polarization and diffuse functions.

Correlation-consistent basis sets: Finally, the correlation-consistent basis set family, developed by Dunning [83], was designed specifically for systematically converging correlated quantum chemistry methods to the CBS limit. For a given basis size, Gaussian exponents and pre-factors are optimized to reproduce as closely as possible the converged MP2 (2^{nd} order Møller-Plesset) total energy in the CBS limit for isolated atoms. It is denoted as $cc-pVxZ$, where $cc-pV$ stands for *correlation-consistent polarized valence-only* and $x = D, T, Q, 5$ denotes the number of basis functions for each orbital; $z = 2, 3, 4, 5$ respectively. Their augmented versions with added diffuse functions have the prefix $aug-$. For example $aug-cc-pVQZ$ is a quadruple-zeta basis set with polarization and diffuse functions.

[83] Dunning Jr (1989).

3.3 Many-Body Perturbation Theory: *GW* & Bethe-Salpeter Formalisms

Spectroscopic experiments such as photoemission, inverse photoemission or optical absorption experiments perturb the system under investigation by promoting it into an excited state (see Chapter 2). Such a process therefore requires to go beyond the mean-field treatment of the ground state provided by Density-Functional Theory, and this is precisely the goal of theoretical spectroscopy. Many-Body Perturbation Theory (MBPT) provides a formally exact and systematic framework to describe the spectral properties of a system by combining Green's functions approaches with perturbation theory. In particular, the description of photoemission experiments requires a one-quasiparticle formulation that is enclosed in the one-particle Green's function. On the other hand, as anticipated in Sec. 2.3, the description of optical absorption experiments requires a two-quasiparticle formulation that is enclosed in the two-particle Green's function. After briefly introducing the Green's functions formalism, we will present two frameworks for the description of photoemission and optical absorption processes: respectively, the *GW* approximation and the Bethe-Salpeter Equation (BSE).

All the *GW* and Bethe-Salpeter calculations done throughout this Thesis use the Fiesta code, developed in our group [84, 85]. It is a Gaussian basis implementation of these formalisms that uses a resolution of the identity scheme and contour deformation. Notably, thanks to the recently developed space-time approach, it features an excellent scaling $\mathcal{O}(N^3)$ with system size as well as the possibility to be interfaced in order to perform *embedded* calculations, see Sec. 3.5.

3.3.1 Introduction to Green's functions

The concept of Green's functions was first introduced by British physicist and mathematician George Green in the context of inhomogeneous linear differential equations

$$\hat{\mathcal{L}}(\mathbf{r}, \nabla, \nabla^2, \ldots)\phi(\mathbf{r}) = S(\mathbf{r}) \tag{3.53}$$

where $\phi(\mathbf{r})$ is the function to be found and $S(\mathbf{r})$ is the source term. The Green's function of the linear operator $\hat{\mathcal{L}}$ is defined as its impulse response, i.e. the solution of Eq. (3.53) when the source is a delta function

$$\hat{\mathcal{L}}(\mathbf{r}, \nabla, \nabla^2, \ldots)G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'). \tag{3.54}$$

When $G(\mathbf{r}, \mathbf{r}')$ is known, the solution of Eq. (3.53) can be found for any source term as [21]

$$\phi(\mathbf{r}) = \int d\mathbf{r}'\, G(\mathbf{r}, \mathbf{r}')S(\mathbf{r}'). \tag{3.55}$$

Whenever $G(\mathbf{r}, \mathbf{r}')$ depends solely on the difference $\mathbf{r} - \mathbf{r}'$, this last operation is simply the convolution between $G(\mathbf{r} - \mathbf{r}')$ and $S(\mathbf{r}')$, which can readily be computed as a multiplication in Fourier space.

21. Consider for instance the Poisson equation for the electric potential

$$\nabla^2 V(\mathbf{r}) = -\frac{S(\mathbf{r})}{\epsilon_0}.$$

The equation for the Green's function is

$$\nabla^2 G(\mathbf{r} - \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$$

which readily gives upon Fourier transform $G(\mathbf{r}, \mathbf{r}') = -1/4\pi|\mathbf{r} - \mathbf{r}'|$. The electric potential is then recovered

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int d\mathbf{r}'\, \frac{S(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}$$

We are interested in applying this framework in the context of Quantum Mechanics, whereby the time evolution of the wavefunction is governed by the Schrödinger equation. Specifically, the action of the time evolution operator on single-particle states $|\phi\rangle$ is defined as

$$\left|\phi(t')\right\rangle = \hat{U}(t', t)\left|\phi(t)\right\rangle. \tag{3.56}$$

In other words, $\hat{U}(t', t)$ propagates the probability amplitude $|\phi\rangle$ from time t to time t'. The conservation of probability ensures that the time evolution operator is unitary, and when the Hamiltonian is stationary it takes the form

$$\hat{U}(t', t) = e^{-i(t'-t)\hat{H}}. \tag{3.57}$$

If one is interested in the propagation in both space and time, it is convenient to define the propagator as the transition amplitude of the time evolution operator between $\mathbf{r}$

and $\mathbf{r}'$, namely

$$K(\mathbf{r}'t', \mathbf{r}t) = \langle \mathbf{r}' | \hat{U}(t', t) | \mathbf{r} \rangle . \tag{3.58}$$

Therefore, the propagator allows to know the probability amplitude for the system to be in the state $|\phi(\mathbf{r}'t')\rangle$ when the one for the system in the state $|\phi(\mathbf{r}t)\rangle$ is known

$$\phi(\mathbf{r}'t') = \int d\mathbf{r}\, K(\mathbf{r}'t', \mathbf{r}t)\phi(\mathbf{r}t). \tag{3.59}$$

We can further enforce the operators with the largest time to be at the left of those with the smallest time, which naturally defines the time-ordered propagator as

$$K^T(\mathbf{r}'t', \mathbf{r}t) = \langle \mathbf{r}' | \hat{T}\left[\hat{U}(t', t)\right] | \mathbf{r} \rangle \tag{3.60}$$

where $\hat{T}$ is Wick's time ordering operator[22]. It is then possible to show that this propagator, when divided by i, is indeed the Green's function of the Schrödinger equation with an adequate source term

$$\left(i\frac{\partial}{\partial t'} - \hat{H}(\mathbf{r}')\right) K^T(\mathbf{r}'t', \mathbf{r}t) = i\hbar\delta(\mathbf{r} - \mathbf{r}')\delta(t - t'). \tag{3.61}$$

The above demonstration, established for systems described by one-body wavefunctions, can be generalized to systems defined by a many-body wavefunction. This is done by defining analogously the time-ordered one-particle Green's function as the correlation function of the field operators[23]

$$\boxed{iG_1(\mathbf{r}t, \mathbf{r}'t') = \left\langle N \middle| \hat{T}\left[\hat{\psi}(\mathbf{r}t)\hat{\psi}^\dagger(\mathbf{r}'t')\right] \middle| N \right\rangle ,} \tag{3.62}$$

where $\hat{\psi}$ and $\hat{\psi}^\dagger$ are the field operators in the Heisenberg picture (which contains the whole time dependence), and $|N\rangle$ stands for the normalized many-body ground-state (which does not contain any time-dependence). From Eq. (3.62) we can picture the physical interpretation of G_1 as the probability to find an electron in $\mathbf{r}$ at time t when an electron was added to the N-electron ground state in $\mathbf{r}'$ at time t', if $t > t'$. Instead, when $t < t'$, it describes the probability to find a hole in $\mathbf{r}'$ at time t' when a hole was added to the N-electron ground state in $\mathbf{r}$ at time t, see Fig. 3.4.

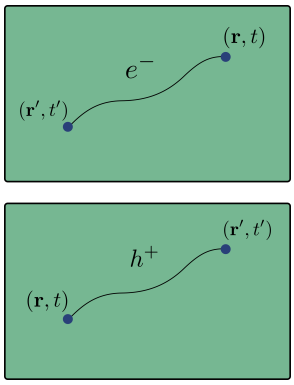

Figure 3.4: Physical interpretation of the one-particle Green's function G_1: when $t > t'$, an electron is added at $(\mathbf{r}', t')$ and is propagated to $(\mathbf{r}, t)$; when $t < t'$ a hole is added at $(\mathbf{r}, t)$ and is propagated to $(\mathbf{r}', t')$.

We anticipated that the single-particle Green's function contains all the information required to describe single-particle processes. Indeed, the one-particle density matrix is related to the diagonal elements of G_1

$$G_1(\mathbf{r}t, \mathbf{r}'t^+) = i n(\mathbf{r}, \mathbf{r}') \tag{3.63}$$

where $t^+ = t + 0^+$ denotes a infinitesimally positive shift above t. This has the effect of selecting the electron channel[24] as well as the time diagonal of G_1. Thus, the diagonal in both space and time of the Green's function is related to the electron density. This implies that the expectation value of any one-body operator $\hat{O}_1$ in the ground state can be obtained as

$$\langle N|\hat{O}_1|N\rangle = -i\int d\mathbf{r}\, dt\, O_1(\mathbf{r}t)G_1(\mathbf{r}t, \mathbf{r}t^+). \tag{3.64}$$

To further illustrate the usefulness of the Green's function, the ground-state total energy is given by the Galitskii-Migdal formula [86]

$$E_0 = -\frac{i}{2}\int d\mathbf{r} \lim_{\mathbf{r}'\to\mathbf{r}} \lim_{t'\to t^+} \left[i\frac{\partial}{\partial t} + H_0(\mathbf{r})\right] G_1(\mathbf{r}t, \mathbf{r}'t') \tag{3.65}$$

where $H_0(\mathbf{r}) = -\nabla^2/2 + V_{\text{ion}}(\mathbf{r})$ is the one-body part of the Hamiltonian. However, in practice this formula is hardly used as its efficiency and accuracy do not clearly supersede that of DFT.

22. The action of Wick's time ordering operator on the creation and annihilation field operators is

$$\hat{T}\left[\hat{\psi}(\mathbf{r}t)\hat{\psi}^\dagger(\mathbf{r}'t')\right] = \Theta(t - t')\hat{\psi}(\mathbf{r}t)\hat{\psi}^\dagger(\mathbf{r}'t') - \Theta(t' - t)\hat{\psi}^\dagger(\mathbf{r}'t')\hat{\psi}(\mathbf{r}t)$$

23. This definition is for the equilibrium Green's function at zero temperature, but it is possible to generalize it as to include the non-equilibrium and finite temperature cases.

24. The electron and hole channels are also called respectively the greater and lower Green's functions $G^>$ and $G^<$.

[86] Galitskii and Migdal (1958).

3.3.2 Lehmann representation

The Lehmann (or spectral) representation has many practical advantages as well as allowing an insightful physical interpretation of the one-particle Green's function in terms of photoemission processes. When the Hamiltonian does not explicitly depend on time, the Green's function depends only on the time difference $\tau = t - t'$. By introducing the closure relation in the Fock space in Eq. (3.62), we get

$$\begin{aligned} iG_1(\mathbf{r},\mathbf{r}',\tau) = \Theta(\tau)\sum_n \langle N|\hat{\psi}(\mathbf{r}t)|N+1,n\rangle\ \langle N+1,n|\hat{\psi}^\dagger(\mathbf{r}'t')|N\rangle \\ -\Theta(-\tau)\sum_n \langle N|\hat{\psi}^\dagger(\mathbf{r}'t')|N-1,n\rangle\ \langle N-1,n|\hat{\psi}(\mathbf{r}t)|N\rangle \end{aligned} \tag{3.66}$$

where the sum over n spans all states of the given particle number. The explicit time dependence of the matrix elements is obtained by switching to the Schrödinger picture, giving [25]

$$\begin{aligned} iG_1(\mathbf{r},\mathbf{r}',\tau) = \Theta(\tau)\sum_n \langle N|\psi(\mathbf{r})|N+1,n\rangle\ \langle N+1,n|\psi^\dagger(\mathbf{r}')|N\rangle\, e^{i(E_N - E_{N+1,n})\tau} \\ -\Theta(-\tau)\sum_n \langle N|\psi^\dagger(\mathbf{r}')|N-1,n\rangle\ \langle N-1,n|\psi(\mathbf{r})|N\rangle\, e^{-i(E_N - E_{N-1,n})\tau}, \end{aligned} \tag{3.67}$$

25. The relation between operators in the Heisenberg picture and those in the Schrödinger picture is $\psi(\hat{\mathbf{r}}t) = e^{i\hat{H}t}\psi(\mathbf{r})e^{-i\hat{H}t}$

where E_N denotes the N-electron ground-state energy and $E_{N\pm1,i}$ the $N\pm1$-electron energy in excited state i. In this form, we readily recognize the total energy differences that correspond to single-particle excitation energies: $E_{N+1,i} - E_N$ are the inverse photoemission levels and $E_N - E_{N-1,i}$ the direct photoemission levels. In a metal, taking $i = 0$ for ground-state total energies, these addition and removal energies are equal to the chemical potential μ, while in a semiconductor the latter lies within the band gap. We can thus define the excitation energies ϵ_n as

$$\epsilon_n = \begin{cases} E_{N+1,n} - E_N, & \epsilon_n > \mu \\ E_N - E_{N-1,n}, & \epsilon_n < \mu \end{cases} \tag{3.68}$$

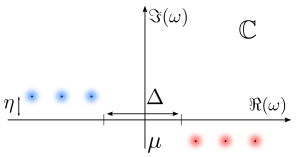

Figure 3.5: Pole structure of Lehmann's representation of the time-ordered Green's function in the complex plane.

and similarly we can define the so-called Lehmann amplitudes as

$$f_n(\mathbf{r}) = \begin{cases} \langle N|\psi(\mathbf{r})|N+1,n\rangle, & \epsilon_n > \mu \\ \langle N-1,n|\psi(\mathbf{r})|N\rangle, & \epsilon_n < \mu. \end{cases} \tag{3.69}$$

Then, the expression of the one-particle Green's function in real space is recast as

$$iG_1(\mathbf{r},\mathbf{r}',\tau) = \sum_n \left[\Theta(\tau)\Theta(\epsilon_n - \mu) - \Theta(-\tau)\Theta(\mu - \epsilon_n)\right] f_n(\mathbf{r}) f_n^*(\mathbf{r}') e^{-i\epsilon_n\tau} \tag{3.70}$$

Finally, the Lehmann representation of the Green's function is obtained by applying a Fourier transform in the time domain [26]

$$\boxed{iG_1(\mathbf{r},\mathbf{r}',\omega) = \sum_n \frac{f_n(\mathbf{r}) f_n^*(\mathbf{r}')}{\omega - \epsilon_n + i\eta\,\mathrm{sign}(\epsilon_n - \mu)},} \tag{3.71}$$

where η is a positive infinitesimal. As depicted in Fig. 3.5, this spectral representation highlights the polar structure of the Green's function in the complex plane. The poles located in the upper half-plane ($\epsilon_n < \mu$, blue) correspond to electron removal (hole addition) energies, while those located in the lower half-plane ($\epsilon_n > \mu$, red) correspond to electron addition (hole removal) energies [27].

26. This is done by adopting the Fourier transform convention

$$f(\omega) = \int d\tau\, f(\tau) e^{i\omega\tau}.$$

Then, the Fourier transform of the Heaviside function is

$$\Theta(\pm\tau) = \mp \lim_{\eta\to0} \frac{1}{2\pi i}\int_{-\infty}^{\infty} d\omega\, \frac{e^{i\omega\tau}}{\omega \pm i\eta}$$

27. Moreover, the electron density can be recovered directly by integrating the diagonal of G over the upper half-plane

$$\frac{1}{2\pi i}\int_{\mathcal{C}} d\omega\, e^{i\omega\eta} G(\mathbf{r},\mathbf{r},\omega) = n(\mathbf{r}), \tag{3.72}$$

where $\eta = 0^+$ is an infinitesimal positive number.

3.3.3 Spectral function

The application of the Sokhotski–Plemelj theorem on the Lehmann representation of the Green's function allows to extract its imaginary part [28]

$$\operatorname{Im} G_1(\mathbf{r},\mathbf{r}',\omega) = \pi \sum_n f_n(\mathbf{r}) f_n^*(\mathbf{r}') \delta(\omega - \epsilon_n) \operatorname{sign}(\mu - \epsilon_n). \tag{3.73}$$

We can therefore define the spectral function as

$$A(\mathbf{r},\mathbf{r}',\omega) = \frac{1}{\pi} \operatorname{sign}(\mu-\omega) \operatorname{Im} G_1(\mathbf{r},\mathbf{r}',\omega) = \sum_n f_n(\mathbf{r}) f_n^*(\mathbf{r}') \delta(\omega - \epsilon_n). \tag{3.74}$$

It contains the same information as the Green's function, with the advantage of being a real and positive function of ω. Indeed, the Cauchy relation on the domain of analyticity of G_1 allows to fully recover it from A

$$G_1(\mathbf{r},\mathbf{r}',\omega) = \int_{-\infty}^{\infty} d\omega' \, \frac{A(\mathbf{r},\mathbf{r}',\omega')}{\omega - \omega' + i\eta \operatorname{sign}(\omega' - \mu)}. \tag{3.75}$$

Furthermore, the closure relation on the Lehmann amplitudes implies a sum-rule on spectral function

$$\int_{-\infty}^{\infty} d\omega \, A(\mathbf{r},\mathbf{r}',\omega) = \delta(\mathbf{r}-\mathbf{r}'), \tag{3.76}$$

since

$$\sum_n f_n(\mathbf{r}) f_n^*(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'). \tag{3.77}$$

The spectral function provides information about the nature of the allowed electronic states, regardless whether they are occupied or not, and can be viewed as a many-body density of states. In fact its trace (diagonal) is the (local) density of states

$$n(\mathbf{r},\omega) = A(\mathbf{r},\mathbf{r},\omega), \qquad n(\omega) = \int d\mathbf{r} \, A(\mathbf{r},\mathbf{r},\omega). \tag{3.78}$$

The spectral function also directly yields the ground-state density

$$\int_{-\infty}^{\mu} d\omega \, A(\mathbf{r},\mathbf{r},\omega) = n(\mathbf{r}). \tag{3.79}$$

The name of the spectral function comes from its direct link to the photoemission spectrum. Indeed, the experimental observable that is measured in photoemission spectroscopy is the photocurrent, which is the probability of ejecting a photoelectron per unit time. It is given to a first approximation by Fermi's golden rule

$$J_{\mathbf{k}}(\omega) = \sum_n |\langle N|\hat{\mathbf{p}}|N-1,n,\mathbf{k}\rangle|^2 \delta(E_{\mathbf{k},n}(N-1) - E_0(N) - \omega) \tag{3.80}$$

where $\hat{\mathbf{p}}$ is the dipole operator, $|N\rangle$, $E_0(N)$ describe the initial ground-state with N electrons and $|N-1,n,\mathbf{k}\rangle$, $E_{\mathbf{k},n}(N-1)$ describe the final state with $N-1$ electrons and an ejected photoelectron with momentum $\mathbf{k}$. It can be shown that under the sudden approximation, the photocurrent is [87]

$$J_{\mathbf{k}}(\omega) = \sum_n |\hat{\mathbf{p}}_{\mathbf{k},n}|^2 A_{nn}(E_{\mathbf{k},n} - \omega). \tag{3.81}$$

Therefore, the photocurrent is related to the diagonal of the spectral function which is itself related to the local density of states that feature an energy difference equal to the photon energy.

28. The theorem can be stated as
$$\lim_{\eta\to 0} \frac{f(x)}{x+i\eta} = \mathcal{P}\frac{f(x)}{x} - i\pi f(x)\delta(x).$$
where $\mathcal{P}$ denotes the Cauchy principal value.

[87] Onida et al. (2002).

3.3.4 Dyson equation and Self-Energy

We are now in position to reformulate the many-body problem in terms of G_1 by stating the equation of motion of the Green's function. The conjunction of the linear response of the Green's function to an external perturbation with the hierarchy of equations of motion will allow to write an effective one-body equation for the Lehmann weights and the excitation energies. We begin by giving the Heisenberg equations of motion of the field operators

$$i\frac{\partial\hat{\psi}(\mathbf{r},t)}{\partial t} = \left[\hat{\psi}(\mathbf{r},t),\hat{H}\right], \tag{3.82}$$

where $\hat{H}$ is the Hamiltonian operator. The latter is expressed in second quantization as

$$\hat{H} = \int d\mathbf{r}\,\hat{\psi}^\dagger(\mathbf{r})H_0(\mathbf{r})\hat{\psi}(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}\,d\mathbf{r}'\,\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r}')V(\mathbf{r},\mathbf{r}')\hat{\psi}(\mathbf{r}')\hat{\psi}(\mathbf{r}), \tag{3.83}$$

where $H_0(\mathbf{r}) = -\nabla^2/2 + V_{\text{ion}}(\mathbf{r})$ is the one-body part of the Hamiltonian and V is the electron-electron potential. Combining Eq. (3.83) and Eq. (3.82) yields, after a few pages of calculations, the equation of motion of the Green's function

$$\left[i\frac{\partial}{\partial t_1} - H_0(\mathbf{r}_1)\right]G_1(1,2) + i\int d3\,V(1,3)G_2(1,3^+,2,3^{++}) = \delta(1,2), \tag{3.84}$$

where the $(+)$ and $(++)$ account for the proper ordering of the field operators in the (time-ordered) two-particle Green's function G_2, defined as

$$i^2G_2(1,2,1',2') = \left\langle N\middle|\hat{T}\left[\hat{\psi}(1)\hat{\psi}(2)\hat{\psi}^\dagger(2')\hat{\psi}^\dagger(1')\right]\middle|N\right\rangle. \tag{3.85}$$

Thus, in order to build G_1, we need the two-particle Green's function G_2, which itself depends on the three-particle one, and so on up to all orders. This is the merciless hierarchy of the Green's functions: we have merely restated the many-body problem, and we have not yet gained anything by doing so. The idea of Many-Body Perturbation Theory is that if one is interested in one-body processes (such as photoemission), only the one-particle Green's function is needed and so one must try approximate the two-particle Green's function in terms of it. Likewise, if one is interested in two-body processes (such as absorption), one has to work with the two-particle Green's function and find approximations for the higher-order ones in terms of it.

The key idea for breaking the hierarchy of Green's function is to relate the two-particle Green's function with the linear response of the system to an external perturbation. Following Schwinger's functional derivative technique, we introduce a non-local external perturbation [29] $U(1,2)$ that will eventually be set to 0 at the end of the derivation. Then, it can be shown that the variation of G_1 with respect to the perturbation $U(1,2)$ is [88]

$$\frac{\delta G_1(1,2)}{\delta U(3,4)} = G_1(1,2)G_1(4,3) - G_2(1,4,2,3). \tag{3.86}$$

The two-particle Green's function G_2 now appears in both the equation of motion and in the linear response of the one-particle Green's function. We can thus eliminate G_2 in the equation of motion for G_1 in Eq. (3.84) by replacing it with two terms involving only the one-particle Green's function. If we consider a local perturbation $U(3)\delta(3,4)$, the equation of motion of the Green's function in Eq. (3.84) can be restated as

$$\left[i\frac{\partial}{\partial t_1} - H_0(\mathbf{r}_1) + i\int d3\,V(1,3)G_1(3,3^+)\right]G_1(1,2) - \int d5\,\Sigma(1,5)G_1(5,2) = \delta(1,2) \tag{3.87}$$

29. The introduction of a *non-local* perturbation is necessary because any variation of the external potential induces a variation of G_2, hence making the total perturbation non-local as well.

[88] Martin and Schwinger (1959).

where we defined the Self-Energy as

$$\Sigma(1,2) = i\int d3\, d4\, V(1^+,3)\frac{\delta G(1,4)}{\delta U(3)}G_1^{-1}(4,2). \tag{3.88}$$

Since $-iG_1(3,3^+)$ is the electron density, the last term in the brackets of Eq. (3.87) is simply the Hartree potential V_H. Hence, we can identify the Self-Energy Σ to the quantity that accounts for all many-body effects beyond the Hartree term; in other words it is the generalization of the Exchange-Correlation potential of DFT. In fact, it is possible to recover a one-body Schrödinger-like equation by inserting the Lehmann representation of the one-particle Green's function of Eq. (3.71) into Eq. (3.87):

$$\boxed{\Big[H_0(\mathbf{r}) + V_H(\mathbf{r})\Big] f_n(\mathbf{r}) + \int d\mathbf{r}'\, \Sigma(\mathbf{r},\mathbf{r}',\epsilon_n) f_n(\mathbf{r}') = \epsilon_n f_n(\mathbf{r}),} \tag{3.89}$$

where ϵ_n are true addition or removal energies that are related to the photoemission levels, and f_n are the Lehmann amplitudes that play a similar role to molecular orbitals. This equation is referred to as the quasiparticle equation; it is an eigenvalue problem that describes the energetics of electron addition when $\epsilon_n > \mu$ or hole addition when $\epsilon_n < \mu$. The difference with the one-body formulation of DFT being that now the Self-Energy is a frequency-dependent non-local operator $\Sigma(\mathbf{r},\mathbf{r}',\omega)$ [30].

30. The resemblance with Eq. (3.30) is led by the substitution

$$\Sigma(\mathbf{r},\mathbf{r}',\epsilon_n) \to V_{xc}[n](\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')$$

The self-energy is also in general a non-hermitian operator providing an imaginary part to the ϵ_n, namely the lifetime of the quasiparticles with respect to electron-electron scattering.

Moreover, the self-energy allows to relate the interacting Green's function to the non-interacting one-body Green's function G_0, defined from

$$\left[i\frac{\partial}{\partial t_1} - H_0(\mathbf{r}_1)\right] G_0(1,2) = \delta(1,2), \tag{3.90}$$

where H_0 is the non-interacting Hamiltonian defined above. We see that G_0 is merely the functional inverse of $\omega - H_0$ in frequency space. By inserting this definition into Eq. (3.87), we obtain

$$\Big[G_0^{-1}(1,3) - \Sigma(1,3) - V_H(1)\Big] G_1(3,2) = \delta(1,2). \tag{3.91}$$

Or, in matrix notation, $G_1 = G_0 + G_0\Sigma G_1$. This is a self-consistent Dyson equation that can be inverted: by multiplying by G_1^{-1} to the right and by G_0^{-1} to the left one obtains, in matrix notation,

$$G_1^{-1} = G_0^{-1} - \Sigma - V_H. \tag{3.92}$$

Eq. (3.92) allows to interpret the self-energy as the operator bridging from the non-interacting system to the fully interacting one.

For the time being, we have just formulated the evolution the one-particle Green's function in terms of the Self-Energy operator without making any approximation. As such, computing G_1 is as difficult as the original many-body problem, and the next step is therefore to design efficient approximations for the Self-Energy that will allow to solve Eq. (3.89).

3.3.5 Hedin's equations

[89] Hedin (1965).

In a seminal paper of 1965, Hedin [89] used the Schwinger functional derivative technique to perform a perturbative expansion of the Self-Energy. Since former expansions in terms of the non-interacting Green's function G_0 and the bare Coulomb potential V were known to lead to convergence difficulties with respect to adding

higher order terms, Hedin devised an expansion in terms of the interacting Green's function G_1 and screened Coulomb potential W. Indeed, the latter is naturally better suited for performing a perturbative expansion as it is reduced with respect to the bare Coulomb potential by the medium's dielectric function.

Following again Schwinger's functional derivative technique, we introduce a local potential $U(1)$ that will eventually be set to zero at the end of the derivation. We know how to obtain G_1 in terms of G_0 and Σ thanks to the Dyson equation of Eq. (3.92). Then, we proceed in defining the total electrostatic potential as the sum of U and the Hartree potential V_H

$$F(1) = U(1) - i\int d2\, V(1,2)G_1(2,2^+). \tag{3.93}$$

We then define the dielectric function as

$$\varepsilon^{-1}(1,2) = \frac{\delta F(1)}{\delta U(2)} = \delta(1,2) + \int d3\, V(1,3)\chi(3,2), \tag{3.94}$$

where

$$\chi(1,2) = -i\frac{\delta G_1(1,1^+)}{\delta U(2)} \tag{3.95}$$

is the *reducible* polarizability, in the sense that it corresponds to a variation of the charge density with respect to the bare external potential U. We can indeed define the *irreducible* polarizability as the density response to the total electrostatic potential F

$$\tilde{\chi} = -i\frac{\delta G_1(1,1^+)}{\delta F(2)}. \tag{3.96}$$

The relation between the two is obtained by making use of the chain rule

$$\chi(1,2) = \tilde{\chi}(1,2) + \int d34\, \tilde{\chi}(1,3)V(3,4)\chi(4,2), \tag{3.97}$$

leading to a self-consistent equation, also called a Dyson equation. These definitions lead to the natural expression of the screened Coulomb potential as the product of the bare Coulomb potential and the inverse dielectric function

$$W(1,2) = \int d3\, V(1,3)\varepsilon^{-1}(3,2). \tag{3.98}$$

The latter can be recast in terms of the reducible and irreducible polarizabilities as

$$W(1,2) = V(1,2) + \int d34\, V(1,3)\chi(4,2)V(3,4) = V(1,2) + \int d3\, d4\, V(1,3)\tilde{\chi}(3,4)W(4,2). \tag{3.99}$$

Furthermore, let us define the *irreducible* vertex function, a three-body term, as

$$\tilde{\Gamma}(1,2,3) = -\frac{\delta G_1^{-1}(1,2)}{\delta F(3)}. \tag{3.100}$$

Combining this expression with the one for the screened Coulomb potential, the Self-Energy defined in Eq. (3.88) can be recast as [31]

$$\Sigma(1,2) = i\int d34\, G_1(1,4)W(3,1^+)\tilde{\Gamma}(4,2,3). \tag{3.101}$$

The last equation allows to express Σ in terms of W and $\tilde{\Gamma}$. An expression of the irreducible vertex function can be obtained by using the Dyson equation of Eq. (3.92) under a perturbation, namely $G_1^{-1} = G_0^{-1} - \Sigma - F$, where G_0 denotes the non-interacting

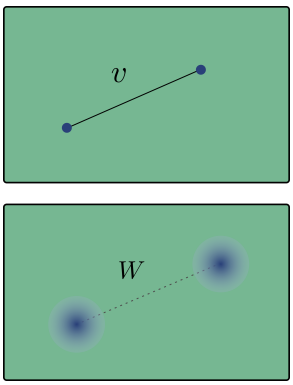

Figure 3.6: Reformulation of two electrons interacting though the bare Coulomb potential into two quasi-particles (dressed by their exchange-correlation hole) weakly interacting through the screened Coulomb potential.

31. We made use of the functional derivative of the inverse

$$\frac{\delta G_1(1,4)}{\delta F(3)} G_1^{-1}(4,2) = -G_1(1,4)\frac{\delta G_1^{-1}(4,2)}{\delta F(3)},$$

and the chain rule was inserted

$$\frac{\delta}{\delta U(3)} = \frac{\delta}{\delta F(4)}\varepsilon^{-1}(4,3).$$

Green's function,

$$\begin{aligned}\widetilde{\Gamma}(1,2,3) &= \delta(1,2)\delta(1,3) + \frac{\delta\Sigma(1,2)}{\delta V(3)} \\ &= \delta(1,2)\delta(1,3) + \int d4567\, \frac{\delta\Sigma(1,2)}{\delta G_1(4,5)} G_1(4,6) G_1(7,5)\widetilde{\Gamma}(6,7,3).\end{aligned} \tag{3.102}$$

This is a closed integral equation for the vertex function which brings into play the 4-point kernel $\delta\Sigma/\delta G_1$. We are now granted with a set of self-consistent equations that can be solved cyclically. Taking the limit of $U \to 0$, we arrive at Hedin's equations

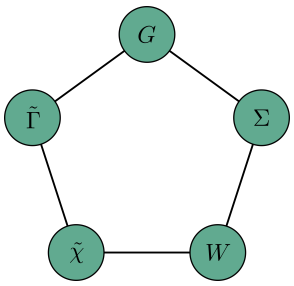

Figure 3.7: Hedin's pentagram.

$$\boxed{\begin{aligned}
G_1(1,2) &= G_0(1,2) + \int d34\, G_0(1,3)\Sigma(3,4)G(4,2) && \text{(3.103a)}\\
\widetilde{\Gamma}(1,2,3) &= \delta(1,2)\delta(1,3) + \int d4567\, \frac{\delta\Sigma(1,2)}{\delta G(4,5)} G(4,6)G(7,5)\widetilde{\Gamma}(6,7,3) && \text{(3.103b)}\\
\tilde{\chi}(1,2) &= -i\int d34\, G(2,3)G(4,2)\widetilde{\Gamma}(3,4,1) && \text{(3.103c)}\\
W(1,2) &= V(1,2) + \int d34\, V(1,3)\tilde{\chi}(3,4)W(4,2) && \text{(3.103d)}\\
\Sigma(1,2) &= i\int d34\, G(1,4)W(3,1^+)\widetilde{\Gamma}(4,2,3) && \text{(3.103e)}
\end{aligned}}$$

Hedin's equations are an iterative solution to the many-body problem that allow to construct the Green's function G_1 from an initial guess of the non-interacting Green's function G_0, as depicted in Fig. 3.7.

3.3.6 *GW* approximation

A. Derivation

As was mentioned earlier, building a perturbation theory in terms of W instead of V allows for a better convergence thanks to the sizable screening described by the dielectric function. The so-called *GW* approximation consists in retaining only first-order terms in W within Hedin's equations, or equivalently in iterating Hedin's equations only once [89]. Indeed, doing so by initializing $G_1 = G_0$ and $\Sigma = 0$ yields

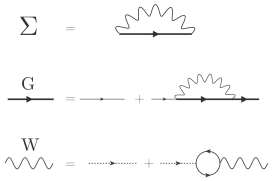

Figure 3.8: Feynmann diagrams corresponding to the *GW* equations. Σ is reduced to the product of G and W, G_0 is connected to G_1 by means of Σ and similarly V is connected to W through $\tilde{\chi}$.

$$\boxed{\begin{aligned}
\widetilde{\Gamma}(1,2,3) &= \delta(1,2)\delta(1,3) && \text{(3.104a)}\\
\tilde{\chi}(1,2) &= -iG_0(2,1)G_0(1,2) && \text{(3.104b)}\\
W(1,2) &= V(1,2) + \int d34\, V(1,3)\tilde{\chi}(3,4)W(4,2) && \text{(3.104c)}\\
\Sigma(1,2) &= iG_0(1,2)W(2,1^+). && \text{(3.104d)}
\end{aligned}}$$

In the frequency domain, the self-energy is written as a convolution of the Green's function with the screened Coulomb potential

$$\Sigma(\mathbf{r},\mathbf{r}',\epsilon) = -\frac{1}{2\pi i}\int d\omega\, e^{i\omega 0^+} G(\mathbf{r},\mathbf{r}',\epsilon+\omega)W(\mathbf{r},\mathbf{r}',\omega). \tag{3.105}$$

This set of equations constitutes the *GW* approximation, which can be seen as a first-order perturbation expansion of the self-energy in terms of W: injecting this result back into Hedin's equations in Eq. (3.103) would lead to a vertex function depending on W, and therefore a second-order dependence in W for Σ.

In practice, Eqs. (3.104) are solved by taking as a starting point the solution of a Hamiltonian that is not the Hartree Hamiltonian, which is a rather poor description

of the N-electron system, but rather a mean-field one which can be either that of DFT or HF. The initial Green's function G_0 stems from the mean-field energies ϵ_n^0 and molecular orbitals ϕ_n, and the initial irreducible polarizability χ_0 is built from G_0, resulting in the following expressions in frequency space

$$G_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_n \frac{\phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')}{\omega - \epsilon_n^0 + i\eta \mathrm{sign}(\epsilon_n^0 - \mu)} \tag{3.106a}$$

$$\chi_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_{mn} (f_n - f_m) \frac{\phi_n^*(\mathbf{r})\phi_m(\mathbf{r})\phi_m^*(\mathbf{r}')\phi_n(\mathbf{r}')}{\epsilon_n^0 - \epsilon_m^0 - \omega + i\eta}, \tag{3.106b}$$

where f_n are the occupation numbers with respect to the chemical potential μ. Then, one can state the quasiparticle equation of Eq. (3.89) with a one-body part that is replaced by that of the chosen mean-field Hamiltonian. Its projection onto the ϕ_n yields

$$\epsilon_n = \epsilon_n^0 + \langle \phi_n | \Sigma(\epsilon_n) - V_{xc} | \phi_n \rangle . \tag{3.107}$$

This expression thus allows to correct the mean-field energies by explicitly replacing the mean-field exchange-correlation potential by the self-energy. Because Σ has to be evaluated at the *GW* quasiparticle energies, Eq. (3.107) is a self-consistent equation.

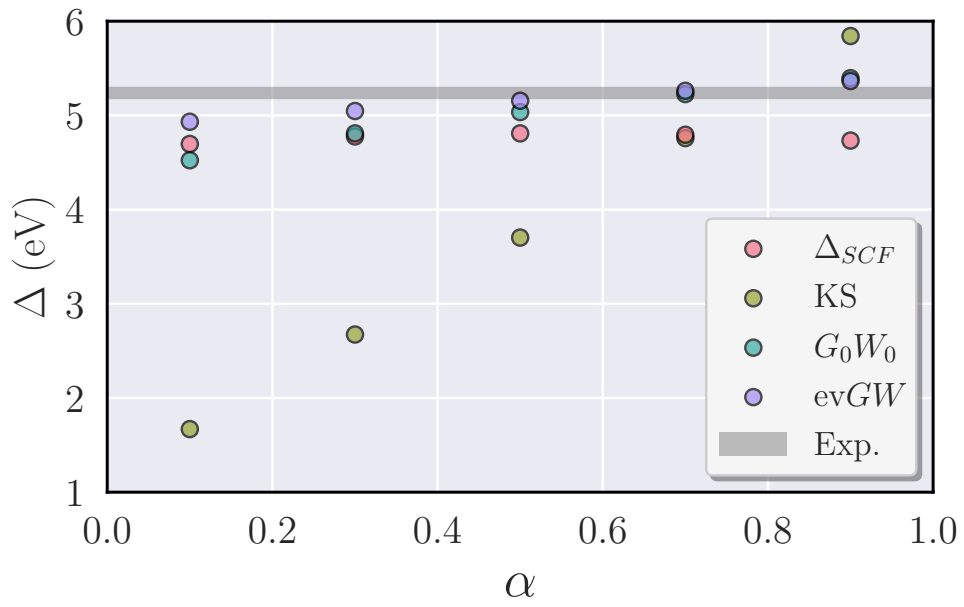

Figure 3.9: Predicted photoemission gap of a single Pentacene molecule as a function of the portion of exact exchange α in the hybrid PBEh functional.

The procedure of applying Eqs. (3.106) directly to build the self-energy and correcting the *GW* levels is called the G_0W_0 procedure, in reference to the fact that the susceptibility and the Green's function are taken to be those of the chosen mean-field starting point (e.g. DFT or HF). In Fig. 3.9, the gap of a single Pentacene molecule is presented as a function of the amount of exact exchange in the starting PBEh(α) hybrid functional. The frightening linear trend of the Kohn-Sham gap with respect to α hinders the hope of predicting the gap without an additional tuning strategy for the empirical parameter of the functional. Instead, the Δ-SCF procedure consists in computing the gap as $E_0[N-1] + E_0[N+1] - 2E_0[N]$, where E_0 denotes the ground-state DFT total energy of the cation, anion and neutral system respectively. As it is a difference of total energies, this strategy features a much reduced dependence on the amount of exact exchange, although it never crosses the experimental value. Because the G_0W_0 procedure does not update the susceptibility and the Green's function, it also features a small dependence on the initial guess with the added benefit of being globally closer to the experimental value.

The partially self-consistent *GW* procedure, dubbed ev*GW*, consists in updating the energy levels ϵ_n^0 by cycling through Eqs. (3.106) and Eq. (3.107), while keeping the molecular orbitals ϕ_n fixed. In particular, this strategy features a better accuracy while allowing to get rid of most of the dependence of the *GW* levels on the initial guess, as can be seen in Fig. 3.9. It is the ev*GW* scheme that has been used throughout this Thesis, starting typically with Kohn-Sham eigenstates with the PBE0 functional. I would like to finally mention the quasiparticle self-consistent qs*GW* procedure [90], in which the Lehmann weights are updated as well in order to assess the *GW* corrections on the density. In the case of organic systems, the simple ev*GW* approach has been shown to provide results very close to the much more expensive qs*GW* scheme [91].

[90] Schilfgaarde et al. (2006).

[91] Kaplan et al. (2016).

Overall, the *GW* approximation provides very good estimates of the photoemission spectra and more importantly it overcomes the strong dependence of DFT results on the functional [92]. Systems containing up to several hundreds of atoms can be dealt with at the *GW* level using efficient implementations that we describe below.

[92] Golze et al. (2019).

B. *Implementation*

The solution of Eq. (3.107) for a particular level can be found at the intersect of $f(\omega) = \omega - \epsilon_n^0 + \langle\phi_n|V_{xc}|\phi_n\rangle$ and $g(\omega) = \langle\phi_n|\Sigma(\omega)|\phi_n\rangle$. The frequency integration of the self-energy in Eq. (3.105) is computationally demanding, and can be performed via several methods. Contour Deformation [93, 94] proceeds by integrating Σ along the imaginary frequency axis, plus the residual of W associated with the poles of $G_1(E+\omega)$ that can be found in the 2nd and 4th quadrant, hence avoiding the numerically unstable region along the real axis. Analytic Continuation of Σ from the imaginary to the real frequency exploits instead the smoothness of the self-energy along the imaginary frequency axis, with the difficulty that it is a notably unstable technique, since the self-energy presents poles on the real axis (see e.g. the difficulties in [95]). Both techniques are widely used in *GW* implementations and allows for full-frequency integration. In this Thesis, and as originally implemented in the Fiesta package, we use contour deformation techniques.

[93] Godby et al. (1988).

[94] March (1999).

[95] Setten et al. (2015).

As in the case of DFT, the *GW* implementation expands the molecular orbitals ϕ_n (or the quasiparticle wavefunctions f_n expressed in their terms) onto a set of normalized basis functions g_m

$$\phi_n(\mathbf{r}) = \sum_m \mathcal{C}_m(\phi_n) g_m(\mathbf{r}), \tag{3.108}$$

where $\mathcal{C}_m$ are the expansion coefficients to be determined. The basis set choice is generally guided by the type of problem under investigation: plane waves are commonly used for periodic systems while Gaussian functions are preferred for finite molecular systems. The construction of the Coulomb and exchange matrices in a Gaussian basis involves the computation of costly four-center integrals such as $V_{klnm} = \int d\mathbf{r}\, d\mathbf{r}'\, g_k(\mathbf{r}) g_l(\mathbf{r}') V(\mathbf{r}-\mathbf{r}') g_n(\mathbf{r}) g_m(\mathbf{r}')$ that can be strongly accelerated by projecting any product of Gaussian orbitals $g_n(\mathbf{r}) g_m(\mathbf{r})$ onto an auxiliary basis $p_\mu(\mathbf{r})$

$$g_n(\mathbf{r}) g_m(\mathbf{r}) = \sum_\mu \mathcal{F}_\mu(g_n g_m) p_\mu(\mathbf{r}). \tag{3.109}$$

where $\mathcal{F}_\lambda$ are the expansion coefficients for the auxiliary basis. Then, the aforementioned four-center integrals can be computed as two-center integrals in this auxiliary basis. The coefficients are determined by employing the resolution of the identity (RI) scheme, whereby an operator $\hat{O}$ is used to calculate

$$\mathcal{F}_\mu(g_n g_m) = \sum_\nu \hat{O}^{-1}_{\mu\nu} \langle p_\nu|\hat{O}|g_m g_n\rangle\,. \tag{3.110}$$

The latter can be either the overlap matrix (RI-I), the Coulomb matrix (RI-V) or a hybrid operator mixing the two (RI-H) [96, 97]. We will be using the more accurate RI-V Coulomb fitting method in this Thesis. Another important point to be aware of is the basis set convergence of the quasiparticle energies. It is well established that correlated electronic structure are difficult to converge with respect to the number of basis functions in order to give reliable results [92]. Within a plane wave basis set, reaching the complete basis set limit is straightforward as the energy cutoff can be smoothly increased (though at a price that may be prohibitive). Gaussian basis sets feature instead only a limited number of sizes and face the problem of non-orthogonality, i.e. linear redundancy when adding basis functions, making the extrapolation to the complete basis set limit more delicate. Typically, the quasiparticle energies are fitted empirically with an exponential function of the cardinality of the basis set (X=1,2,3 for single-, double- and triple-zeta basis sets), as shown in Fig. 3.10.

[96] Ren et al. (2012).

[97] Duchemin et al. (2017).

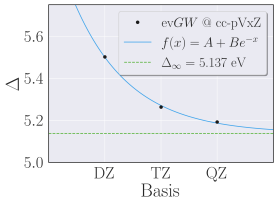

Figure 3.10: Donor-Acceptor gap of the Perylene-TBPA complex [98] as a function of the basis set size. The horizontal dashed line corresponds to the extrapolated value, assuming an exponential trend with the basis set cardinality.

3.3.7 COHSEX decomposition

Our treatment of environmental dielectric effects in Sec. 3.5 will invite us to use a simplified version of the self-energy. Within the *GW* approximation, the self-energy is expressed as the direct product or a convolution of the Green's function and the screened Coulomb potential, as in Eq. (3.103) and Eq. (3.105) respectively.

In order to gain physical insight on the *GW* approximation, the screened Coulomb potential can be split as $W(\mathbf{r},\mathbf{r}',\omega) = V(\mathbf{r},\mathbf{r}') + W_p(\mathbf{r},\mathbf{r}',\omega)$, where V is the bare Coulomb potential and $W_p = V\chi V$ is the so-called reaction field associated with the polarization of the system under the effect of an added or removed charge. By employing the Lehmann representation of G_1 in Eq. (3.71) and the spectral representation of W_p, it is possible to recast Σ from the residue theorem as

$$\Sigma(\mathbf{r},\mathbf{r}',\omega) = \Sigma_{SEX}(\mathbf{r},\mathbf{r}',\omega) + \Sigma_{COH}(\mathbf{r},\mathbf{r}',\omega) \tag{3.111a}$$

$$\Sigma_{SEX}(\mathbf{r},\mathbf{r}',\omega) = -\sum_n^{occ} f_n(\mathbf{r}) f_n^*(\mathbf{r}') W(\mathbf{r},\mathbf{r}',\epsilon_n - \omega) \tag{3.111b}$$

$$\Sigma_{COH}(\mathbf{r},\mathbf{r}',\omega) = \sum_n f_n(\mathbf{r}) f_n^*(\mathbf{r}') \mathcal{P}\int_0^\infty d\omega' \frac{B(\mathbf{r},\mathbf{r}',\omega')}{\omega - \epsilon_n - \omega'}, \tag{3.111c}$$

where $B(\mathbf{r},\mathbf{r}',\omega)$ is the spectral function of W_p, and $\mathcal{P}$ denotes the principal value. The first term originates from the poles of G_1 and can readily be recognized as a dynamically screened exchange (SEX) since in a non-polarizable system $W = V$ and Σ_{SEX} would simply become the regular Fock exchange operator on the Lehmann amplitudes [32]. The second term originates from the poles of W and corresponds instead to the so-called Coulomb hole (COH) since it is related to the interaction of an electron with its adiabatically built correlation hole [89, 99].

32. Indeed, the bare exchange is given by
$$\Sigma_X(\mathbf{r},\mathbf{r}') = -\sum_n f_n(\mathbf{r}) f_n^*(\mathbf{r}') V(\mathbf{r},\mathbf{r}')$$

[99] Hedin and Lundqvist (1970).

The former decomposition of the self-energy is exact, but one can derive the static approximation to the exact COHSEX decomposition assuming that $\epsilon_n - \omega \simeq 0\ , \forall n$, leading to

$$\Sigma_{SEX}(\mathbf{r},\mathbf{r}') = -\sum_n^{occ} f_n(\mathbf{r}) f_n^*(\mathbf{r}') W(\mathbf{r},\mathbf{r}',\omega = 0) \tag{3.112a}$$

$$\Sigma_{COH}(\mathbf{r},\mathbf{r}') = \frac{1}{2}\delta(\mathbf{r}-\mathbf{r}') W_p(\mathbf{r},\mathbf{r}',\omega = 0). \tag{3.112b}$$

The static COHSEX approximation is often simply referred to as the COHSEX approximation, as will be the case thereafter. Within this static approximation of the full *GW* self-energy, the Coulomb hole is not only static but also local.

The latter can be interpreted in a classical picture as the energy required to build an electron or a hole in the polarizable system. Indeed, when a point charge $\pm\delta(\mathbf{r}-\mathbf{r}_0)$ is added to the system, it generates a potential $\pm V(\mathbf{r},\mathbf{r}_0)$ that in turn induces a density variation

$$\delta n(\mathbf{r}) = \pm \int d\mathbf{r}'\, \chi(\mathbf{r},\mathbf{r}')V(\mathbf{r}',\mathbf{r}_0). \tag{3.113}$$

This density variation creates an induced potential

$$\delta V_{ind}(\mathbf{r}'') = \pm \int d\mathbf{r}\, V(\mathbf{r}'',\mathbf{r})\delta n(\mathbf{r}) = \pm \int d\mathbf{r}\, d\mathbf{r}'\, V(\mathbf{r}'',\mathbf{r})\chi(\mathbf{r},\mathbf{r}')V(\mathbf{r}',\mathbf{r}_0), \tag{3.114}$$

where we note the appearance of the reaction field $W_p(\mathbf{r}'',\mathbf{r}_0)$ in the last term. Then, the energy required to adiabatically build the charge distribution $\pm\delta(\mathbf{r}-\mathbf{r}_0)$ is

$$E_{ind} = \pm \int d\lambda\, \lambda \int d\mathbf{r}''\, \delta(\mathbf{r}-\mathbf{r}_0)\delta V_{ind}(\mathbf{r}'') = \frac{1}{2} W_p(\mathbf{r}_0,\mathbf{r}_0), \tag{3.115}$$

where the parameter $\lambda \in [0,1]$ builds adiabatically the added charge. This energy is precisely the the Coulomb hole self-energy.

While the COHSEX approximation is known to produce band gaps that are too large, it will be especially useful in the context of embedded Many-Body Perturbation Theory, whereby the Green's function formalism is interfaced with a coarser classical theory that considers only the low-frequency ($\omega \to 0$) dielectric response of the classical system (electronic degrees of freedom), see Sec. 3.5.

3.3.8 The Bethe-Salpeter Equation

So far, the Green's function formalism allowed us to describe the charged excitations in a many-body system as the propagation of quasi-electrons and quasi-holes, namely particles dressed by their exchange-correlation hole, since these are really the quasi-particles that are added or removed from the system in a photoemission experiment. However, optical absorption experiments involve pairs of *interacting* quasi-electrons and quasi-holes, which are called excitons. Describing these two-particle excitations with a one-particle Green's function would result in treating them as non-interacting quasiparticles. Instead, the Bethe-Salpeter Equation (BSE) characterizes the optical excitations of the many-body system by considering the two-particle Green's function G_2. The Bethe-Salpeter formalism is of central importance in this Thesis as we will be concerned with charge-transfer excitations whereby long-range electron-hole interactions must properly be taken into account, a feature that is ill-described within time-dependent DFT with functionals containing little amount of exact exchange. We will not describe here time-dependent DFT since we chose to work with the Bethe-Salpeter formalism, whose success arises from its impressive ability to reproduce absorption peaks with 10% precision and peak strengths with 20% precision, while being computationally affordable [87].

A. Derivation

In analogy with the definition of the reducible polarizability of Eq. (3.96) [33], let us introduce the four-point response function to a non-local perturbation $U(2',2)$

$$L(1,2,1',2') = -i\frac{\delta G_1(1,1')}{\delta U(2',2)}. \tag{3.116}$$

Again, the central result of linear response in the Green's function formalism of Eq. (3.86) can be applied to Eq. (3.116), namely

$$iL(1,2,1',2') = G_1(1,1')G_1(2,2') - G_2(1,2,1',2'). \tag{3.117}$$

33. The response function is actually a different notation for the four-point susceptibility

$$\chi(1,1',2,2') = -i\frac{\delta G_1(1,1')}{\delta U(2,2')}.$$

Since the electron density is $n(1) = -iG_1(1,1^+)$, the two-point susceptibility is related to the four-point one by

$$\chi(1,2) = \chi(1,1^+,2,2).$$

In this form, L appears to describe the correlation part of the two-particle Green's function i.e. the difference between the motion of the independent quasiparticles (an electron from 1 to $1'$ and a hole from $2'$ to 2) and that of the coupled electron-hole pair. If one writes the derivative of G_1 in Eq. (3.116) in terms of the derivative of G_1^{-1}, by making use of the same functional properties as in the derivation of Hedin's equations, one is left with

$$L(1,2,1',2') = i\int d34\, G_1(1,3)\frac{\delta G_1^{-1}(3,4)}{\delta U(2',2)}G_1(4,1'). \tag{3.118}$$

The derivative of G_1^{-1} is then conveniently carried out from the Dyson equation $G_1^{-1} = G_0^{-1} - \Sigma - V_H - U$. Using the fact that G_0 is independent of U, and that $\delta U(3,4)/U(2',2) = \delta(3,2')\delta(4,2)$, the response function reads

$$L(1,2,1',2') = -iG_1(1,2')G_1(2,1') - i\int d34\, G_1(1,3)\frac{\delta\big(V_H(3)\delta(3,4)+\Sigma(3,4)\big)}{\delta U(2',2)}G_1(4,1'). \tag{3.119}$$

Furthermore, we define the non-interacting response function as

$$L_0(1,2,1',2') = -iG_1(1,2')G_1(2,1') \tag{3.120}$$

and the Bethe-Salpeter kernel as

$$\Xi(1,2,3,4) = i\frac{\delta}{\delta G_1(4,2)}\Big(V_H(1)\delta(1,3)+\Sigma(1,3)\Big). \tag{3.121}$$

Injecting the chain rule of derivative in order to obtain derivatives with respect to G_1 in Eq. (3.119) one obtains a factor $-i\,\delta G(5,6)/\delta U(2,2')$ that is readily identified to $L(5,2,6,2')$, thus leading to the Dyson equation for L

$$\boxed{L(1,2,1',2') = L_0(1,2,1',2') + \int d3456\, L_0(1,4,1',3)\Xi(3,6,4,5)L(5,2,6,2').} \tag{3.122}$$

This is the Bethe-Salpeter equation for the four-point polarizability, which is simply $L = L_0 + L_0\Xi L$ in matrix form. The way the kernel Ξ links L_0 to L is strictly analogous to how the self-energy Σ links G_0 to G_1. Besides, Eq. (3.122) yields its TD-HF and TD-DFT analogues when replacing Ξ by their respective kernels.

While Eq. (3.122) is formally exact, it is necessary in practice to specify the Bethe-Salpeter kernel within certain approximations. Since the latter depends explicitly on the self-energy, it is tempting to apply the *GW* approximation directly. The derivative of the Hartree term yields in general

$$\frac{\delta V_H(1)}{\delta G_1(4,2)} = -iV(1,4)\delta(4,2). \tag{3.123}$$

On the other hand, the derivative of the self-energy contains a term $\delta W/\delta G_1$ which is to higher order in *W* and is thus traditionally neglected, an approximation that has often been considered harmless (even though its formal validation seems difficult to find in the literature). Within this approximation, the derivative of the self-energy reads

$$\frac{\delta\Sigma(3,4)}{\delta G_1(5,6)} = i\delta(3,5)\delta(4,6)W(4,3^+). \tag{3.124}$$

Finally, the screened Coulomb potential is generally restricted to its static limit $\omega \to 0$ in the so-called adiabatic approximation. As such, the Bethe-Salpeter kernel is

$$\Xi(1,2,3,4) = V(1,4)\delta(1,3)\delta(2,4) - W(1^+,3)\delta(t_1-t_3)\delta(1,4)\delta(2,3). \tag{3.125}$$

The Bethe-Salpeter equation simplifies accordingly

$$\boxed{L(1,2,1',2') = L_0(1,2,1',2') + \int d34\, L_0(1,3,1',3)\Big[V(3,4)L(4,2,4,2') - W(3^+,4)\delta(t_3-t_4)L(3,2,4,2')\Big].} \tag{3.126}$$

This approximation is the most commonly used in practical implementations of the BSE [87], but it is rather difficult to have insights on the physical picture it describes when formulated as such.

B. *Implementation*

In principle, the solution of the Dyson equation Eq. (3.126) in its spectral representation requires to invert the four-point response function $L(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_3,\omega)$ for each frequency ω [34]. Similarly to the way the one-body quasiparticle equation Eq. (3.89) was derived, the four-point susceptibility can be projected onto the molecular orbitals $\phi_\mu(\mathbf{r})$ in order to obtain a two-body quasiparticle equation [87]. The difference is that the quasiparticles are now two-body objects, and thus the excitonic Hamiltonian is expressed in the product basis $\{\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}')\}$, also called the *transition* basis. Any four-point function $F(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4)$ can be projected onto this basis as

$$F(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4) = \sum_{\mu\nu\lambda\xi} \phi^*_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_2)\phi_\lambda(\mathbf{r}_3)\phi^*_\xi(\mathbf{r}_4)F_{\mu\nu\lambda\xi}. \tag{3.127}$$

In particular, consider the non-interacting response function L_0 obtained from the Lehmann representation of G_1

$$L_0(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3,\mathbf{r}_4,\omega) = \sum_{\mu\nu}(f_\mu - f_\nu)\frac{\phi^*_\mu(\mathbf{r}_1)\phi_\nu(\mathbf{r}_2)\phi^*_\nu(\mathbf{r}_4)\phi_\mu(\mathbf{r}_3)}{\epsilon_\mu - \epsilon_\nu - \omega}, \tag{3.128}$$

where ϵ_n are the independent-quasiparticle energies i.e. the *GW* eigenvalues of Eq. (3.89). L_0 is readily diagonal in the transition basis and reads

$$L_0^{\mu\nu\lambda\xi}(\omega) = (f_\nu - f_\mu)\frac{\delta_{\mu\lambda}\delta_{\nu\xi}}{\epsilon_\nu - \epsilon_\mu - \omega}. \tag{3.129}$$

Then, by writing the Bethe-Salpeter equation as $L = \left(I - L_0\Xi\right)^{-1}L_0$, one can show that the response function becomes [87, 100]

$$L_{\mu\nu\lambda\xi}(\omega) = \left(H^{2p} - \omega I\right)^{-1}_{\mu\nu\lambda\xi}\delta_{\mu\lambda}\delta_{\nu\xi}(f_\nu - f_\mu), \tag{3.130}$$

where the *effective* two-particle Hamiltonian H^{2p} is defined as

$$H^{2p}_{\mu\nu\lambda\xi} = (\epsilon_\nu - \epsilon_\mu)\delta_{\mu\lambda}\delta_{\nu\xi} - (f_\nu - f_\mu)\Xi_{\mu\nu\lambda\xi}. \tag{3.131}$$

Because of the occupation factors $(f_\nu - f_\mu)$, the cases where $f_\mu = f_\nu$ or $f_\lambda = f_\xi$ (i.e. when any two indices of an element of the transition basis are both occupied or unoccupied) are irrelevant for the calculation of $L_{\mu\nu\lambda\xi}(\omega)$. Physically, this means that only the transitions between occupied and unoccupied states contribute to the response function. This formulation in terms of H^{2p} allows to cast the Bethe-Salpeter equation as an eigenvalue problem in the transition basis. Indeed, the inversion of the effective Hamiltonian leads to the following spectral decomposition of the response function

$$L_{\mu\nu\lambda\xi}(\omega) = \sum_n \frac{A^{\mu\nu}_n A^{*\lambda\xi}_n}{E_n - \omega}, \tag{3.132}$$

where E_n and $A^{\mu\nu}_n$ are respectively the eigenvalues and eigenvectors of H^{2p}. The latter are obtained by solving the generally non-Hermitian eigenvalue problem

34. Note that the $\delta(t_3 - t_4)$ in Eq. (3.126) makes the response function dependent on a single frequency.

[100] Bussi (2004).

$H^{2p} A_n = E_n A_n$, which can be stated in matrix form in the basis of occupied and empty states as [101]

$$\begin{pmatrix} R & C \\ -C^* & -R^* \end{pmatrix} \cdot \begin{pmatrix} X_n \\ Y_n \end{pmatrix} = E_n \begin{pmatrix} X_n \\ Y_n \end{pmatrix}, \tag{3.133}$$

where

$$R_{\mu\nu\lambda\xi} = \delta_{\mu\lambda}\delta_{\nu\xi}(\epsilon_\mu - \epsilon_\nu) + \Xi_{\mu\nu\lambda\xi} \tag{3.134a}$$

$$C_{\mu\nu\lambda\xi} = \Xi_{\mu\nu\xi\lambda}, \tag{3.134b}$$

and X_n, Y_n are the amplitudes of the electron-hole eigenstates, involved in the electron-hole wavefunction

$$\Psi_n(\mathbf{r}_e, \mathbf{r}_h) = \sum_{\mu\nu} \left(X_n^{\mu\nu} \phi_\mu(\mathbf{r}_h)\phi_\nu(\mathbf{r}_e) + Y_n^{\mu\nu} \phi_\mu(\mathbf{r}_e)\phi_\nu(\mathbf{r}_h) \right). \tag{3.135}$$

Here, R is the *resonant* block of H^{2p} where $f_\nu - f_\mu = f_\xi - f_\lambda = 1$, corresponding to transitions from occupied to empty states. Conversely, $-R^*$ is the *anti-resonant* block where $f_\nu - f_\mu = f_\xi - f_\lambda = -1$, corresponding to transitions from empty to occupied states. Finally, the *coupling* block C couples the resonant and anti-resonant parts, and neglecting it results in the Tamm-Dancoff approximation (TDA) [102, 103]. Note that the resonant block is Hermitian while the coupling block is symmetric, thus making the effective Hamiltonian non-Hermitian in general. It is particularly enlightening to specify the resonant part for singlet excitations as [101]

[102] Tamm (1991).

[103] Dancoff (1950).

$$\begin{aligned} R_{\mu\nu\lambda\xi} = \delta_{\mu\lambda}\delta_{\nu\xi}(\epsilon_\mu - \epsilon_\nu) + 2\int d\mathbf{r}\, d\mathbf{r}'\, \phi_\mu(\mathbf{r})\phi_\lambda(\mathbf{r})V(\mathbf{r},\mathbf{r}')\phi_\nu(\mathbf{r}')\phi_\xi(\mathbf{r}') \\ - \int d\mathbf{r}\, d\mathbf{r}'\, \phi_\mu(\mathbf{r})\phi_\lambda(\mathbf{r}')W(\mathbf{r},\mathbf{r}')\phi_\nu(\mathbf{r})\phi_\xi(\mathbf{r}'). \end{aligned} \tag{3.136}$$

Eq. (3.136) shows that the BSE approach corrects the *GW* gap $(\epsilon_\mu - \epsilon_\nu)$ with matrix elements $\langle 2V - W \rangle$ that account for the electron-hole binding energy. The latter has two contributions: the first is the electron-hole exchange while the second is the screened electron-hole interaction. Together, they lead to a reduction of the optical gap with respect to the photoemission gap. In fact, this is witnessed experimentally by the observation of optical absorption peaks within the photoemission gap measured e.g. by ARPES.

3.3.9 Application of the BSE to Charge-Transfer excitations

Throughout this Thesis, we will be specifically interested in charge-transfer (CT) excitations, whereby a quasiparticle undergoes a HOMO-LUMO [35] transition that coincides with a net charge transfer from one molecule to another. In the limit of a vanishing overlap between the HOMO and the LUMO the electron-hole exchange in Eq. (3.136) vanishes as well, hence the exciton binding energy is precisely the screened Coulomb potential between the electron and the hole W_{eh}, namely

$$V_{HL,HL} = \int d\mathbf{r}\, d\mathbf{r}'\, \phi_H(\mathbf{r})\phi_L(\mathbf{r})V(\mathbf{r},\mathbf{r}')\phi_H(\mathbf{r}')\phi_L(\mathbf{r}') = 0 \tag{3.137a}$$

$$W_{HH,LL} = \int d\mathbf{r}\, d\mathbf{r}'\, \phi_H(\mathbf{r})\phi_L(\mathbf{r}')W(\mathbf{r},\mathbf{r}')\phi_H(\mathbf{r})\phi_L(\mathbf{r}') = W_{eh}, \tag{3.137b}$$

since the co-densities $\phi_H(\mathbf{r})\phi_L(\mathbf{r})$ are identically zero for CT excitations.

35. Highest Occupied Molecular Orbital (HOMO), Lowest Unoccupied Molecular Orbital (LUMO).

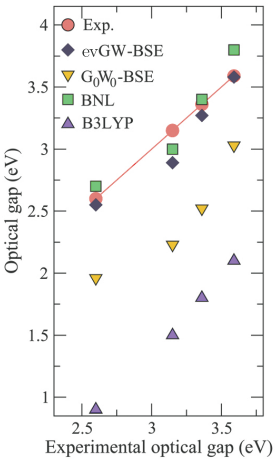

Figure 3.11: First singlet excitation energies of gas-phase donor-acceptor molecular complexes (benzene, naphthalene, and anthracene derivatives with the tetracyanoethylene acceptor).

For such excitations, standard TD-DFT with local kernels is known to fail [104], since the exchange-correlation kernel f_{xc} cannot couple electron and holes in the long-range. Again, only the integration of some amount of exact exchange can provide a long-range (even though unscreened) electron-hole interaction. For the sake of illustration we show in Fig. 3.11 the energy of charger-transfer excitations in a gas-phase donor-acceptor complex obtained by Blase and Attaccalite in the early days of the Fiesta initiative [85]. The standard TD-DFT calculations, such as TD-B3LYP, dramatically fail in predicting the correct CT energies, while both Bethe-Salpeter or a tuned range-separated hybrid functional are much more accurate.

In practice, any four-point quantity can be retrieved upon diagonalization of H^{2p}. In particular, the macroscopic dielectric function $\varepsilon_M(\omega)$, whose imaginary part yields the absorption spectrum, is given by

$$\varepsilon_M(\omega) = 1 - \lim_{\mathbf{q}\to 0} V(\mathbf{q}) \int d\mathbf{r}\, d\mathbf{r}'\, e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{r}')}\bar{L}(\mathbf{r},\mathbf{r},\mathbf{r}',\mathbf{r}',\omega), \tag{3.138}$$

where $\bar{L}$ verifies $\bar{L} = L + LV\bar{L}$ [36].

[104] Dreuw and Head-Gordon (2004).

36. See [87, 100] for more details about the relationship between microscopic and macroscopic functions.

[105] Loos et al. (2021).

I would like to briefly mention a recent article in which we compared the accuracy of the BSE formalism with respect to more advanced quantum chemistry methods [105]. In this joint work with P-F. Loos, X. Blase and D. Jacquemin, a series of highly accurate vertical transition energies for partial to complete intramolecular charge-transfer excitations were obtained for 17 (π-conjugated) molecular compounds. The reference theoretical best estimates were computed at the Coupled-Cluster Single-Double-Triple (CCSDT) level, and my contribution was to perform the BSE calculations. The error distribution of TD-DFT calculations with three different functionals (PBE0, CAM-B3LYP, LC-ωHPBE) with respect to the reference CCSDT calculations is shown in Fig. 3.12. While CAM-B3LYP performs very well, we observe a large variability of the accuracy of TD-DFT with respect to the exchange-correlation functional. This is an indication that the performances of a given functional are problem-specific, meaning that we cannot expect a given functional to perform better than others in all problems. In particular, the LC-ωHPBE functional - a tuned range-separated hybrid with 100% of exact exchange in the long-range - significantly overestimates transition energies.

The error distribution of BSE calculations starting with ev*GW*@PBE0 energy levels with respect to the CCSDT references are shown in Fig. 3.12 (bottom right). We

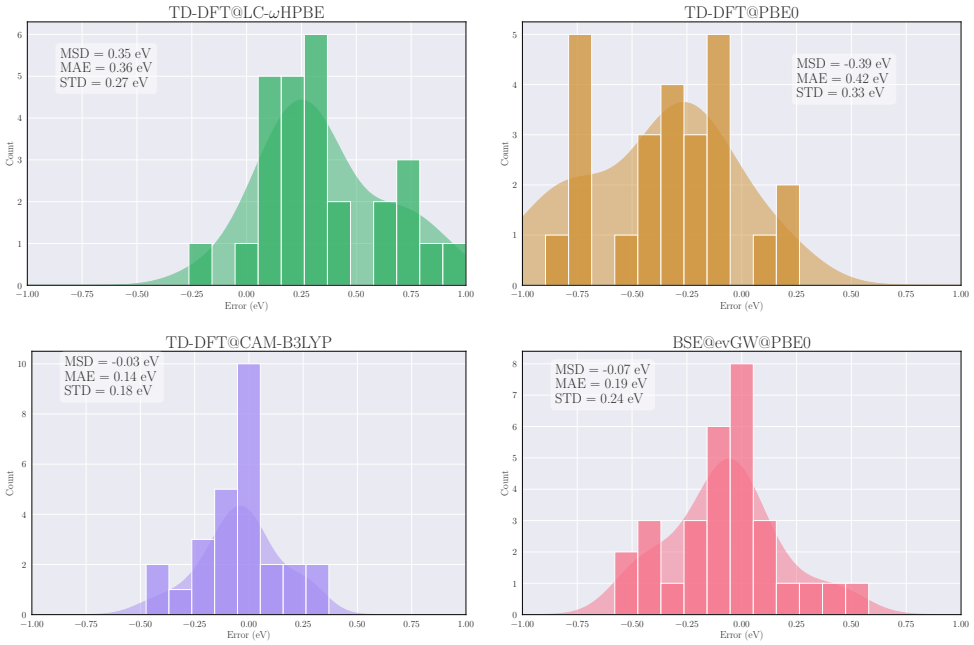

Figure 3.12: Error distribution of TD-DFT calculations for 30 optical transitions on 17 molecular compounds against the theoretical best estimate for three different exchange-correlation functionals. (top left) LC-ωHPBE, (top right) PBE0, (bottom left) CAM-B3LYP, (bottom right) BSE@ev*GW*@PBE0.

see that BSE@ev*GW*@PBE0 performs much better than TD-DFT@PBE0, with an accuracy similar to that of CAM-B3LYP. The robustness of the BSE formalism is a general feature, namely it performs similarly than a problem-tuned functional without having to find empirically the parameters that fit a particular system. Overall, the BSE approach allows for a consistent accuracy among a large variety of problems by lifting the need for empirical tuning of the parameters. This results in a more predictive theory, while keeping a favorable $\mathcal{O}(N^4)$ scaling, the same as TD-DFT. Such a robustness allows to explore novel systems and to make predictions on their optical excitations with more confidence.

3.4 ATOMISTIC ELECTROSTATIC MODELS: MICRO-ELECTROSTATICS

THE VARIOUS *ab initio* FORMALISMS we have discussed so far allowed us to tackle the many-body problem at various levels of theory. In the context of organic semiconductors, the bulk energy landscape has a strong impact (1-2 eV) on the charged and optical excitations. We now turn our attention towards a theoretical approach that allows to evaluate such environmental contributions by properly accounting for the long-range intermolecular interactions. The Micro-Electrostatics (ME) formalism relies on classical polarizable points for the description of electronic polarization, and was first proposed by Mott and Littleton for atomic lattices [45]. It permits on the one hand to compute the macroscopic dielectric properties of molecular systems and on the other hand to assess the environmental contributions to photoemission and absorption spectra, possibly by being integrated into *embedded* Quantum-Classical schemes (which will be discussed in Sec. 3.5).

All Micro-Electrostatics calculations in the present Thesis have been performed with either the MESCal or the Tequila in-house codes.

[45] Mott and Littleton (1938).

3.4.1 Micro-Electrostatics

The cohesion of non-covalent molecular solids is mostly due to the Van der Waals forces that arise from the interactions between instantaneous multipoles in molecules. Intermolecular interactions (0.02-1.3 eV) are typically much weaker than metallic (4-5 eV), ionic (7-9 eV) or covalent (1.5-9 eV) interactions which translates into low dielectric constants ($\varepsilon \sim 3$) and a weak intermolecular overlap. As a result, intermolecular interactions ($\sim$ eV) are a few orders of magnitude stronger than charge-transfer couplings ($\sim$1-100 meV) associated to the narrow electronic bands, and they act on a faster time scale (fs) than charge dynamics (ps), leading to typical shifts in charged excitations of a few eV [106, 107, 108].

[106] Tsiper and Soos (2001).

[107] D'Avino et al. (2014).

[108] D'Avino et al. (2016).

As before, we are interested in the one-particle charged excitations and two-particle neutral excitations, as these are indeed the relevant quasiparticles that are involved in charge transport in semiconductors. The effects of intermolecular interactions in the solid state on these energy levels can be quantified as the difference between their value in the solid and in the gas phase (isolated molecules)

$$\boxed{\Delta^+ = \mathrm{IP}^s - \mathrm{IP}^g} \tag{3.139a}$$

$$\boxed{\Delta^- = \mathrm{EA}^g - \mathrm{EA}^s,} \tag{3.139b}$$

where $\mathrm{IP}^g, \mathrm{EA}^g$ denote the energy levels in the gas phase and $\mathrm{IP}^s, \mathrm{EA}^s$ those in the solid state. Here, Δ^+ and Δ^- denote the *environmental* energies and encompass all the energetic shifts due to the interactions in the solid state[37]. Because of the notably long range of electrostatic intermolecular interactions determining the screening from molecular polarizabilities, large systems comprising $10^2 - 10^3$ molecules must be considered in order to reach convergence with respect to system size. This severely restricts the theoretical methods that can be applied. Fortunately, in the context of molecular systems the weak intermolecular overlap can be neglected to a good approximation[38], resulting in a ground state that is determined by classical electrostatic interactions between quantum objects [109]. Within this approximation, molecules only interact through classical Coulomb forces while intermolecular exchange and correlation effects are discarded. Thus, the total energy of a system of N non-overlapping molecules is a functional of the individual charge densities $\rho_n(\mathbf{r})$

37. The environmental energies were historically called polarization energies because the formation of induced dipoles was first recognized as the main source of stabilization of the energy levels in the solid state. Because of the variety of mechanisms involved in the context of intermolecular interactions, such a denomination is misleading. In order to avoid further confusion, we therefore adopt the term *environmental* energies.

38. Conversely, the overlap must be taken into account when describing electronic bands and charge transport processes. Perturbative approaches where intermolecular overlap is introduced a posteriori via hopping or band models are justified by the large energy difference between the two processes.

[109] Castet et al. (2008).

and the electrostatic potential felt by each molecule $V_n(\mathbf{r})$, namely

$$U[\rho] = \sum_{n=1}^{N} E[\rho_n(\mathbf{r}), V_n(\mathbf{r})] - \frac{1}{2}\int d\mathbf{r}\, \rho_n(\mathbf{r}) V_n(\mathbf{r}), \tag{3.140}$$

where $E[\rho_n(\mathbf{r}), V_n(\mathbf{r})]$ is the total energy of molecule n, and the second term in the right-hand side prevents the double counting of interactions[39]. Moreover, the electrostatic potential felt by molecule n reads

$$V_n(\mathbf{r}) = V^{ext}(\mathbf{r}) + \sum_{i\neq n}\int d\mathbf{r}'\, \frac{\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{3.141}$$

where $V^{ext}(\mathbf{r})$ is the external potential. The minimization of the total energy of the system with respect to the individual charge densities

$$\frac{\partial U[\rho]}{\partial \rho_n} = 0, \quad \forall n \tag{3.142}$$

leads to a self-consistent problem because of the interdependence of the potential and the charge densities in Eqs. (3.140)-(3.141).

39. When an external field is applied to the system, it does not appear in the double counting correction. However, one must integrate it to the total energy functional $E[\rho(\mathbf{r}), V(\mathbf{r})]$

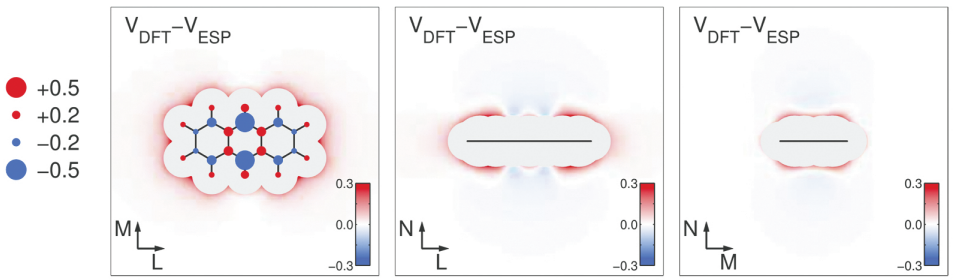

Figure 3.13: Difference between the potential of Anthracene (Volts) as calculated with DFT and as generated by ESP atomic charges, shown in the left panel, both obtained at the B3LYP/6-311++G** level. The gray area shows the projection of the Van der Waals surface in the plane. Adapted from [107].

Because of the weak intermolecular interactions, it is possible to split the molecular charge density $\rho(\mathbf{r})$ into the sum of the isolated gas-phase density $\rho^0(\mathbf{r})$ and the induced charge density $\delta\rho(\mathbf{r})$, namely

$$\rho(\mathbf{r}) = \rho^0(\mathbf{r}) + \delta\rho(\mathbf{r}), \tag{3.143}$$

where $\delta\rho(\mathbf{r})$ is assumed to be small. Then, within Micro-Electrostatics, the aforementioned variational problem is solved by providing discrete approximations of the gas-phase charge densities in terms of partial atomic charges

$$\rho_n(\mathbf{r}) = \sum_{i\in n} q_i \delta(\mathbf{r}-\mathbf{r}_i), \tag{3.144}$$

where q_i and $\mathbf{r}_i$ are the permanent charge and the nuclear position of atom i in the molecule. The set of charges $\{q_i\}$ can be obtained from electronic structure calculations in order to reproduce at best the molecular potential. Many fitting schemes are available in most quantum chemistry packages such the Löwdin [110], Mulliken [111], CHELPG [112] or ESP [113] partial atomic charges. As shown in Fig. 3.13, the electrostatic potential is well reproduced by this mapping, especially in the long range.

The induced charge density $\delta\rho(\mathbf{r})$ is instead described by distributing the molecular polarizability tensor $\underline{\boldsymbol{\alpha}}$ on atomic sites which express the molecular response to the electrostatic fields by the formation of induced dipoles $\{\boldsymbol{\mu}_n\}$. Again, many schemes

[110] Löwdin (1950).

[111] Mulliken (1955).

[112] Breneman and Wiberg (1990).

[113] Besler et al. (1990).

exist such as distributing the polarizability according to the relative atomic numbers or atomic polarizabilities. Empirically, adopting different schemes did not result in any sizable difference on the calculated quantities. We note that the number and location of both permanent charges $\{q_i\}$ and induced dipoles $\{\boldsymbol{\mu}_i\}$ pertaining to each molecule is arbitrary, but placing both at atomic positions has led to more accurate results and practical implementations [108]. Moreover, these quantities can be computed at the desired level of theory, e.g. DFT, thus making the ME scheme a multi-scale approach.

3.4.2 Self-consistent equations

Let us now delve into the derivation of the linear equations describing intermolecular interactions within the ME scheme. They will be presented in the simplified case of atomic lattices, whereby molecules are represented by a single site carrying the full molecular polarizability and net charge. The straightforward generalization to molecular systems with atomic resolution is given in paragraph C.

A. Derivation from Linear Response

Consider the system of N polarizable atomic sites at positions $\mathbf{r}_i$, each carrying a permanent charge q_i and an anisotropic polarizability tensor $\underline{\boldsymbol{\alpha}}_i$. The mutual electrostatic interactions between atomic sites will result in a net polarization. Within linear response, each polarizable site will develop induced dipoles $\boldsymbol{\mu}_i$ proportional to the total field felt by the site, namely

$$\boldsymbol{\mu} = \underline{\boldsymbol{\alpha}} \cdot \mathbf{F} \tag{3.145}$$

where $\boldsymbol{\mu}$ and $\mathbf{F}$ are respectively 3N-column vectors of the induced dipoles and the total electric field on atomic sites and $\underline{\boldsymbol{\alpha}}$ is a block-diagonal matrix consisting of N blocks of 3×3 polarizability tensors. The total dipole of the system also contains a contribution from the permanent dipole of each molecule. However, for molecules with inversion symmetry the permanent dipole must cancel, and we thus neglect the permanent dipoles in the following. The total electric field is the sum of the contributions of the permanent charges $\mathbf{F}^0$, the induced dipoles $\mathbf{F}^\mu$ and the applied external field $\mathbf{F}^{\text{ext}}$, namely

$$\mathbf{F} = \mathbf{F}^0 + \mathbf{F}^\mu + \mathbf{F}^{\text{ext}}, \tag{3.146}$$

The field on dipole i arising from all other permanents charges is

$$\mathbf{F}_i^0 = \sum_{j \neq i} q_j \frac{\mathbf{r}_{ij}}{r_{ij}^3}, \tag{3.147}$$

where $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$. The field on dipole i arising from all other induced dipoles is

$$\mathbf{F}^\mu = \underline{\mathbf{D}} \cdot \boldsymbol{\mu} \tag{3.148}$$

where $\underline{\mathbf{D}}$ is a $3N \times 3N$ matrix of dipole-dipole interactions consisting of N blocks of 3×3 matrices for the interaction between dipoles i and j

$$\underline{\mathbf{D}}_{ij} = (1 - \delta_{ij}) \frac{3\mathbf{r}_{ij} \otimes \mathbf{r}_{ij} - r_{ij}^2 \mathbb{I}}{r_{ij}^5}, \tag{3.149}$$

where $\otimes$ denotes a tensor product and $\mathbb{I}$ the 3×3 identity matrix. Inserting Eq. (3.148) into Eq. (3.145) leads to

$$\boldsymbol{\mu} = \underline{\boldsymbol{\alpha}} \cdot \left(\mathbf{F}^0 + \mathbf{F}^{\text{ext}} + \underline{\mathbf{D}} \cdot \boldsymbol{\mu} \right)$$

Which can be recast into a set of coupled linear equations by factorizing $\boldsymbol{\mu}$

$$\left(\underline{\boldsymbol{\alpha}}^{-1} - \underline{\mathbf{D}}\right) \cdot \boldsymbol{\mu} = \left(\mathbf{F}^0 + \mathbf{F}^{\text{ext}}\right) \tag{3.150}$$

This equation is precisely of the form

$$\underline{\mathbf{T}} \cdot \boldsymbol{\mu} = \mathbf{S} \tag{3.151}$$

where $\underline{\mathbf{T}} = \underline{\boldsymbol{\alpha}}^{-1} - \underline{\mathbf{D}}$ and $\mathbf{S} = \mathbf{F}^0 + \mathbf{F}^{\text{ext}}$. Indexing by ij the 3×3 sub-block corresponding to the interaction between dipoles i and j and attributing the indices kl to the spatial coordinates, the expression of the Interaction Matrix $\underline{\mathbf{T}}$ and the Source Vector $\mathbf{S}$ yields

$$\boxed{\begin{aligned} T_{ij}^{kl} &= \delta_{ij}\left(\underline{\boldsymbol{\alpha}}_i^{-1}\right)^{kl} + (1-\delta_{ij})\frac{r_{ij}\delta_{kl} - 3r_{ij}^k r_{ij}^l}{r_{ij}^5} && \text{(3.152a)} \\ S_i^k &= F_{ik}^{\text{ext}} + \sum_{j\neq i} q_j \frac{r_{ij}^k}{r_{ij}^3} && \text{(3.152b)} \end{aligned}}$$

Hence, the self-consistent equations of a system of N non-overlapping molecules interacting classically via the formation of induced dipoles is now expressed as the inversion of a linear system of 3N equations.

B. *Derivation from the Total Energy*

It is instructive to derive the linear system of Eq. (3.151) from total energy considerations. Let us consider again the system of N polarizable atomic sites at positions $\mathbf{r}_i$ each carrying a permanent charge q_i and an induced dipole $\boldsymbol{\mu}_i$. The total electric field at site i is given by Eq. (3.146). Likewise, the total potential at site i is the sum of the contributions of the permanent charges, the induced dipoles and the applied external potential

$$\mathbf{V} = \mathbf{V}^0 + \mathbf{V}^{\mu} + \mathbf{V}^{\text{ext}} \tag{3.153}$$

where $\mathbf{Q}$ and $\mathbf{V}$ are N-column vectors whose entries are respectively the charge q_i and potential V_i at each atomic site. Moreover, within linear response, the work necessary to adiabatically polarize the system by forming induced dipoles U_p is

$$U_p = \int_0^{\mathbf{F}} \boldsymbol{\mu} \cdot d\mathbf{F} = \frac{1}{2}\mathbf{F} \cdot \underline{\boldsymbol{\alpha}} \cdot \mathbf{F} = \frac{1}{2}\boldsymbol{\mu} \cdot \underline{\boldsymbol{\alpha}}^{-1} \cdot \boldsymbol{\mu} = \frac{1}{2}\boldsymbol{\mu} \cdot \mathbf{F}. \tag{3.154}$$

Therefore, the total energy of the system is given by the potential energy stored in the point charges and the induced dipoles plus the polarization work, namely

$$\boxed{U = \frac{1}{2}\mathbf{Q} \cdot \mathbf{V} + \frac{1}{2}\mathbf{Q} \cdot \mathbf{V}^{\text{ext}} - \frac{1}{2}\boldsymbol{\mu} \cdot \mathbf{F} - \frac{1}{2}\mu\mathbf{F}^{\text{ext}} + \frac{1}{2}\boldsymbol{\mu} \cdot \boldsymbol{\alpha}^{-1} \cdot \boldsymbol{\mu},} \tag{3.155}$$

where the $1/2$ factors arise from the double counting correction of the interactions, as in the second term of Eq. (3.140). The factor $1/2$ in front of the interaction energies with the external potential and field comes from our definitions of the total potentials and fields in Eq. (3.153) and Eq. (3.146). Besides, we note that in the absence of an external field, the use of the relation (3.154) results in a compact expression for the total energy

$$\boxed{U = \frac{1}{2}\mathbf{Q} \cdot \mathbf{V}.} \tag{3.156}$$

Expanding Eq. (3.155) in terms of sums over atomic sites, we find

In the case of an atomic lattice, the total energy can further be split in two separate contributions. By noticing from Eq. (3.159) that the charge-dipole and dipole-charge interactions are the negative of each other, i.e. $\mathbf{Q} \cdot \mathbf{V}^{\mu} = -\boldsymbol{\mu} \cdot \mathbf{F}^0$, the total energy is written as

$$U = \frac{1}{2}\mathbf{Q} \cdot \mathbf{V}^0 - \frac{1}{2}\boldsymbol{\mu} \cdot \mathbf{F}^0.$$

This expression highlights the two contributions to the total energy, namely

$$U_E = \frac{1}{2}\mathbf{Q} \cdot \mathbf{V}^0 \tag{3.157a}$$

$$U_I = -\frac{1}{2}\boldsymbol{\mu} \cdot \mathbf{F}^0. \tag{3.157b}$$

where E_E is the *electrostatic* energy and E_I the *induction* energy. This separation is not possible in the case of molecular lattices with atomic resolution, and the electrostatic and induction energies in the general case are discussed below.

$$U = \frac{1}{2}\sum_i q_i\left(V_i^0 + V_i^\mu\right) + \sum_i q_i V_i^{\text{ext}} - \frac{1}{2}\sum_i \boldsymbol{\mu}_i \cdot \left(\mathbf{F}_i^0 + \mathbf{F}_i^\mu\right) - \frac{1}{2}\sum_i \boldsymbol{\mu}_i \cdot \mathbf{F}_i^{\text{ext}} + \frac{1}{2}\sum_i \boldsymbol{\mu}_i \cdot \boldsymbol{\alpha}_i^{-1} \cdot \boldsymbol{\mu}_i. \tag{3.158}$$

In the above formula we recognize different factors corresponding to the interactions of the various components present in the system: the charge-charge, charge-dipole and charge-potential interactions and analogously the dipole-charge, dipole-dipole and dipole-field interactions. The last term corresponds to the elastic energy required to induce such dipoles, which is analogous to that of a harmonic oscillator where the force constant is $\boldsymbol{\alpha}^{-1}$. Finally, the full expression of the total potential energy in terms of the positions, charges, dipoles and polarizabilities of the system reads

$$U = \left[\frac{1}{2}\sum_{i\neq j}\left(\frac{q_i q_j}{r_{ij}} + q_i\frac{\boldsymbol{\mu}_j\cdot\mathbf{r}_{ij}}{r_{ij}^3} - q_j\frac{\boldsymbol{\mu}_i\cdot\mathbf{r}_{ij}}{r_{ij}^3} - \frac{3(\boldsymbol{\mu}_i\cdot\mathbf{r}_{ij})(\boldsymbol{\mu}_j\cdot\mathbf{r}_{ij}) - \boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j}{r_{ij}^5}\right) + \sum_i\left(q_i V_i^{\text{ext}} - \boldsymbol{\mu}_i\cdot\mathbf{F}_i^{\text{ext}}\right) + \frac{1}{2}\sum_i \boldsymbol{\mu}_i\cdot\underline{\boldsymbol{\alpha}}_i^{-1}\cdot\boldsymbol{\mu}_i\right]. \tag{3.159}$$

By differentiating the above expression with respect to $\boldsymbol{\mu}$ one recovers exactly the linear system of Eq. (3.151). Furthermore, we notice that U is a quadratic form in terms of $\boldsymbol{\mu}$ i.e. it has the form

$$U = U_0 + \nabla_{\boldsymbol{\mu}} U\cdot\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{\mu}\cdot\nabla_{\boldsymbol{\mu}}^2 U\cdot\boldsymbol{\mu}. \tag{3.160}$$

Then, the stationary condition $\nabla_{\boldsymbol{\mu}} U = 0$ corresponding to the minimization of the total energy with respect to the induced dipoles yields

$$\boxed{\nabla_{\boldsymbol{\mu}}^2 U\cdot\boldsymbol{\mu} = \nabla_{\boldsymbol{\mu}}\, U\Big|_{\boldsymbol{\mu}=0}}. \tag{3.161}$$

By identification[40] to Eq. (3.151), this last expression highlights that the Interaction Matrix $\underline{\mathbf{T}}$ is nothing else than the Hessian of U of the system with respect to the induced dipoles and the Source Vector $\mathbf{S}$ its gradient i.e.

$$\boxed{\begin{aligned}\underline{\mathbf{T}} &= \nabla_{\boldsymbol{\mu}}^2 U \qquad (3.162a)\\ \mathbf{S} &= \nabla_{\boldsymbol{\mu}}\, U\Big|_{\boldsymbol{\mu}=0}. \qquad (3.162b)\end{aligned}}$$

When the Hessian Matrix $\underline{\mathbf{T}}$ is positive-definite, i.e. $\nabla_{\boldsymbol{\mu}}^2 U > 0$, the system is in a global minimum of the total energy and is therefore stable. Whenever $\underline{\mathbf{T}}$ is not positive-definite, negative eigenvalues corresponding to unstable modes developing a dipole even in the absence of an external field (spontaneous polarization) appear. The advent of such modes will be covered in more detail in Chapter 5.

40. Performing explicitly the derivative with respect to $\boldsymbol{\mu}$ of Eq. (3.159) and setting it to 0 leads to (3.152).

C. *Generalization to the atomic resolution of molecular systems*

The adaptation of the ME equations to molecular systems with atomic resolution is rather straightforward. Indeed, in this context a molecule is an ensemble of atomic sites representing its electrostatic field and polarizability (by means of distributed atomic charges and polarizabilities). However, when two polarizable atoms are coming closer than a certain limit, instabilities appear because of a diverging polarizability. Various solutions were proposed to tackle this problem, the Thole model [114] being a popular one. We instead choose to fully retain intermolecular interactions and

[114] Thole (1981).

completely discard intramolecular interactions [107], a solution that has the advantage of taking into account the molecular identity and ensures the transferability of polarizability tensors. As such, atomic sites pertaining to the same molecule do not interact with each other, and the the dipole field tensor $\underline{\mathbf{D}}$ is modified according to

$$D_{ij}^{kl} = f_{ij}(1-\delta_{ij})\frac{r_{ij}\delta_{kl} - 3r_{ij}^k r_{ij}^l}{r_{ij}^5}, \tag{3.163}$$

where $f_{ij} = 0$ if atom i and j belong to the same molecule and $f_{ij} = 1$ otherwise.

3.4.3 Environmental Energies of Charged Excitations

In this Thesis, we will be particularly concerned with the the environmental contributions to the energy levels of charged excitations, namely the IP and EA, which can readily be evaluated in the framework of Micro-Electrostatics. Such environmental energies are differences of total energies between a neutral and a charged system where a probe molecule, i.e. the one for which energy levels are computed, is explicitly charged. The total energies that are involved are calculated at the ME level, only accounting for intermolecular interactions, which results in the environmental energies defined in Eq. (3.139). Owing to Eq. (3.156), the total energies of the neutral and charged systems can be written in terms of the permanent charges and induced dipoles as [41]

$$U^0 = \frac{1}{2}\sum_{mi} q^0(V^0 + \tilde{V}^0) \tag{3.164a}$$

$$U^{\pm} = \frac{1}{2}\sum_{mi} (q^0 + q^\delta)(V^0 + V^\delta + \tilde{V}^0 + \tilde{V}^\delta), \tag{3.164b}$$

41. Indices m, i are dropped for clarity, as they apply to all the quantities in Eq. (3.164).

where q^0_{mi} ($q^0_{mi} + q^\delta_{mi}$) is the permanent charge on atom i of molecule m in the neutral (charged) state, V^0_{mi} (V^δ_{mi}) is the *electrostatic* potential felt on atom i of molecule m from all other molecules in the neutral (charged) system, and $\tilde{V}^0_{mi}$ ($\tilde{V}^\delta_{mi}$) is the *induced* potential felt on atom i of molecule m from all other molecules for the neutral (charged) system. In other words, the differential charges and potentials between the charged and neutral systems are denoted with a δ superscript and the induced potentials as $\tilde{V}$. Owing to these definitions, an exact partitioning of the environmental energies can be done as

$$\boxed{\Delta = \Delta_E + \Delta_I,} \tag{3.165}$$

where Δ_E and Δ_I are called respectively the *electrostatic* and the *induction* contributions to the environmental energies. We show in this Appendix 3.4.5 that the two contributions can be expressed as

$$\boxed{\begin{aligned} \Delta_E &= \sum_{i\in\alpha} q^\delta_{\alpha i}\left(V^0_{\alpha i} + \tilde{V}^0_{\alpha i}\right) && (3.166a)\\ \Delta_I &= \frac{1}{2}\sum_{i\in\alpha} q^\delta_{\alpha i}\tilde{V}^\delta_{\alpha i}. && (3.166b) \end{aligned}}$$

The above result holds under the assumption that the polarizability of the charged and neutral molecules are the same, an approximation that is in most cases harmless [107]. This partitioning allows for a sound physical interpretation of the two contributions to the environmental energies, as well as providing a practical formula to compute them. We note that this partitioning is slightly different from that proposed in Ref. [107]. It also allows to compute separately the neutral and the charged systems, applying

the appropriate scaling laws to extrapolate the infinite system limit according to its dimensionality [108], see Sec. 3.4.4.

The electrostatic term corresponds to the energy shift arising from the average of the electrostatic potential in neutral system, including contributions from permanent and induced multipoles, weighted by the differential charge distribution of the molecular ion. Whenever hole and electron densities occupy the same region of space, one has $\Delta_E^+ = \Delta_E^-$, i.e. the electrostatic term acts as an approximately rigid shift of the frontier energy levels, barely affecting the gap.

The induction term corresponds instead to the energy shift due to the differential charge density of the charged molecule and the charge rearrangement this induces in the polarizable medium. The induction term has similar magnitude and opposite sign for positive (hole) and negative (electron) carriers, which has the effect of closing the transport gap. A rough estimate of Δ_I is given by the Born equation for the polarization energy of a charge in a spherical cavity of radius R inside a medium of relative permittivity ε

$$\Delta_I \approx \frac{1}{2R}\left(1 - \frac{1}{\varepsilon}\right). \tag{3.167}$$

By inserting typical values of molecular solids, $R = 5$ Å and $\varepsilon = 3.5$, we obtain $\Delta_I \approx -1eV$. Thus, we can expect the environmental energies to close the gap by approximately 2 eV.

3.4.4 Practical implementation

A. Periodic Boundary Conditions and Electrostatic Convergence

The calculation of the dielectric response of molecular solids in the context of Micro-Electrostatics requires to consider large systems in order to converge the electrostatic interactions at play. Moreover, because the environmental energy shifts are *per se* a macroscopic correction to the microscopic energy levels, it is desirable to calculate them in the infinite system limit as well. This is achieved by applying Periodic Boundary Conditions (PBC) to replicate the original unit cell. It is accounted for in the definitions of the Hessian matrix and the source vector of Eqs. (3.152) as

$$T_{ij}^{kl} = \delta_{ij}\left(\underline{\boldsymbol{\alpha}}_i^{-1}\right)^{kl} + (1-\delta_{ij})D_{ij}^{kl} + \sum_{\rho(j)} D_{i\rho(j)}^{kl} \tag{3.168a}$$

$$S_i^k = F_{ik}^{ext} + \sum_{J\neq i} q_J \frac{r_{iJ}^k}{r_{iJ}^3} \tag{3.168b}$$

where the index $\rho(j)$ spans over all periodic replica of j (excluding j itself), and the index J spans overs all sites and all periodic replica. Note that the diagonal of T_{ij}^{kl} now contains the dipole-dipole interaction between site i and its replica.

It is well-known that the evaluation of electrostatic interaction energies requires some care. The electrostatic sums that are involved are difficult to convergence as a result of the long range of Coulomb interactions. Moreover, common electrostatic sums (e.g. the potential in a crystal of quadrupolar molecules or the dipole fields in a polar crystal) only converge conditionally in 3D since their value depends on the order of summation and thus on the macroscopic shape of the system. This makes the electrostatic energy in the limit of an infinite sphere (*bulk* limit) and an infinite slab of arbitrary thickness (*film* limit) different in general. In neutral periodic systems, this

convergence can be achieved exactly by taking advantage of translational symmetry and performing reciprocal space summations such as Ewald's method [115, 116, 117]. The induction energy can also be evaluated exactly in periodic neutral systems, but when an aperiodic charged excitation breaks translational symmetry the induced dipoles have to be computed within a finite cutoff distance. In general, the treatment of disordered or charged systems requires to perform several calculations on finite samples of different sizes and extrapolate the values of Δ_E and Δ_I independently to the infinite system limit.

[115] Ewald (1921).

[116] Soos et al. (2001).

[117] Poelking and Andrienko (2016).

Because the charge-charge and charge-dipole interaction energies scale respectively as $1/r$ and $1/r^2$, the electrostatic and induction energies feature different system size scalings, presented in Table 3.2 as a function of the system's dimensionality and size. The electrostatic energy Δ_E scales as R^0 (i.e. oscillatory convergence) and R^{-1} in the bulk (3D) and film (2D) limits respectively. The induction energy Δ_I scales as R^{-1} and R^{-2} in the bulk and film limits respectively [108].

	Bulk	Film
Δ_E	R^0	R^{-1}
Δ_I	R^{-1}	R^{-2}

Table 3.2: Scaling with respect to the system dimensionality, e.g. 2D/3D for a cylinder/sphere, and radius R of the electrostatic and induction environmental energies for the bulk and film limits [108].

B. *Numerical resolution methods*

The self-consistent polarization state of N atoms is obtained by solving the system of 3N coupled linear equations of Eq. (3.151) for the induced dipoles $\boldsymbol{\mu}$. There are several ways to tackle this problem that are characterized by different stability and computational efficiency. One way this can be achieved is by inverting the Hessian matrix, namely

$$\boldsymbol{\mu} = \underline{\mathbf{T}}^{-1} \cdot \mathbf{S}. \tag{3.169}$$

The inversion is performed either by LU Factorization of the Hessian matrix or by Full Diagonalization. In the former case, the Hessian is decomposed as

$$\underline{\mathbf{T}} = \underline{\mathbf{L}} \cdot \underline{\mathbf{U}}, \tag{3.170}$$

where $\underline{\mathbf{L}}$ and $\underline{\mathbf{U}}$ are respectively lower and upper triangular matrices. The computation of $\boldsymbol{\mu}$ in Eq. (3.169) is then computed directly from this factorization, regardless of the spectral properties of the Hessian. In the case of Full Diagonalization, the Hessian is decomposed as

$$\underline{\mathbf{T}} = \underline{\mathbf{O}} \cdot \underline{\boldsymbol{\Lambda}} \cdot \underline{\mathbf{O}}^T, \tag{3.171}$$

where $\underline{\mathbf{O}}$ ($\underline{\boldsymbol{\Lambda}}$) is the matrix whose columns (diagonal elements) are the eigenvectors (eigenvalues) of $\underline{\mathbf{T}}$. From this factorization, the inversion is straightforward $\underline{\mathbf{T}}^{-1} = \underline{\mathbf{O}} \cdot \underline{\boldsymbol{\Lambda}}^{-1} \cdot \underline{\mathbf{O}}^T$, with the benefit of having access to the eigenvalues and eigenmodes of the Hessian that can provide further insights on the system's response. In particular, negative eigenvalues corresponding to unstable polarization modes can be detected and analyzed. The above methods of solving the ME equations are numerically expensive, with a scaling of $\mathcal{O}(N^3)$ in current implementations of the LU factorization and full diagonalization [42]. They feature however the advantage of solving the problem irrespectively to the presence of negative Hessian eigenvalues, i.e. of the arousal of physical instabilities to the polarization problem.

42. The LAPACK routines corresponding the LU Factorization and Full Diagonalization, featuring the beautiful names of `DSYSV` and `DSYEV` respectively, have a memory requirement $\mathcal{O}(N^2)$ since they need to store the $3N \times 3N$ Hessian Matrix.

Another strategy to evaluate the induced dipoles as solutions of the self-consistent equations is to follow an iterative scheme. In this case, the induced dipoles are initialized to a given value $\boldsymbol{\mu}_0$, typically taken to be $\underline{\boldsymbol{\alpha}} \cdot (\mathbf{F}^{ext} + \mathbf{F}^0)$. Then, the induced dipoles are updated by propagating the self-consistent equations as

$$\mathbf{F}^{\mu}_{n+1} = \underline{\mathbf{D}} \cdot \boldsymbol{\mu}_n \tag{3.172a}$$

$$\mathbf{F}_{n+1} = \mathbf{F}^{ext} + \mathbf{F}^0 + \mathbf{F}^{\mu}_{n+1} \tag{3.172b}$$

$$\boldsymbol{\mu}_{n+1} = \underline{\boldsymbol{\alpha}} \cdot \mathbf{F}_{n+1}. \tag{3.172c}$$

This scheme is iterated up to convergence of the total energy to a typical tolerance of 10^{-5} eV, that is typically reached in ~ 50 iterations for atomistic systems. Besides, it features a $\mathcal{O}(N^2)$ time complexity and less stringent memory requirements than matrix inversion methods as the Hessian matrix is never stored. This scheme is thus better suited for large systems. As for many other iterative algorithms, the convergence can be stabilized by introducing a damping parameter β in the updates of the dipoles

$$\boldsymbol{\mu}_{n+1} = \beta\boldsymbol{\mu}_n + (1-\beta)\underline{\boldsymbol{\alpha}} \cdot \mathbf{F}_{n+1}, \tag{3.173}$$

where $\beta \in [0,1]$. This strategy prevents the strong oscillations that are sometimes observed in the first iterations and allows for a smoother convergence, while keeping results independent of β. Many other optimization schemes exist, for example considering per-parameter damping factors, Direct Inversion of the Iterative Subspace (DIIS), or the Newton-Raphson procedure. Nevertheless, the convergence of such algorithms is not guaranteed, notably in systems with small or negative Hessian eigenvalues [118].

[118] Lipparini et al. (2014).

These resolution schemes have been implemented in two in-house codes. MESCal [107] is a code written in Fortran 90 that has been developed expressly for large molecular systems and adopts the iterative resolution method. It is thus well-suited for large systems e.g. when used in embedded calculations. The companion code Tequila [119] was developed by myself during this Thesis, it is written in C++ and has been thought specifically for lattices of polarizable points. The main need was to have an iteration-free method, able to reveal the intrinsic instabilities of dipole systems, since iterative solutions may have convergence issues even when all solutions are stable. To this aim, Tequila adopts a resolution scheme based on direct matrix inversion or full diagonalization. This allows to obtain the spectral decomposition of the polarization modes of the system under investigation, a feature that will be of central importance in Chapter 5.

[119] Comin (2022).

3.4.5 Appendix: Derivation of the Electrostatic and Induction partitioning

We show in this Appendix the expression of the *electrostatic* and *induction* contributions to the environmental energies stated in Eq. (3.165) and Eqs. (3.166). We begin by recalling that q^0_{mi} ($q^0_{mi} + q^\delta_{mi}$) is the permanent charge on atom i of molecule m in the neutral (charged) system, V^0_{mi} (V^δ_{mi}) is the *electrostatic* potential felt on atom i of molecule m from all other molecules in the neutral (charged) system, and $\tilde{V}^0_{mi}$ ($\tilde{V}^\delta_{mi}$) is the *induced* potential felt on atom i of molecule m from all other molecules for the neutral (charged) system. In other words, the differential charges and potentials between the charged and neutral systems are denoted with a δ superscript and the induced potentials as $\tilde{V}$. Similarly, if we let μ^0_{mi} (μ^δ_{mi}) be the induced dipole on atom i of molecule m for the neutral (charged) system, we can express the potentials as

$$V^0_{mi} = \sum_{nm,j} (1-\delta_{nm}) q^0_{nj} V(\mathbf{r}_{nj}, \mathbf{r}_{mi}) \tag{3.174a}$$

$$V^\delta_{mi} = \sum_{nm,j} (1-\delta_{nm}) q^\delta_{nj} V(\mathbf{r}_{nj}, \mathbf{r}_{mi}) = \sum_{j} q^\delta_{\alpha j} V(\mathbf{r}_{\alpha j}, \mathbf{r}_{mi}) \tag{3.174b}$$

$$\tilde{V}^0_{mi} = \sum_{nm,j} (1-\delta_{nm}) \boldsymbol{\mu}^0_{mi} \cdot \nabla V(\mathbf{r}_{nj}, \mathbf{r}_{mi}) \tag{3.174c}$$

$$\tilde{V}^\delta_{mi} = \sum_{nm,j} (1-\delta_{nm}) \boldsymbol{\mu}^\delta_{mi} \cdot \nabla V(\mathbf{r}_{nj}, \mathbf{r}_{mi}), \tag{3.174d}$$

where $V(\mathbf{r}_{nj}, \mathbf{r}_{mi})$ is the Coulomb potential between atom j of molecule n on atom i of molecule m, and molecule α is the probe that is possibly charged. Notice that the summations $\sum_{n\neq m}$ have been reformulated as $\sum_{nm}(1-\delta_{nm})$.

In order to prove Eqs. (3.166), we will need to prove the following related properties.

Property 1:

$$\sum_{mi} q^0_{mi} V^\delta_{mi} = \sum_{mi} q^\delta_{mi} V^0_{mi} \tag{3.175}$$

Proof:

$$\sum_{mi} q^0_{mi} V^\delta_{mi} = \sum_{mi,nj} q^0_{mi} q^\delta_{nj} V(\mathbf{r}_{nj}, \mathbf{r}_{mi}) = \sum_{nj} q^\delta_{nj} V^0_{nj} \tag{3.176}$$

Property 2:

$$\sum_{mi} q^0_{mi} \tilde{V}^\delta_{mi} = \sum_{mi} q^\delta_{mi} \tilde{V}^0_{mi} \tag{3.177}$$

Proof:
First notice from Eqs. (3.174) that $\tilde{V}$ depends linearly on $\boldsymbol{\mu}$. On the other hand, the linear response equation for the induced dipoles in Eq. (3.145) yields

$$\boldsymbol{\mu} = -\underline{\boldsymbol{\alpha}} \cdot \nabla \cdot \Big(V + \tilde{V} \Big), \tag{3.178}$$

i.e. the induced dipoles depend linearly on V and $\tilde{V}$. This results in a linear recursive relation for $\tilde{V}$ so that it can be expressed as a linear functional of V, namely

$$\tilde{V} = \mathcal{L}[V]. \tag{3.179}$$

Taking advantage of the bilinearity of $\mathcal{L}$ and using Property 1 then gives

$$\begin{aligned} \sum_{mi} q^0_{mi} \tilde{V}^\delta_{mi} &= \sum q^0 \mathcal{L}[V^\delta] = \sum \mathcal{L}[q^0 V^\delta] = \sum \mathcal{L}[q^\delta V^0] = \sum q^\delta \mathcal{L}[V^0] \\ &= \sum_{mi} q^\delta_{mi} \tilde{V}^0_{mi}. \end{aligned} \tag{3.180}$$

We now proceed to prove Eqs. (3.166) by first noticing that $q^\delta_{mi} = 0 \quad \forall m \neq \alpha$ since we assume that only the probe molecule α can have a different charge. Moreover, we have that $V^\delta_{\alpha i} = 0$ since the charged molecule does not feel the potential due to its own charging. Going back to Eqs. (3.164) for the total energies of the neutral and charged systems, we can now write (mi indices are dropped, any change will be made explicit)

$$U^\pm = \frac{1}{2} \sum_{mi} \left(q^0 + q^\delta\right)\left(V^0 + V^\delta + \tilde{V}^0 + \tilde{V}^\delta\right) \tag{3.181a}$$

$$= \frac{1}{2} \sum_{mi} q^0 \left(V^0 + \tilde{V}^0\right) + q^\delta \left(2V^0 + 2\tilde{V}^0 + V^\delta + \tilde{V}^\delta\right) \tag{3.181b}$$

$$= U^0 + \frac{1}{2} \sum_{mi} q^\delta \left(2V^0 + 2\tilde{\phi}^0 + V^\delta + \tilde{V}^\delta\right). \tag{3.181c}$$

Therefore, their difference reads

$$\Delta^{\pm} = U^{\pm} - U^{0} = \frac{1}{2}\sum_{mi} q^{\delta}\left(V^{\delta} + \tilde{V}^{\delta}\right) + \sum_{mi} q^{\delta}\left(V^{0} + \tilde{V}^{0}\right) \tag{3.182a}$$

$$= \frac{1}{2}\sum_{i} q^{\delta}_{1i}\left(V^{\delta}_{1i} + \tilde{V}^{\delta}_{1i}\right) + \sum_{i} q^{\delta}_{1i}\left(V^{0}_{1i} + \tilde{V}^{0}_{1i}\right) \tag{3.182b}$$

$$= \frac{1}{2}\sum_{i} q^{\delta}_{1i}\tilde{V}^{\delta}_{1i} + \Delta_{E} \tag{3.182c}$$

$$= \Delta_{I} + \Delta_{E}, \tag{3.182d}$$

which proves the statement.

3.5 Embedded DFT & MBPT

The achievement of quantitative accuracy on the photoemission energy levels and optical excitations in the solid state is challenging task, in particular because of the complex effect of long-range intermolecular interactions. As discussed in the previous section, in the context of molecular systems the macroscopic screening effects arising from electrostatic and dielectric interactions in the solid state strongly impact the quasiparticle levels (1-2 eV) [108]. Moreover, while the IP and EA are well-defined properties of an isolated molecule, these quantities can present variations that can exceed 0.6 eV between different solid samples of the same compound [120, 121, 122], and intermolecular interactions in different host semiconductors are found to lower the EA of a dopant impurity by more than 1 eV with respect to the pure dopant crystal [38].

[120] Duhm et al. (2008).

[121] Salzmann et al. (2008).

[122] D'Avino et al. (2020).

[38] Li et al. (2019).

These observations call for a theoretical treatment that captures both the accurate quantum-mechanical nature of quasiparticle excitations and the bulk screening effects. While periodic bulk *GW* and BSE calculations have proven their ability in describing pristine organic semiconductors [123, 124, 125, 126], they cannot easily attain the absolute values of charged excitations due to the missing internal reference for the vacuum level. Moreover, these approaches cannot properly describe systems lacking periodicity that are disordered or chemically inhomogeneous systems, such a doped OSCs. The heavy computational cost of periodic *GW* and BSE calculations with large supercells requires the consideration of alternative approaches for the study of large and chemically or energetically disordered solids.

[123] Tiago et al. (2003).

[124] Hummer et al. (2004).

[125] Sharifzadeh et al. (2012).

[126] Cudazzo et al. (2012).

First introduced in 1976 by Warshel and Levitt [127] (2013 Chemistry Nobel Prize) in the context of complex biological systems [128], Hybrid Quantum Mechanics / Molecular Mechanics (QM/MM) approaches describe the structural and electronic properties of extended molecular systems by partitioning the system under investigation into regions that are assigned different levels of theory. The region that is described at a quantum-mechanical level, the QM part, is embedded in a larger region that is treated at a coarser level of theory such as molecular mechanics, the MM part. Such multiscale approaches are nowadays a standard computational tool for describing organic electronics [108] as they are especially well-suited to the theoretical study of supramolecular systems featuring disorder and heterogeneous chemical composition.

[127] Karplus et al. (2013).

[128] Warshel and Levitt (1976).

3.5.1 Embedded Density-Functional Theory

In this section, we introduce embedded Density-Functional Theory by discussing the partitioning of the system within this multiscale approach, and then derive the general formula for the total energy of the QM/MM system. This allows to account for electrostatic environmental shifts when applied to a neutral system and induction environmental shifts when differences of total energies between charged and neutral systems are considered, resulting in an evaluation of ionization potentials and electronic affinities.

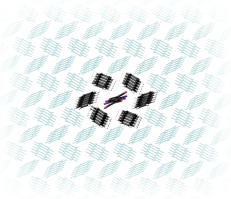

Figure 3.14: Illustration of the embedded DFT procedure: the ground state of the QM region (black) is described at the DFT level, while the electrostatic response of the MM region (cyan) is described at the ME level.

Let us consider a system in which we want only a subregion to be described quantum-mechanically, whereas the remaining is modeled at a coarser level of theory, as shown in Fig. 3.14. We label the former subregion the QM part and the latter the MM part of the system. For the purpose of this section, we are interested in the case where the QM

region is described at the DFT level, while the MM region is described at the Micro-Electrostatic level. Such a division is made possible by assuming that the molecular orbitals weakly overlap between the regions but not within each region. Owing to the forced orthogonality of the QM and MM systems, the general expression for the total energy of a collection of non-overlapping molecules given in Eq. (3.140) holds true for the QM/MM system as well. In this context, the QM region can be interpreted as a single molecular complex, since it is not separable into non-overlapping units. The total energy then yields

$$U^{QM/MM} = \sum_n E[\rho_n(\mathbf{r}), V_n(\mathbf{r})] - \frac{1}{2}\int d\mathbf{r}\, \rho_n(\mathbf{r}) V_n(\mathbf{r}), \tag{3.183}$$

where $E[\rho_n(\mathbf{r}), V_n(\mathbf{r})]$ denotes the total energy functional of molecule n with density $\rho_n(\mathbf{r})$, and $V_n(\mathbf{r})$ is the field *felt* by molecule n from the other, namely

$$V_n(\mathbf{r}) = \sum_{m\neq n} \int d\mathbf{r}'\, \frac{\rho_n(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.184}$$

The last term in the right-hand side of Eq. (3.183) rectifies the double counting of the interactions in the total energy since each molecule feels the potential due to all the others. We stress that the above formula is valid irrespective on the level of theory employed for the evaluation of the individual energies $E[\rho, V]$.

We now assume the MM region to be described at the Micro-Electrostatic level of theory, where the total energy of the MM region is obtained as the sum of the individual gas-phase energies of its constituents $E[\rho_n(\mathbf{r}), 0]$ plus the total electrostatic energy arising from their mutual interactions. Hence, the total energy of the MM region, in the absence of external fields, is obtained from the total electrostatic energy from Eq. (3.140) as

$$U^{MM} = \sum_{n\in MM} E[\rho_n(\mathbf{r}), 0] + \frac{1}{2}\mathbf{Q}\cdot\mathbf{V}, \tag{3.185}$$

where $\mathbf{Q}$ and $\mathbf{V}$ are respectively the ME permanent charges and the potentials at atomic sites, accounting for all molecules in the system, including the QM region. Therefore, the total energy of the QM/MM system becomes

$$\boxed{U^{QM/MM} = E[\rho^{QM}(\mathbf{r}), V^{QM}(\mathbf{r})] + \sum_{n\in MM} E[\rho_n(\mathbf{r}), 0] + \frac{1}{2}\mathbf{Q}\cdot\mathbf{V} - \frac{1}{2}\int d\mathbf{r}\, \rho^{QM}(\mathbf{r}) V^{MM}(\mathbf{r}).} \tag{3.186}$$

The first two terms of Eq. (3.186) are readily available from the outcome a DFT/ME calculation. The last integral corresponds to the electrostatic energy of the QM subsystem in the field of the MM subsystem. In practice, it is obtained within DFT as the following difference:

$$\int d\mathbf{r}\, \rho^{QM}(\mathbf{r}) V^{MM}(\mathbf{r}) = E_0[\rho^{QM}(\mathbf{r}), V^{MM}(\mathbf{r})] - E_0[\rho^{QM}(\mathbf{r}), 0]. \tag{3.187}$$

Here, $E_0[\rho^{QM}(\mathbf{r}), 0]$ is the total energy of the QM region in the gas phase and $E_0[\rho^{QM}(\mathbf{r}), V^{MM}(\mathbf{r})]$ is the same total energy functional evaluated in the self-consistent field of the MM region. Notice that $\rho^{QM}(\mathbf{r})$ is the electron density that minimizes the total energy functional in the gas phase, so that the first term in the right-hand side of the above equation is not minimal.

The total energy functional $E[\rho, V]$ can account for the nuclear repulsion energies within each molecule and between them. However, in most cases one is interested

in differences of total energies of QM/MM systems, such as in the calculation of Ionization Energies or Electron Affinities. When the ionic positions are unchanged between the neutral and ionized system, the IP and EA are calculated within the frozen core approximation (i.e. neglecting relaxation effects), resulting in the calculation of direct transitions. In this case, the nuclear contributions to the total energy of the QM/MM system cancel exactly, as is the case for the individual gas-phase energies of the molecules in the MM region.

Moreover, the total energy functional can also account for dispersion interactions within each molecule, such as those provided by the empirical D3 corrections in DFT, and between each molecule via empirical Lennard-Jones interactions. When considering differences of total energy, the Lennard-Jones energy within the MM region cancels exactly, since the interaction depends solely on nuclear positions. However, the inclusion of such interactions affects the ground state of the QM region, and their inclusion is thus not negligible.

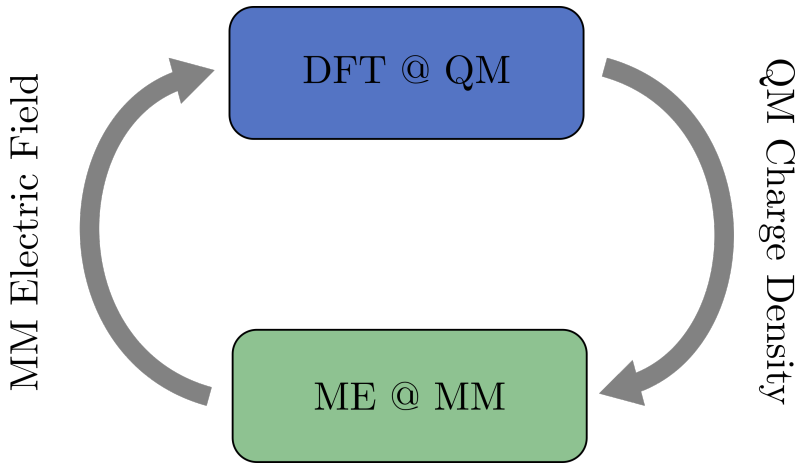

Figure 3.15: Self-consistency flowchart of a DFT/ME calculation.

In practice, a DFT/ME calculation proceeds by a self-consistent cycle between the QM and MM subsystems, as illustrated in Fig. 3.15. The QM gas-phase density is obtained from a first DFT calculation, and the electrostatic potential arising from it is directly calculated at the position of each MM atom. The potential of the QM region is then taken as an external field in the ME calculation. Afterward, the self-consistent ME calculation for the MM subsystem provides the set of induced dipoles that characterize its electrostatic response. The latter are injected back into the DFT calculation by describing each dipole as two point charges with a small, fixed separation (e.g. 10^{-2} Å) closely reproducing the dipole field. This procedure is then repeated in a self-consistent fashion till convergence of the total energy of both the QM and MM subsystems. Typically, 5 iterations between the two regions are enough to convergence the total energy up to a few meV.

3.5.2 Embedded MBPT

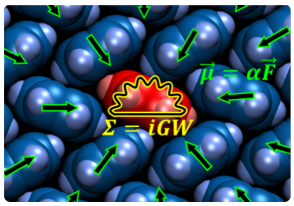

Figure 3.16: Illustration of the embedded *GW* procedure: the charged excitations in the QM region (red) are described at the *GW* level (yellow), while the induced dipoles in the MM region (blue) are described at the ME level (green).

We now present the merging of the many-body Green's function *GW* and BSE formalisms with the classical polarizable model ME approach. These hybrid approaches provide an effective way to compute the full spectrum of charged and neutral excitations in complex molecular systems, retaining an accurate quantum description for the subsystem of interest while fully accounting for the electrostatic interactions with the polarizable environment. While the effects of the electrostatic field in the neutral ground state are accounted for in the embedded DFT calculation, the main concern here is to describe the screening of charged and neutral excitations within the QM region due to the dielectric response on the MM environment.

A. The Reaction Field Matrix

Let us assume again a system that can be divided into a QM and an MM region, such that the molecular overlap between the two can be neglected. The QM region is now described by Many-Body Perturbation Theory, while the MM region is still described by Micro-Electrostatics, as illustrated in Fig. 3.16. The neglect of the overlap between the QM and MM regions allows to assess the contribution of the dielectric response of the MM region to the *GW* addition and removal energies. As a direct consequence, the susceptibility defined in Eq. (3.106) is zero whenever a pair of positions $\mathbf{r}, \mathbf{r}'$ are in the QM and MM system respectively. In other words, χ_0 adopts a block-diagonal form with two blocks coinciding with the QM and MM susceptibility respectively [43]. By denoting the QM and MM blocks by indexes 1 and 2 respectively, the equation for screened Coulomb potential W of the QM subsystem reads [129, 130, 131]

$$W_{11} = V_{11} + V_{11} \cdot \chi^0_{11} \cdot W_{11} + V_{12} \cdot \chi^0_{22} \cdot W_{21} \tag{3.188a}$$
$$W_{21} = V_{21} + V_{21} \cdot \chi^0_{11} \cdot W_{11} + V_{12} \cdot \chi^0_{22} \cdot W_{21}. \tag{3.188b}$$

Let us further define the Coulomb potential screened by the MM subsystem only as

$$\widetilde{V}_{11} = V_{11} + V_{12} \cdot \chi^*_{22} \cdot V_{21} \tag{3.189a}$$
$$\chi^*_{22} = \chi^0_{22} + \chi^0_{22} \cdot V_{22} \cdot \chi^*_{22}, \tag{3.189b}$$

where χ^*_{22} is the interacting polarizability of the MM subsystem in the absence of the QM subsystem [44]. These definitions lead to the expression of the screened Coulomb potential of the QM region in terms of the renormalized bare Coulomb potential $\widetilde{V}$, namely

$$W_{11} = \widetilde{V}_{11} + \widetilde{V}_{11} \cdot \chi^0_{11} \cdot W_{11}. \tag{3.190}$$

This is the same formulation as the one we would get for a gas-phase QM system, but with a bare Coulomb potential dressed by the reaction field originating from the MM environment. The MM-screened Coulomb potential can be written as

$$\widetilde{V}_{11} = V_{11} + V^{\text{reac}}_{11}, \tag{3.191}$$

where $V^{\text{reac}}_{11} = V_{12} \cdot \chi^*_{22} \cdot V_{21}$ is the *reaction field*. Since the interacting polarizability relates the variation of the charge density as a response to an applied perturbation, the quantity

$$\Delta\rho(\mathbf{r}_2, \omega) = \int \mathrm{d}\mathbf{r}'_2 \, \chi^*(\mathbf{r}_2, \mathbf{r}'_2, \omega) V(\mathbf{r}_1, \mathbf{r}'_2) \tag{3.192}$$

describes the variation of the charge density in $\mathbf{r}_2 \in$ MM by a unit test charge in $\mathbf{r}_1 \in$ QM. Thus, the Reaction Field Matrix

$$V^{\text{reac}}(\mathbf{r}_1, \mathbf{r}'_1, \omega) = \int \mathrm{d}\mathbf{r}_2 \, \mathrm{d}\mathbf{r}'_2 \, V(\mathbf{r}'_1, \mathbf{r}_2) \chi^*(\mathbf{r}_2, \mathbf{r}'_2, \omega) V(\mathbf{r}_1, \mathbf{r}'_2) \tag{3.193}$$

43. This formulation was first developed in the case of a Polarizable Continuum Model (PCM) for the MM subsystem [129]. This derivation is however general and assumes the MM subsystem to be described at the ME level.

[130] Li et al. (2016).

[131] Li et al. (2018).

44. This interacting polarizability should not be confused with the restriction of χ to the MM subsystem.

is the reaction potential on $\mathbf{r}_1' \in$ QM due to the charge density $\Delta\rho(\mathbf{r}_2) \in$ MM induced by a test charge in $\mathbf{r}_1 \in$ QM. In this way, the dielectric response of the MM subsystem are captured within the *GW* framework simply replacing the bare Coulomb potential by the MM-screened Coulomb potential of Eq. (3.191). The latter is thus the key quantity through which molecular polarizabilities of the MM region screen the quasiparticles in the QM region.

The construction of the self-energy requires the knowledge of the frequency-dependent reaction field of Eq. (3.193) as well as the dynamically screened Coulomb potential $W(\omega)$, accounting for the frequency dependence of the dielectric function in the optical range. Although such an energy dependence constitutes an important feature of the *GW* self-energy in extended solids or organic systems [24], Micro-Electrostatic models rely on the low frequency limit $\omega \to 0$ of the dielectric response[45]. The static formulation of the *GW* formalism corresponds to the COHSEX approximation discussed in Sec. 3.3.7. Here, the COHSEX approximation is adopted only for the contribution of the MM environment to the quasiparticle excitations by approximating the QM/MM self-energy operator as

[24] Dressel et al. (2008).

45. This limitation can in principle be overcome, although it has not been done in practice.

$$\begin{aligned}\Sigma^{\mathrm{GW/MM}} &= \Sigma^{\mathrm{GW}} + \left(\Sigma^{\mathrm{GW/MM}} - \Sigma^{\mathrm{GW}}\right) \\ &\simeq \Sigma^{\mathrm{GW}} + \left(\Sigma^{\mathrm{COHSEX/MM}} - \Sigma^{\mathrm{COHSEX}}\right),\end{aligned} \tag{3.194}$$

where $\Sigma^{\mathrm{COHSEX/MM}}$ and Σ^{COHSEX} are respectively the self-energy constructed with and without the inclusion of V^{reac}, both at the COHSEX level. Hence, while dynamic screening effects are taken into account in the QM region within the full *GW* framework, the contribution of the environment to the quasiparticle energies is evaluated in the static limit. The underlying assumption is that the contribution of the dynamical screening largely cancels out in the difference between COHSEX self-energies (ΔCOHSEX). The approximated *GW*/MM quasiparticle energies are then

$$E_n^{\mathrm{GW/MM}} = E_n^{\mathrm{GW}} + \Delta_n^{\mathrm{COHSEX}}, \tag{3.195}$$

where $\Delta_n^{\mathrm{COHSEX}}$ is the state-specific induction energy accounting for the screening provided by the induced dipoles in the MM region. In this formulation, only the induced multipoles in the MM region contribute to the quasiparticle energies.

We note that within the Gaussian basis implementation of the *GW* approximation, V^{reac} is not computed on a real-space grid but rather on a basis of auxiliary functions $\beta(\mathbf{r})$ describing co-densities within the RI formalism (see Sec. 3.3.6). These are atom centered Gaussian functions, leading to

$$V^{\mathrm{reac}}_{\beta\beta'} = \int \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,\beta(\mathbf{r})V^{\mathrm{reac}}(\mathbf{r},\mathbf{r}')\beta(\mathbf{r}'), \tag{3.196}$$

where $V^{\mathrm{reac}}_{\beta\beta'}$ is the Reaction Field Matrix (RFM). In practice, the dipoles induced in the MM subsystem at $\mathbf{r}' \in$ MM by the potential of the charge density associated to each orbital β' in the auxiliary basis are calculated as

$$V_{\beta'}(\mathbf{r}') = \int \mathrm{d}\mathbf{r}\,\frac{\beta'(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.197}$$

Then, these induced dipoles generate in return the reaction potential $V^{\mathrm{reac}}(\mathbf{r})$ on the QM subsystem. The energy of the probe auxiliary orbital β in the field of the MM subsystem polarized by the source charge density in β' is finally given by

$$V^{\mathrm{reac}}_{\beta\beta'} = \int \mathrm{d}\mathbf{r}'\,V^{\mathrm{reac}}_{\beta'}(\mathbf{r}')\beta(\mathbf{r}). \tag{3.198}$$

Such a self-consistent calculation of the induced polarization in the MM subsystem for each auxiliary orbital is performed before the *GW* calculation. For each element of the auxiliary basis, the response of the MM is calculated and the reaction potential on all basis elements is obtained. The computation of the RFM scales as $\mathcal{O}\left(N_{MM}^2 N_{QM}\right)$ where N_{QM} and N_{MM} are the number of atoms in the QM and MM subsystems respectively [131].

The calculation of the Reaction Field Matrix is done in practice on finite-size systems. As discussed in Sec. 3.4.4, the environmental energies can be extrapolated by performing several calculations of Δ_n^{COHSEX} with different sizes of MM system in order to reach the limit of an infinite system. For instance, the induction energies Δ_n^{COHSEX} for spherical MM clusters of increasing radius R scale as $1/R$ [131]. However, this strategy necessitates to perform several COHSEX calculations. Another strategy is to extrapolate the matrix elements $V^{\text{reac}}_{\beta\beta'}$ beforehand and then performing a single COHSEX calculation that directly targets the infinite crystal limit. In practice, the RFM matrix elements β, β' corresponding to orbitals with angular momenta l, l' scale as $R^{-1+l+l'}$, thus allowing a straightforward extrapolation [46] [131]. In particular, such a scaling can be very fast for high angular momenta resulting in converged values in clusters of relatively small size.

Analogously to the embedding of a *GW* calculation, the embedding of a Bethe-Salpeter calculation consists solely in adding the RFM to the bare Coulomb potential. Since the Bethe-Salpeter kernel is already approximated in the static limit, as discussed previously in Sec. 3.3.8, the static RFM can be used as such [101, 132]. Such corrections to the gas-phase optical levels typically result in a red-shift of the excitation energies that originates from the screening provided by the polarizable MM subsystem. Such a stabilization is expected also to reduce the electron-hole binding energy, which for pure Charge-Transfer excitations coincides with the screened Coulomb potential between the electron and the hole W_{eh}, a feature that may be of importance in the context of doping or organic photovoltaics.

46. This dependence on the radius of the spherical cluster stems from the fact that the auxiliary basis functions $\{\beta\}$ describe charge densities corresponding to electrical multipoles. Under the assumption that the size of the QM subsystem is small with respect to R, such multipoles scale as $R^{-1+l+l'}$.

[132] Duchemin et al. (2018).

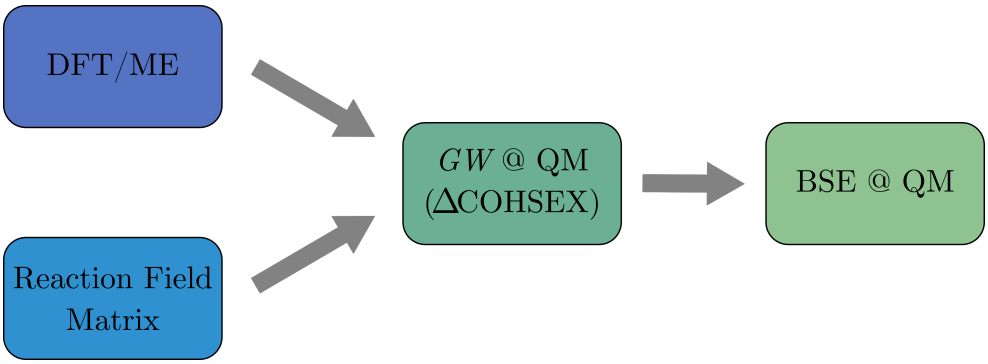

Figure 3.17: Flowchart of Embedded *GW*/BSE calculations.

[133] Siegrist et al. (2001).

[20] Troisi and Orlandi (2005).

[22] Hatch et al. (2009).

[134] Sakanoue and Sirringhaus (2010).

[12] Salzmann et al. (2012).

[29] Kleemann et al. (2012).

[28] Harada et al. (2010).

[21] Doi et al. (2005).

B. *Application to the charged excitations in the Pentacene crystal*

As an illustrative example of the *GW*/ME approach, we discuss here the different contributions from intermolecular interactions to the quasiparticle levels of the Pentacene crystal highlighting the role of the different contributions, from the work of Li et al. [131]. This system has been extensively studied experimentally [133, 20, 22, 134, 12, 29, 28] and theoretically [21] which makes it an ideal case study. The evolution of the IP and EA from the isolated gas phase to the crystal phase is depicted in Fig. 3.18, where the effects of the environment are progressively added from left to right. First, we observe that the calculated ev*GW* gas-phase levels of a single Pentacene molecule

are in excellent agreement with the associated experimental data.

The effect of the environment to the neutral ground state of the system is accounted for in the embedded DFT/ME calculation, and results in an approximately rigid shift of the quasiparticle levels. This contribution from the crystal field corresponds to the electrostatic environmental energy shift Δ_E, discussed in Sec. 3.4.3 in the context of ME. It is physically equivalent to the charge-quadrupole interaction in classical electrostatics.

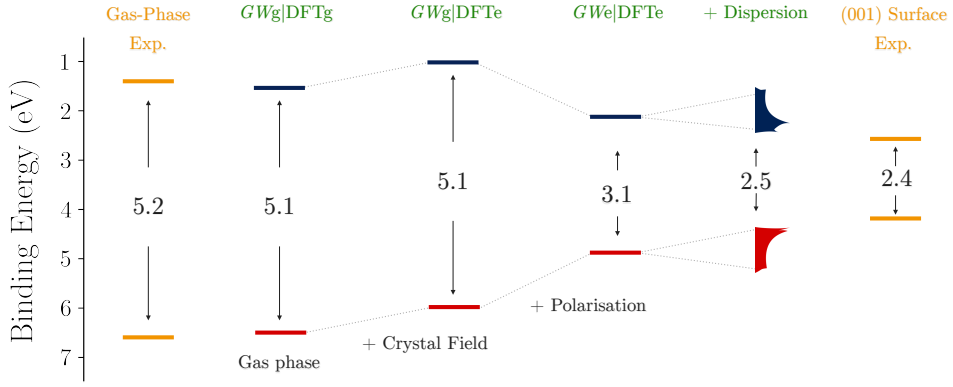

Figure 3.18: Evolution of the IP and EA of Pentacene from the gas phase to crystal surfaces. Theoretical results (blue) are presented by progressively introducing the different contributions from intermolecular interactions, i.e. polarization, crystal field and dispersion.

The polarization of the environment due to the presence of charged excitations in the system is accounted for in the GW/ME calculation, and corresponds to the induction environmental energy shift Δ_I, which is physically equivalent to the charge-dipole interaction in classical electrostatics. As roughly estimated in Sec. 3.4.3, it closes the gap by approximately 2 eV.

Finally, a fair comparison to the bulk experimental levels requires to account for band dispersion. The latter are introduced by means of a tight-binding model, with first nearest neighbors hopping parameters, that is fully parameterized *ab initio*. As shown in Fig. 3.18 edges of the resulting density of states for the HOMO and LUMO bands show an agreement within 0.2 eV for both the experimental IP and EA. This difference is comparable to the experimental uncertainty as well as the spread of values obtained from different UPS experiments arising from structural and energetic disorder [121, 135].

[135] Yoshida et al. (2015).

This example highlights the quantitative accuracy of this hybrid QM/MM scheme for the evaluation of electron addition and removal energy levels in the solid state from first-principles. Such multi-scale methods provide a general and computationally efficient framework that can be applied for the calculation of the charge transport levels of complex and disordered molecular systems. Moreover, they also provide the starting point for the evaluation of optical excitations within the framework of embedded Bethe-Salpeter formalism.

Chapter 4

Charge Transfer in Doped Organic Polymers

Charge Transfer in Doped Organic Polymers

4

Caressez un cercle, il deviendra vicieux.

Eugène Ionesco

Contents

This work has been published in the Journal of Materials Chemistry C [136].

[136] Comin et al. (2022).

MOLECULAR DOPING is one of the most important methods to control charge carrier concentration in Organic Semiconductors (OSCs). The insertion of dopant impurities in the host semiconductor allows to deliberately tune the band alignment at interfaces and is known to increase the conductivity by orders of magnitude, for both molecular systems and conjugated polymers [11, 13, 30]. The doping mechanism is understood microscopically as a two-step process: (i) the transfer of an electron between the dopant and the host material (Charge Transfer, CT) and (ii) the subsequent separation of this charge from its parent ionized dopant resulting in a free carrier (charge release). These two elementary steps will be the subject of the studies presented in the present and the forthcoming Chapters.

[11] Walzer et al. (2007).

[13] Salzmann et al. (2016).

[30] Jacobs and Moulé (2017).

In contrast to inorganic semiconductors where each dopant provides one free charge, the doping efficiency in organic semiconductors is limited by a number of factors including the intrinsic structural disorder, dopant aggregation, electron-hole Coulomb interactions (discussed in Chapter 5) and the degree of charge transfer. In particular, fundamental questions regarding the charge-transfer nature of the ground state of host-dopant complexes in organic semiconductors are still open. Evidences for both Fractional Charge Transfer (FCT) [1] and Integer Charge Transfer (ICT) ground states have been found in optical and vibrational spectra of molecular and polymer systems. In the FCT scenario, a hybridization between the host and dopant wavefunctions takes place because of their orbital overlap, resulting in a ground state in which only a fraction of an electron is transferred to (or shared with) the semiconductor. This condition seems far from being ideal for efficient charge release. Conversely, an ICT ground state consists of a fully-ionized dopant, with the charge being transferred to the host semiconductor. The formed electron-hole pair is then bound by the Coulomb interaction. It has been argued that ICT is predominant in polymer systems while FCT befalls in molecular OSCs [137]. It is in fact presently unclear which factors among the many possibly in place (charge delocalization, nanostructure, disorder, electron-hole interactions) determine the formation of a FCT or ICT host-dopant

1. Host-dopant complexes featuring FCT are called Charge-Transfer Complexes (CTCs) in the litterature.

[137] Mendez et al. (2015).

complex in the ground state.

A widely used argument for designing doped OSCs is that the doping efficiency would only depend on the ionization energetics of the donor and acceptor. Despite having found plausible experimental confirmations, this hypothesis ignores the important role of interactions in the condensed phase, which make the measurements of Ionization Potential (IP) and Electron Affinity (EA) strongly dependent on the measurement technique and on the molecular organization in the solid state. These interactions have been shown to affect the EA of molecular dopants by as much as 1 eV [38]. Moreover, the important role of excitonic and polaronic effects in the formation of ICT complexes has been highlighted. These findings call for a theoretical treatment of the CT energetics that accounts for both the excitonic and the environmental properties.

[38] Li et al. (2019).

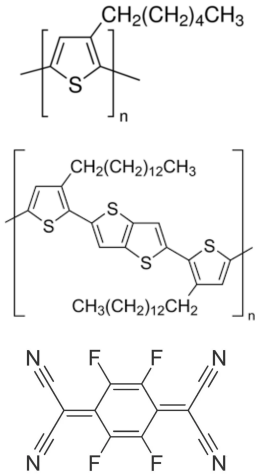

Figure 4.1: Semi-developed chemical formulas of P3HT, PBTTT and F4TCNQ.

The settlement of the relationship between structure and properties is extremely complicated in conjugated polymers. Their intrinsic disorder is characterized by a marked polydispersity — i.e. the dispersion of the distribution of polymer chain lengths — which can attain several orders of magnitude, and a complex structural folding due to the energetic freedom associated with the torsion angles. This leads to a variety of possible nanostructures in pristine systems and even more in doped samples, which makes it difficult to study their properties systematically. For these reasons, many studies focused on highly ordered conjugated polymers such as poly(2,5-bis(3-tetradecylthiophen-2-yl)thieno[3,2-b]thiophene) (PBTTT) and poly(3-hexyltiophene-2,5-diyl) (P3HT), typically doped with the strong electron acceptor 2,3,5,6-tetrafluoro-7,7,8,8-tetracyanoquinodimethane (F4TCNQ), see Fig. 4.1.

As illustrated in Fig. 4.2, the core structure of PBTTT consists in a stacking of π-conjugated chains made of repeated monomers, each formed by a Thiophene-Thienothiophene-Thiophene (T-TT-T) succession, with an Alkyl chain departing from each Thiophene ring. This results in a lamellar structure whose electronic properties arise from the π-conjugated region while the Alkyl region forms an insulating layer that is mostly responsible for the structural properties of the material.

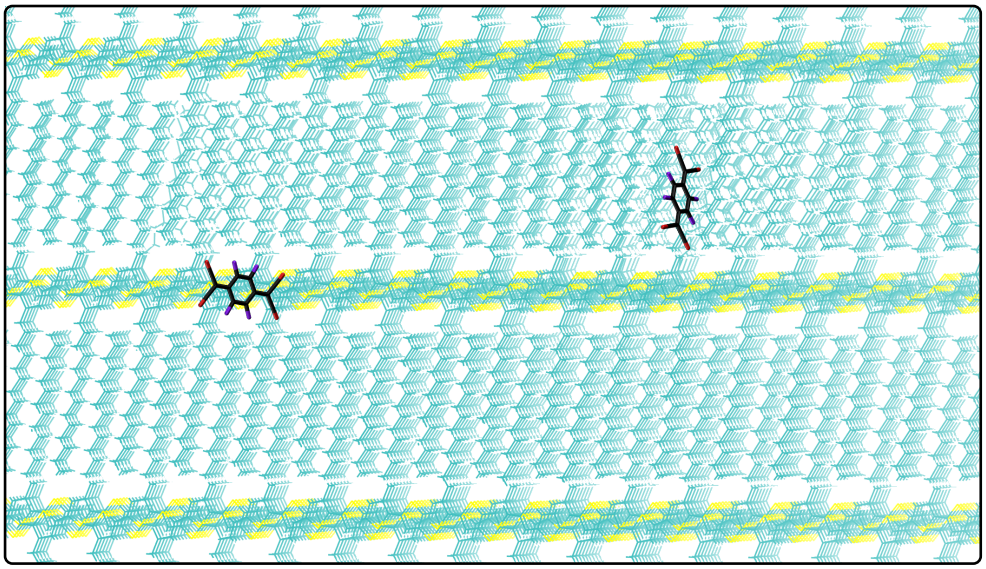

Figure 4.2: Illustration of the PBTTT polymer structure and the two possible dopant locations: (left) α system, the F4TCNQ is placed in the alkyl side chains, (right) π system, the F4TCNQ is next to the π-conjugated backbone.

Even in the case of a semi-crystalline polymer such as PBTTT, experimental reports pointed out to various nanostructures depending on the processing, with

dopants showing a preferential affinity for either the region of the π-conjugated backbones [138, 137, 139] or the inert alkyl regions that are typically grafted to the functional core to facilitate solution processing [14, 140, 141, 142, 143, 144].

Such differences in the nanostructure were shown to have a major impact on charge transport properties. Indeed, room-temperature conductivities differing by 2 orders of magnitude have been related to the transition from a FCT to an ICT in PBTTT doped with the strong electron acceptor F4TCNQ [139]. When the dopant is co-deposited with the polymer solution, dopants are found to be interlaced between the π-stacked conjugated chains of the polymer, as shown on the left of Fig. 4.2. In this system, which we will refer to as the **π** structure, the dopant positioning induces some structural disorder in the π stacking and the resulting conductivity at the optimal doping ratio is found to be relatively moderate, namely 2 S.cm^{-1} [138].

Conversely, when F4TCNQ is evaporated on top of PBTTT layers the dopant is allowed to diffuse and is found to be in the layer of the alkyl side chains, as shown on the right of Fig. 4.2. In this system, which we will refer to as the **α** structure, the polymer retains a highly ordered lamellar nanostructure and the resulting conductivity is much higher, up to 250 S.cm^{-1} [14]. Room-temperature conductivities as high as 1380 S/cm have been reported for aligned films of PBTTT doped with F4TCNQ by means of the incremental concentration doping technique [143]. These profound differences call for an in-depth explanation and serve as a relevant and interesting case study for theoretical calculations based on QM/MM electronic structure calculations.

Therefore, we study in this Chapter the structure-property relationships between the different nanostructures arising in PBTTT doped with F4TCNQ and their associated ionization energetics. We will be interested in whether and how charge transfer takes place in the ground state by applying our Embedded Many-Body Perturbation Theory (EMBPT) approaches on the system in the infinite dilution limit. We begin with a brief digression on a previous study treating the role of the excitonic and polaronic effects in the CT energetics of doped OSCs. Then, we present the preliminary analyses that will set the stage to the subsequent many-body treatment of the ground and excited states of the system.

[138] Cochran et al. (2019).

[139] Jacobs et al. (2018).

[14] Kang et al. (2016).

[140] Hamidi-Sakr et al. (2017).

[141] Scholes et al. (2017).

[142] Hase et al. (2018).

[143] Vijayakumar et al. (2020).

[144] Untilova et al. (2021).

4.1 Preface: Dopant Ionization in F4TCNQ-Pentacene

As a brief digression from the main topic of the present chapter, we discuss the work of Li et al. [37] which sheds lights on the partial versus full charge transfer scenario in a doped molecular semiconductor, before turning to polymer systems. In a work that preceded this Thesis, the authors revisited the paradigmatic case of Pentacene doped with F4TCNQ in the infinite dilution limit, contributing to clarify the nature of the ground state of this system, and in particular its FCT or ICT character.

[37] Li et al. (2017).

On the one hand, the optical properties of Pentacene-F4TCNQ blends first investigated by [12] resulted in novel absorption lines at 1.2-1.4 eV that were absent in the spectra of the two pristine components, and thus arose upon doping. The latter have been ascribed originally to CT transitions between Pentacene and F4TCNQ resulting in a partially ionized dopant, according to gas-phase Density-Functional Theory (DFT) calculations on a cofacial dimer [12]. On the other hand, scanning tunneling microscopy images suggested that isolated F4TCNQ in Pentacene films are fully ionized at room temperature [43], an observation that has been rationalized on the basis of electrostatic modeling [145, 146]. By means of embedded Many-Body Perturbation Theory and model Hamiltonian calculations, Li et al. showed that full dopant ionization takes place at room temperature thanks to the contribution of the electron-hole binding energy, i.e. in virtue of an excitonic effect.

[12] Salzmann et al. (2012).

[43] Ha and Kahn (2009).

[145] Topham and Soos (2011).

[146] Theurer et al. (2021).

The system that was investigated consists of a single F4TCNQ substitutional defect in the Pentacene crystal lattice [43]. Following a multi-scale approach, the system is partitioned into a QM region that comprises the single dopant and its first shell of 6 Pentacene neighbors, embedded into the Pentacene crystal forming the MM region. The QM complex is described at the DFT, *GW* and BSE level of theory while the MM region is treated by the classical Charge Response (CR) model, which is a generalization of Micro-Electrostatics (ME) that accounts for intramolecular charge redistribution.

In addition to the QM/MM approach, an accurately parameterized model Hamiltonian was defined in order to include the potentially important effects associated with the delocalization of the hole transferred to Pentacene molecules beyond the first shell of neighbors and structural relaxation or polaronic effects. This generalized Mulliken model for intermolecular CT describes a single dopant in the crystal ab plane of the host OSC on the basis of a neutral state $|\mathcal{N}\rangle$ and singlet full-CT states $|i_n\rangle$ where the electron populates the dopant's Lowest Unoccupied Molecular Orbital (LUMO) and the hole the Highest Occupied Molecular Orbitals (HOMOs) of OSC site i. The model was parameterized with the CT state energies obtained from *GW*-CR calculations accounting for the environmental energy shifts, and with appropriate CT integrals between molecular sites computed at the DFT level [2]. Additionally, the Hamiltonian was extended in order to account for intramolecular structural relaxation upon charging (polaron formation) within the framework of the Mulliken-Holstein model [37].

2. This model will be used later in this Thesis (Sec. 5.1) to assess the polarizability of host-dopant complexes.

Fig. 4.3 summarizes the main results obtained from embedded ev*GW* and embedded BSE calculations for the complex in its proper dielectric environment. The HOMO-LUMO gap is found to be 0.67 eV, which is dramatically larger than room temperature thermal energy, clearly illustrating the difference with the physics of

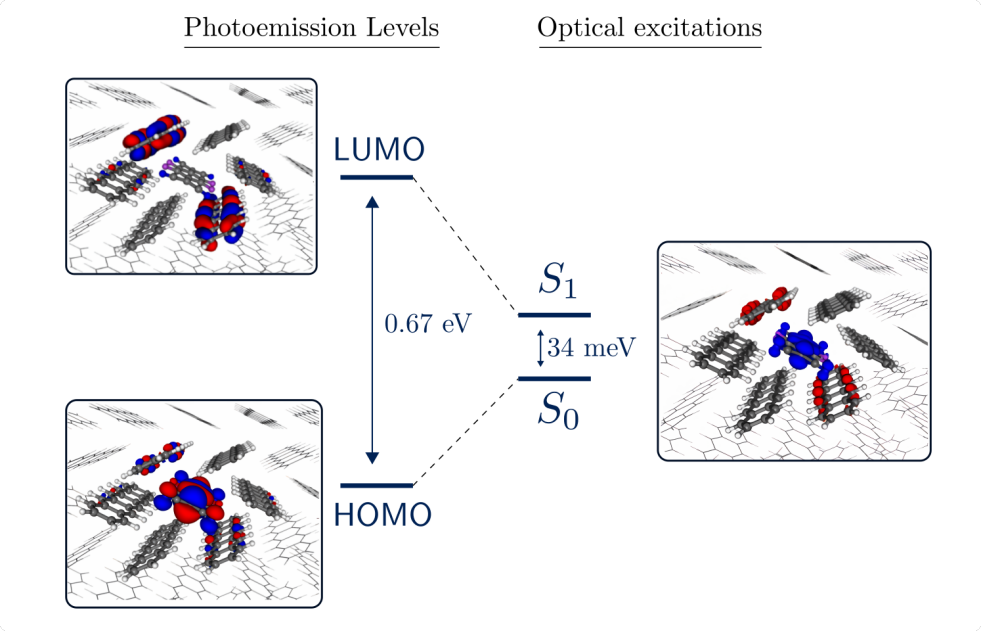

Figure 4.3: Isocontour representation of the frontier molecular orbitals and the associated photoemission levels, with the corresponding optical transition S_1 and the associated hole (red) and electron (blue) densities, for the Pentacene-F4TCNQ complex.

inorganic semiconductors where shallow impurity levels are formed near the band edges. The analysis of the frontier orbital isocontours presented in Fig. 4.3 reveals that the LUMO of the QM supra-molecular complex can be identified to the LUMO of the dopant, while the HOMO and HOMO-1 levels stem from a superposition of the HOMOs of a specific pair of symmetry-equivalent Pentacene molecules.

In the absence of orbital hybridization between host and dopant, the gap was found to be 0.45 eV, which allows to conclude that the HOMO-LUMO gap of the complex is mostly sourced from the energy mismatch between the IP of Pentacene and the EA of F4TCNQ. Such a limited hybridization manifests in a partial charge transfer in the ground state, with a net charge on the dopant $Q_{dop} = 0.25e$, as estimated by Electrostatic Potential Fitting (ESP) atomic charges.

While the photoemission gap of the complex is large, the lowest optical excitation, as obtained within the embedded BSE formalism, is found to be extremely low in energy, at 34 meV above the ground state, and corresponds to a complete charge transfer between the HOMO and the LUMO of the QM complex. Such a large difference between the photoemission and the optical gap in a purely CT excitation reveals an important electron-hole binding energy that comes mostly from the screened Coulomb potential W_{eh}, which is quantitatively different from the Hydrogenic model for inorganic semiconductors where the exciton binding energies are comparable to room-temperature thermal energy.

The optical spectrum closely reproduces the experimental absorption line of pristine Pentacene at 1.85 eV with additional new features in the 1.3 eV–1.6 eV energy range, which may correspond to the sub-gap peaks observed experimentally. BSE calculations suggested that such excitations arise mainly from CT transitions between the manifold of HOMO-1 Pentacene orbitals to the LUMO of F4TCNQ. Moreover, further absorption peaks in the 0.3-0.6 eV energy range were predicted, which fall in the mid-infrared region that has not been investigated experimentally.

The picture emerging from this zero-temperature QM/MM approach is that of a

system very close to the neutral-CT boundary with two competing ground states, namely a quasi-neutral one with modest intermolecular CT, and full-CT one with completely ionized molecules. In this situation, structural relaxation effects associated with polaron formation may play a determinant role by stabilizing ionized species. In fact, the small optical gap of this system is comparable with the polaron binding energies, namely $\lambda^+ + \lambda^- = 0.2\,\mathrm{eV}$, and the inclusion of polaronic effects in the model Hamiltonian indeed leads to a symmetry-broken ground state with a full CT character [37]. This prediction is in excellent agreement with very recent optical and vibrational spectroscopy data demonstrating the majority presence of fully ionized complexes in Pentacene-F4TCNQ and Pentacene-F6TCNNQ blends [146]. The relaxed-structure calculations also allowed to ultimately assign the bands at 1.2-1.4 eV as characteristic transitions of the Pentacene cation, in agreement with the literature data [147, 134].

[147] Brinkmann et al. (2004).

[134] Sakanoue and Sirringhaus (2010).

These theoretical findings, tightly connected to experimental evidences, show a novel picture of charge transfer in molecular semiconductors in which a full ionization of the dopant is made possible by the excitonic e-h interactions, and by a smaller, yet determinant, polaronic effect.

4.2 Preliminary Analyses

We now turn to the main topic of this Chapter, namely the problem of charge transfer in PBTTT doped with F4TCNQ. The ionization energetics are related to the photoemission and optical electronic levels, which are accurately described within the Green's function *GW* and BSE formalisms (see Section 3.3). However, while the IP and EA are well-defined quantities for an isolated molecule, they can present sizable variations between different samples of the same compound, due to their dependence on their environment [121, 120, 38]. Extended molecular solids can in principle be described with periodic *GW* or BSE calculations, but these approaches are rather computationally expensive in that they require a large number of atoms in the unit cell and are strictly limited to periodic systems. Moreover, when describing doped OSCs at low doping concentrations, periodic treatments are subject to spurious interactions between periodic replica. Finite-size QM/MM approaches seem to be more suited for the description of such systems in that they inherently avoid these limitations.

[121] Salzmann et al. (2008).

[120] Duhm et al. (2008).

Such multiscale approaches require to partition the system into a QM region comprising one dopant and a part of the host, that is described by either DFT of MBPT, and an MM region that comprises the surrounding polymer crystal, described within Micro-Electrostatics (ME). In this Section, we discuss the preliminary analyses pertaining to the choice of the QM/MM partitioning and the density functional employed to obtain starting molecular orbitals and energy levels in the subsequent study of the many-body ground state and excited electronic levels.

4.2.1 Convergence of Electronic Properties with the Polymer Length

The QM/MM partitioning of a system requires the QM and MM regions to be non-overlapping, a criterion which can plausibly be satisfied for systems composed of disjoint molecular units by incorporating a finite number of molecules in each of the subregions. The QM/MM partitioning of a virtually infinite conjugated polymer is however more involved and inevitably arbitrary. A practical solution then consists in retaining within the QM region only a physically relevant part of the polymer.

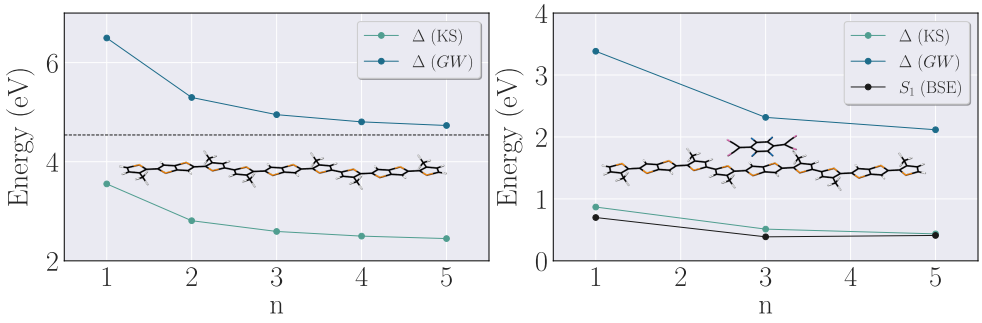

Figure 4.4: Dependence of the electronic properties of nBTTT on the oligomer size n: Kohn-Sham and *GW* gap of nBTTT alone (left) and nBTTT with F4TCNQ (right). All results obtained from gas-phase calculations, using the PBE0 functional and 6-311G* basis and the Coulomb fitting resolution of the identity (RI-V).

Since the alkyl chains contribute mainly to the structural rather than the electronic properties of the material, these will be excluded from the QM region and replaced by methyl (CH_3) groups. This then leaves us with the choice of how many monomers of the π-conjugated backbone to include in the QM region. We lift the arbitrariness of such a choice by selecting the number of monomers that ensures a reasonable convergence of the electronic properties with respect to the oligomer size. To this aim, we consider an isolated polymer chain optimized from a one-dimensional periodic

[148] Dovesi et al. (2018).

DFT calculation at the PBE0/6-31G* level, performed with the CRYSTAL17 code [148]. The planarity of the polymer has been imposed in the structural relaxation in order to reproduce the chain system in the solid state.

We then compute the Kohn-Sham (KS) and *GW* gaps of oligomers including n monomers, which we dub nBTTT, as a function of n. The results are presented in the left panel of Fig. 4.4, and we observe that both the KS and *GW* HOMO-LUMO gap decrease with length, with the latter converging to a constant asymptotic value of 4.54±0.1eV (horizontal line) obtained by extrapolation. The *GW* gaps of 3BTTT and 5BTTT are converged within 0.41 and 0.19 eV from the infinite-chain limit.

The right panel of Fig. 4.4 presents the convergence of the electronic levels for the system including additionally a F4TCNQ dopant stacked on top of the π-conjugated plane of nBTTT. The structure was optimized at the PBE0-D3/6-31G* level for the monomer, and the same position of the dopant was then retained also in calculations for longer oligomers. The KS and *GW* gaps as well as the first optical excitation S_1 obtained from the Bethe-Salpeter equation (BSE) are shown as a function of the oligomer length n. We observe again a convergent behavior of the electronic properties, especially for the first optical excitation S_1, which corresponds to a host-dopant CT transition and differs by only 20 meV between the 3BTTT/F4TCNQ and 5BTTT/F4TCNQ systems. Qualitatively, the hole on the conjugated polymer is Coulombically bound to the electron on the dopant and thus cannot delocalize along the full conjugated chain. The faster convergence of the optical excitations in donor-acceptor systems results from the excitonic electron-hole interaction that tends to localize charges [149]. Thus, this satisfactory convergence justifies our choice to carry all further analyses on 3BTTT.

[149] Niedzialek et al. (2015).

4.2.2 Density Functional Tuning

Another important point is the tuning of the DFT functional as to reproduce at best the gap obtained from *GW* calculations in the condensed phase. Indeed, as we discussed in Sec. 3.2, the KS gap is strongly dependent on the parameters that affect the DFT functional, e.g. the amount of exact exchange in the PBEh(a_x) functional. Even though this influence will be smeared out by the subsequent MBPT calculations, it is desirable to select the best starting point possible, in particular in view of the fact that the KS eigenstates, and hence the hybridization between the polymer and the dopant's molecular orbitals, will not be corrected beyond the DFT level.

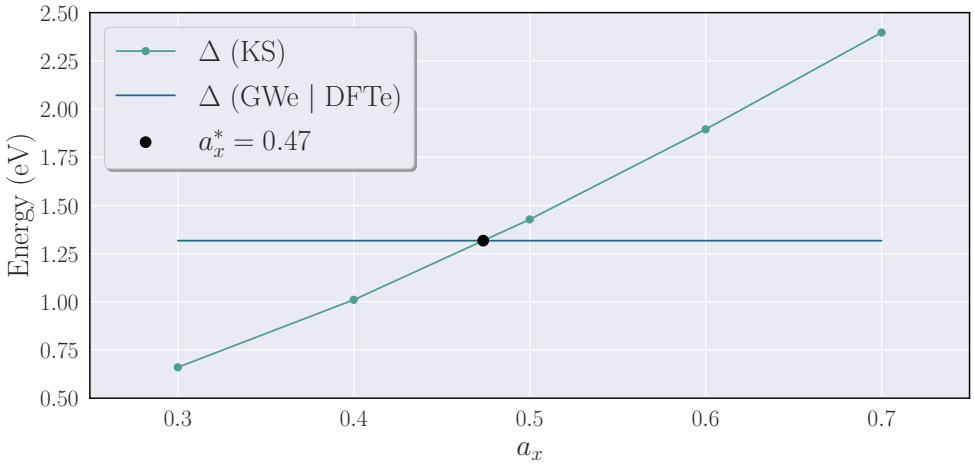

Figure 4.5: Tuning of the PBEh(a_x) DFT functional against an embedded *GW*e | DFTe calculation for 3-BTTT-F4TCNQ complex. The crossing with the KS gap happens at $\alpha = 0.47$.

We thus compute the KS gap of 3BTTT with F4TCNQ as a function of the amount of exact exchange a_x, and compare it to that obtained from an embedded *GW* calculation in a polarizable environment of dielectric constant $\varepsilon = 3.5$. The latter is achieved by defining an MM region consisting of an FCC lattice of polarizable points with lattice constant $a = 1.5\,\text{Å}$ and site polarizability $\alpha = 2.93\,\text{Å}^3$. The separation between the QM and MM regions was taken to be that of the Van der Waals surface of the QM molecules, multiplied by an empirical factor of 1.2, in order to avoid spurious effects arising from polarizable points being too close to the QM region. Fig. 4.5 shows the result of this tuning procedure, giving an optimal value of $a_x^* = 0.47$. In the following of this study, we will consider the rounded value of $a_x^* = 0.45$.

4.2.3 QM/MM Setup

We now define the QM/MM partitioning of the system that will be investigated. The starting structure is obtained from a force field simulation of crystalline PBTTT with a single F4TCNQ dopant located either in the π-conjugated ($\boldsymbol{\pi}$ system) or the Alkyl chains ($\boldsymbol{\alpha}$ system) regions. From this structure, the F4TCNQ molecule and the nearest 3BTTT oligomer are extracted in order to be eventually included in the QM region, while the remaining of the crystal is to form the MM region.

The separation of 3BTTT from the polymer chains requires special care, as it involves parting the oligomer from both the conjugated and the Alkyl chains, the latter not contributing substantially to the electronic properties. Such splitting is achieved by replacing the junction between the Thiophene rings in 3BTTT and the Alkyl chains with a Methyl group (CH_3). The next Carbon atom in the Aklyl chain is removed to avoid spurious interactions between the QM and MM regions originating from atoms being too close to the frontier, and the remaining reduced Alkyl chain ($CH_3(CH_2)_{10}CH_3$) is included in the MM region. Then, the separation of 3BTTT from the conjugated polymer chain is done by breaking the covalent Thiophene-Thiophene bond and capping 3BTTT with a Hydrogen. The Thiophene ring closest to the junction in the MM region has been removed while keeping its associated Alkyl chain by splitting it as described above, in order to avoid unphysical interactions arising from a close contact between the QM and MM regions.

This strategy results in the structures of 3BTTT and F4TCNQ to be included in the QM region, and the MM region is then made up of the remaining conjugated and Alkyl chains. We consider a lamellar MM structure similar to the edge-on alignment of PBTTT, in accordance with Grazing-Incidence Wide-Angle X-ray Scattering (GIWAXS) measurements [150]. The latter will be described at the Micro-Electrostatic level, which requires to be parameterized with the partial atomic charges and polarizability tensors of its molecular constituents. The polymer chains were separated into different segments in order to achieve such parameterization. Specifically, the Alkyl chains, the Thiophene and Thienotiophene rings were assigned independent partial charges and polarizabilities, computed from DFT calculations at the B3LYP/6-311+G** level. We note that Hydrogens have been made implicit within the MM region, leading to a drastic reduction of the number of atoms in the MM region without compromising the accuracy of the ME model. In practice, their polarizabilities have been redistributed over the segment while partial atomic charges are instead obtained from an Electrostatic Potential Fit (ESP) that considers only heavy atoms.

[150] Tanaka et al. (2019).

We are now in position to define the $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$ systems as the structures including 3BTTT with F4TCNQ in the corresponding position. After the QM/MM partition-

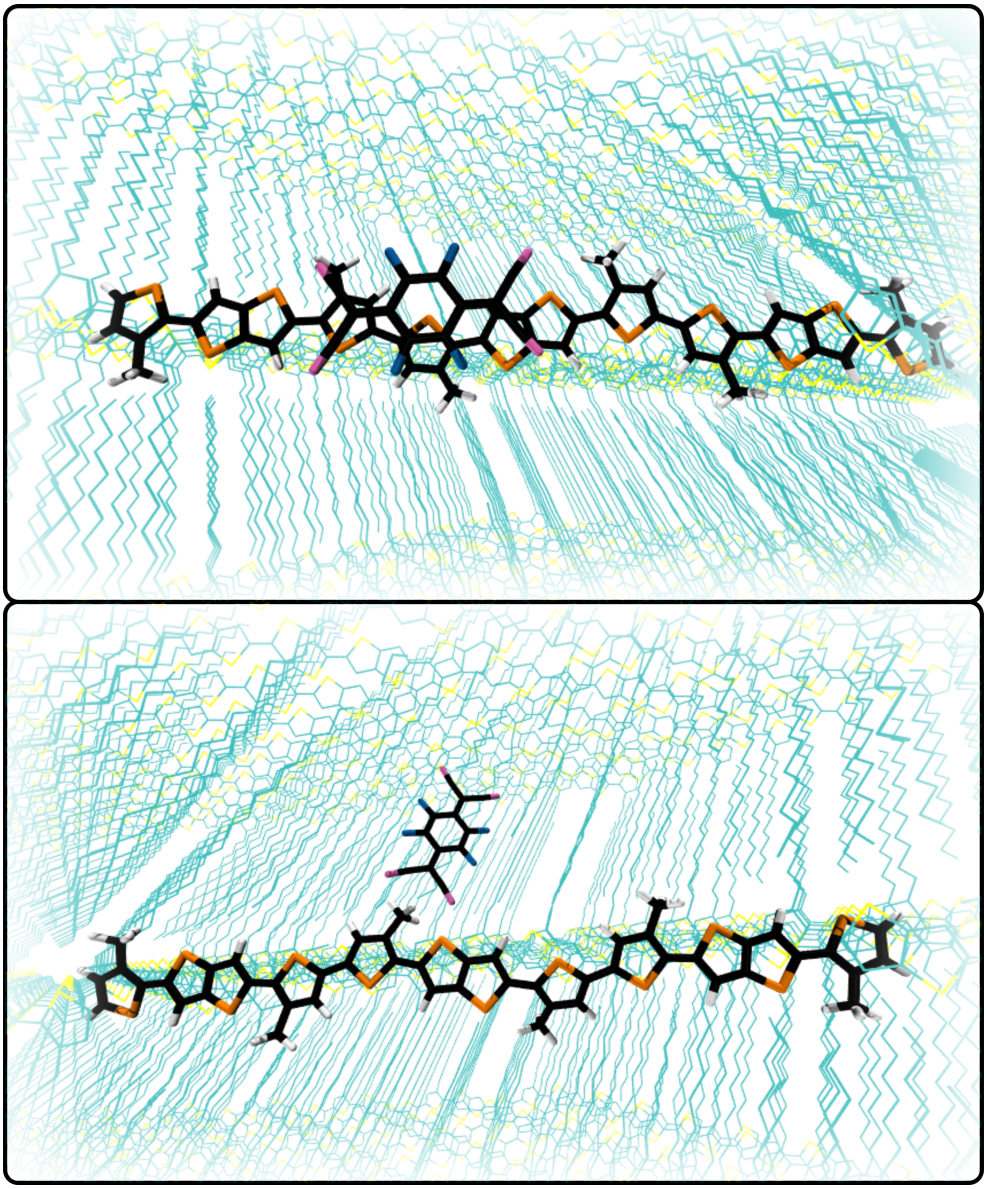

Figure 4.6: Rendering of the π (top) and the α (bottom) structures that will be used throughout this study. In the π structure, the dopant is stacked on top of the π-conjugated Thiophene chains while in the α structure it is found in between the Alkyl chains.

ings, the geometries of each structure are relaxed at the DFT level with the tuned PBEh(α_x^*)-D3 functional with dispersion corrections and the 6-31G* basis, accounting for Lennard-Jones and electrostatic interactions with the MM region. From now on, we will refer to these two geometries, depicted in Fig. 4.6, as the π and α systems.

4.3 Electronic Levels in the Non-Overlapping Limit

Having defined the system under study as well as the optimal DFT functional that will be used as the starting point for our many-body calculations, we now explore as a first step the donor and acceptor electronic levels, considered as non-overlapping molecular fragments. This strategy will hence constrain charged excitations on individual molecular units, allowing us to grasp the electrostatics effects involved in the system in the absence of intermolecular orbital hybridization, which will be treated in the next section.

The non-overlapping limit is reached by considering either the host or the dopant in the QM region when performing embedded QM/MM calculations. According to the preliminary study that we conducted on a single polymer chain, considering the 3BTTT oligomer in place of the polymer chain is an appropriate choice for the convergence of the electronic properties. We are specifically interested in the HOMO and HOMO-1 of 3BTTT and the LUMO of F4TCNQ, as these are the frontier levels that are expected to be mostly involved in charge transfer.

The partitioning of the QM/MM system described above enables us to compute the charged excitations of F4TCNQ-doped PBTTT in the π and α systems. Since we are considering here the electronic levels of the host and the dopant in the non-overlapping limit, the QM region consists either in a 3BTTT or F4TCNQ, while the MM region comprises a large cell of the PBTTT crystal described at the Micro-Electrostatic level.

All DFT calculations are done at the PBEh(α_x^*)/6-311G*, and *GW* calculations consider 5 ev*GW* iterations [3] from the DFT ground state. In addition to the frontier electronic levels, the HOMO-1 level was also corrected; indeed, we expect it to play a role in the level rearrangements because of its proximity to the HOMO level, both in terms of energy and spatial arrangement. Embedded DFT calculations are done at the PBEh(α_x^*)/6-311G* level and consist in 6 self-consistent cycles between the QM and MM regions, so as to guarantee the convergence of the interactions between the two regions. Finally, the embedded *GW* calculations are performed at the Δ^{COHSEX} level. The polarization of the MM region in response to quasiparticle excitations is described by the Reaction Field Matrix (RFM), which is calculated by considering two spherical cuts of the MM region of radii 35 and 55 Å, allowing to extrapolate it to the infinite system limit.

3. We recall that in the ev*GW* scheme, the corrected quasiparticle energy levels are reinjected self-consistently in the construction of the Green's function G and the screened Coulomb potential *W*, see Sec. 3.3.

The shift of the energy levels from the gas to the solid state are defined as the environmental energies (see Sec. 3.5). The latter can be separated in an electrostatic contribution Δ_E arising from the ground-state electrostatic effects, and an induction contribution Δ_I arising from the polarization of the MM region in response to charged excitations. These are captured respectively by embedded DFT (DFTe) calculations and embedded *GW* calculations (*GW*e). As a result, a gas-phase *GW* calculation starting from an embedded DFT calculation (*GW*g | DFTe) will only account for Δ_E, an embedded *GW* calculation starting from a gas-phase DFT calculation (*GW*e | DFTg) will only account for Δ_I, and an embedded *GW* calculation starting from an embedded DFT calculation (*GW*e | DFTe) will account for both, see Table 4.1.

*GW*g | DFTg	ε^n
*GW*g | DFTe	$\varepsilon^n + \Delta_E^n$
*GW*e | DFTg	$\varepsilon^n + \Delta_I^n$
*GW*e | DFTe	$\varepsilon^n + \Delta^n$

Table 4.1: Correspondence between the environmental contributions captured by different embedded *GW* calculations.

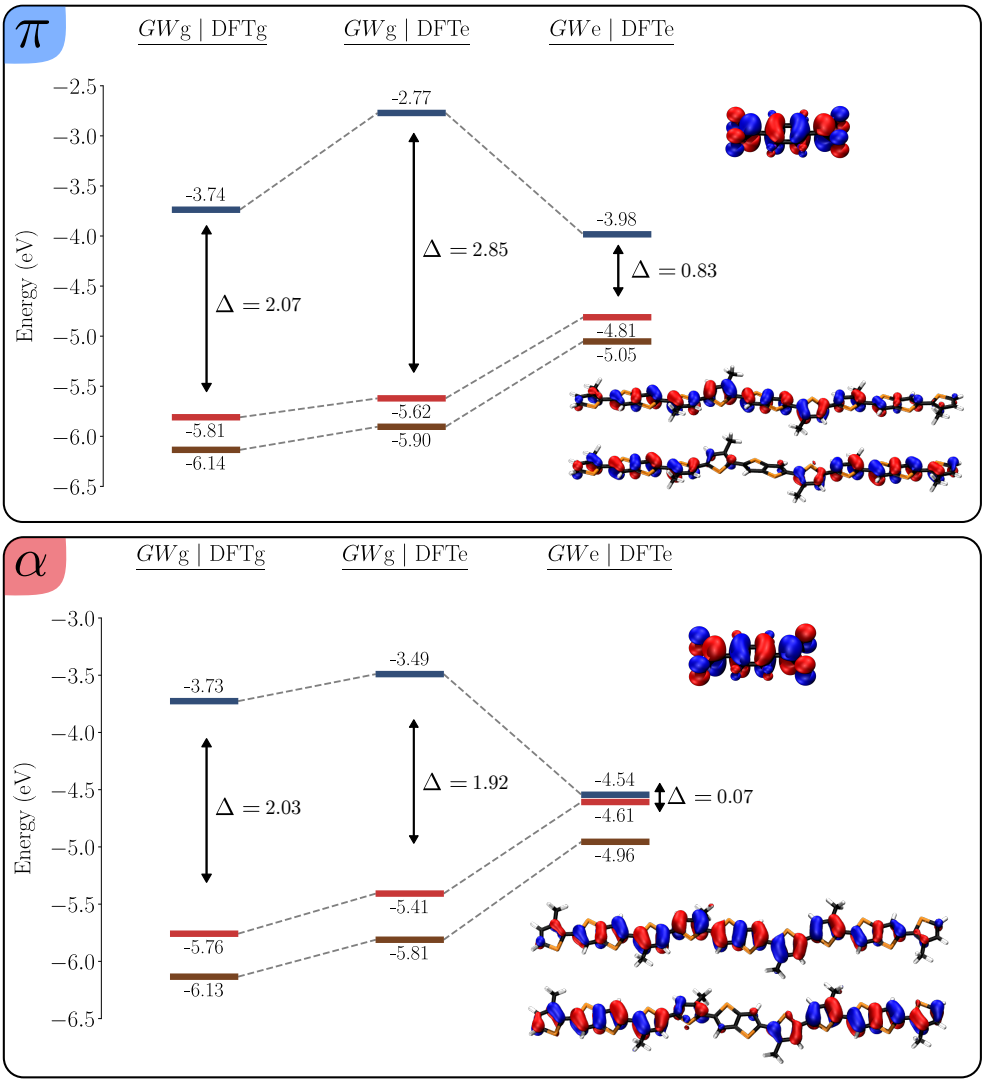

Figure 4.7: Environmental contributions to the *GW* energy levels in 3BTTT and F4TCNQ treated as non-overlapping units in the π structure (top) and the α structure (bottom).

In order to dissect the individual electrostatic and induction contributions to the electronic levels, we present in Fig. 4.7 the frontier electronic levels of isolated 3BTTT and F4TCNQ as obtained from *GW*g | DFTg, *GW*g | DFTe and *GW*e | DFTe calculations, for the π and α systems. A first observation is that the HOMO and LUMO in the gas phase are essentially equal for both systems, leading to a donor-acceptor gap of about 2 eV . This is indeed expected since without the environment around, the molecules in the π and α systems reduce to the same chemical species, with small differences in the energy levels arising from the slightly different geometries, following from the independent QM/MM optimizations (see Sec. 4.2).

Conversely, the electrostatic contribution Δ_E is rather different in the two systems. As noted previously, Δ_E can be seen as a measure of the average electrostatic potential felt by the atoms in the QM region. Since the HOMO and HOMO-1 levels share the same spatial region, they sit in the same potential region and are thus equally shifted in a rigid way. We observe that the value of Δ_E for these levels is different for the π and α systems. This is attributed to the different electrostatic interactions with the dopant in the MM region which has different relative position and orientation in the two systems [39].

[39] Privitera et al. (2020).

The electrostatic contribution to the LUMO levels of the dopant is quite different

from that of the host's levels. Indeed, the LUMO of the π system is 0.72 eV higher than the LUMO of the α system. Such a disparity can be rationalized on the basis of the classical electrostatic potential surface in the polymer matrix. By knowing the polarizability and partial atomic charges of each segment of the polymer chains, it is possible to perform a neutral Micro-Electrostatic calculation that reveals the potential felt by each atom arising from the mutual electrostatic interactions of all components in the sample. This electrostatic potential map is reported in Fig. 4.8. The potential shows a characteristic striped pattern corresponding to the alternation of high and low potential regions in the Alkyl and conjugated chains respectively. In particular, we observe that the π-conjugated chains sit in a region of potential that is approximately 0.2 eV lower than that in the neighboring of the Alkyl chains.

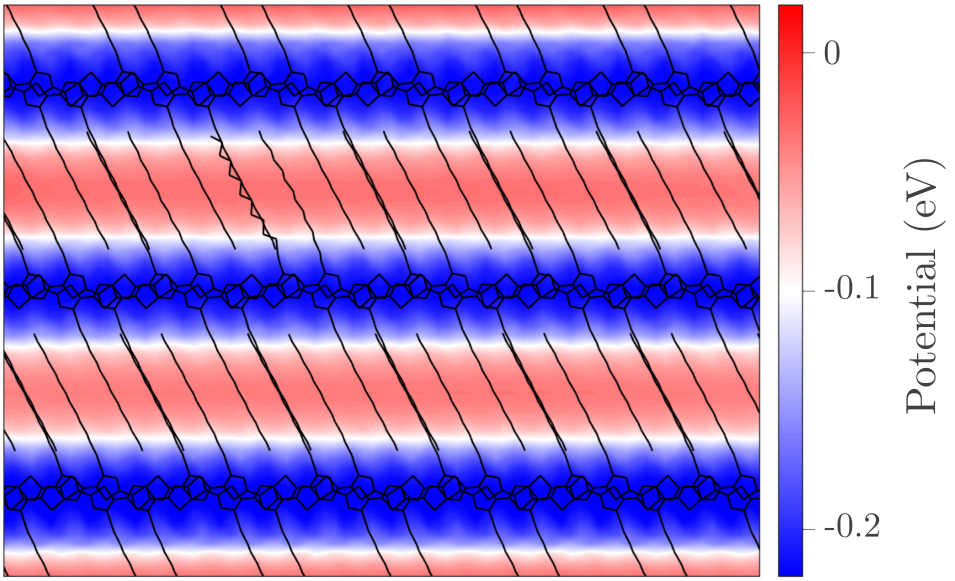

Figure 4.8: Electrostatic potential map of the PB3T crystal as obtained from classical Micro-Electrostatics calculations. The calculation considers Periodic Boundary Conditions (PBC) in 2D along the ab crystal axes, up to a circular cutoff of 510 Å, ensuring the convergence of the involved electrostatic sums.

However, this map is obtained as the convolution of the atomic potentials with a spherical kernel of radius 5 Å, defining the size of the microscopic probe, and then averaged in the out-of-plane c crystal axis. As such, it does not reflect the local inhomogeneities below the kernel resolution and along the c crystal axis. Evaluating instead the total electrostatic potential felt by the electron density of the dopant results in a more precise probe. In the π system this potential is -1.1 eV, which translates into a large destabilization of the dopant's Electron Affinity. On the contrary, in the α structure the considerable separation between the dopant and the conjugated chains results in moderate quadrupolar fields, with a potential of only -0.3 eV on F4TCNQ, leading to a correspondingly limited shift of the LUMO. Therefore, the difference between the LUMO energy in the two systems corresponds quantitatively to that of the electrostatic potential felt by the dopant.

Turning now to the induction contribution Δ_{I}, it is instead quite similar in both systems. The environment's polarization effectively screens the Coulomb field emanating from the charged quasiparticles, thereby stabilizing the excitation energies [108]. This results in a closure of the gap of approximately 1 eV per level, giving a donor-acceptor gap of 0.83 eV in the π system and an almost vanishing gap of 70 meV in the α system.

[108] D'Avino et al. (2016).

The disparate electrostatic environments in which the host and dopant are embedded provided us the key to interpret the difference between the corresponding donor-acceptor gaps. The potential energy map further offers a qualitative reading of

the electrostatic effects at play in PBTTT, and it will serve as a common thread for our analyses in the following sections. As a consequence of the imposed non-overlapping of the host and the dopant, these results do not account for intermolecular hybridization, which we expect to open the gap. Nevertheless, we observe already a conspicuous difference between the donor-acceptor gap of the two structures, with the α structure being seemingly better suited for charge transfer.

4.4 Electronic Structure of the Host-Dopant Complex

4.4.1 Intermolecular Orbital Hybridization

While the energetics of non-overlapping donor and acceptor molecules provide important insights on the charge-transfer processes of a doped OSC, the understanding of the donor-acceptor CT interactions will be key to our forthcoming analysis. These comprise all the quantum-mechanical interactions between the 3BTTT and F4TCNQ molecules, resulting in intermolecular orbital hybridization and ultimately in a certain degree of CT, either fractional of full. The essential physics behind orbital hybridization can be captured from a simple two-state model, where the relevant states are the polymer HOMO and the dopant LUMO. The quantum mixing between these two states (hybridization) is expected to open the gap between the occupied and the empty level of the complex, as shown in Fig. 4.9.

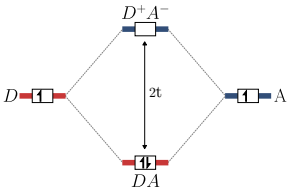

Figure 4.9: Orbital hybridization in a two-state donor-acceptor system. The hybridization between the two states opens a gap that is proportional to the transfer integral t.

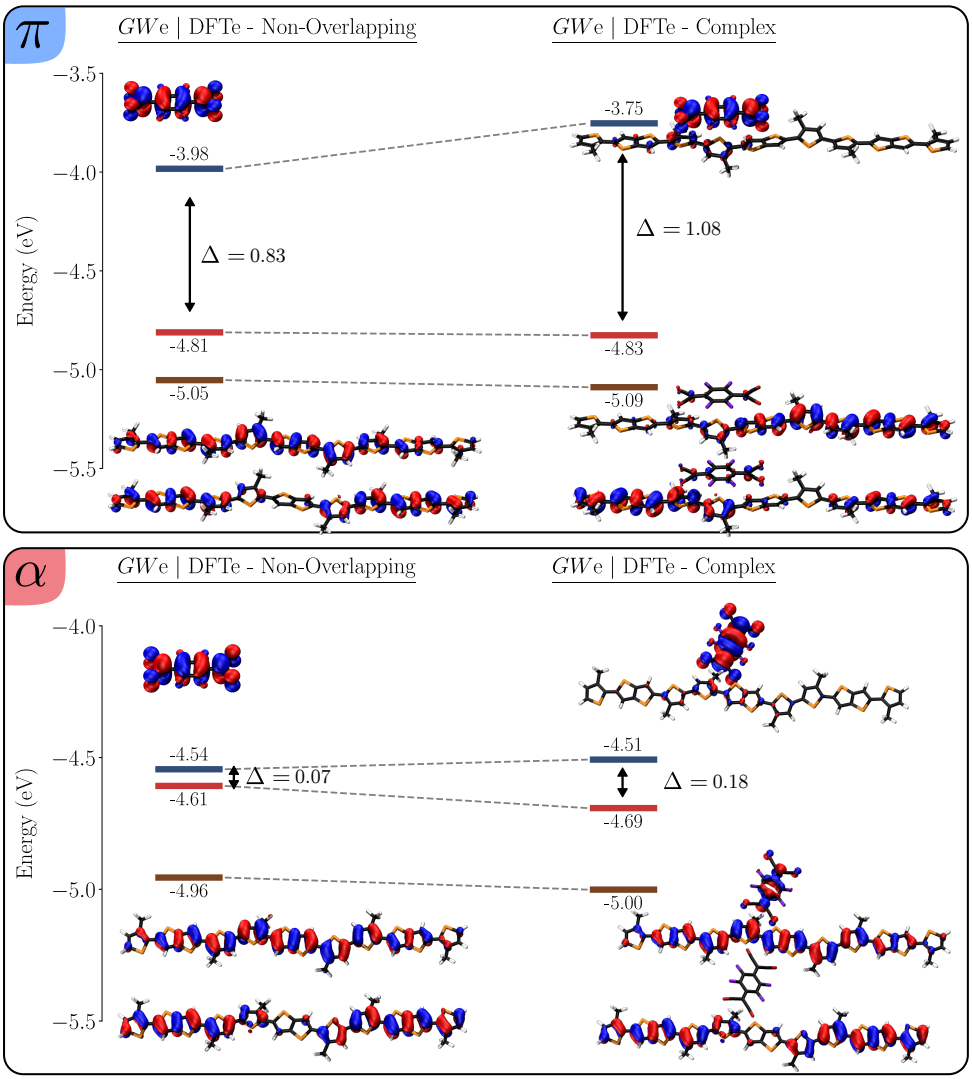

Figure 4.10: Effect of intermolecular interactions on the frontier energy levels of the 3BTTT-F4TCNQ complex in the condensed phase. (top) π system: the gap of the complex is opened by 0.25 eV. (bottom) α system: the gap of the complex is opened by 0.11 eV, reflecting a smaller orbital hybridization.

We assess the effects of hybridization in Fig. 4.10 by comparing the frontier electronic levels of the non-overlapping host and dopant molecules with those of the host-dopant complex in the π and α structures. In particular, we present the outcome of a fully embedded (*GWe* | DFTe) calculation which includes the environmental effects on both the ground and excited states [4]. We observe that in both cases, the orbital

4. We refer the reader to the beginning of Sec. 4.3 for the computational details of the *GWx* | DFTx calculations.

hybridization results in an opening of the donor-acceptor gap.

In the π structure the gap is opened by 0.25 eV, which indicates a sizable effect of intermolecular CT interactions. This change is mostly reflected in the difference between the LUMO energy levels with and without hybridization, while the HOMO and HOMO-1 barely change in energy although they feature an orbital remixing due to the close presence of the dopant. Indeed, the HOMO and HOMO-1 extended over the whole polymer segment in the non-overlapping case, with the latter orbital featuring a node at its center. The shape of the orbitals is consistent with the inversion symmetry of the polymer, which is only slightly lifted by the actual QM/MM geometry. The symmetry is further lowered in the overlapping complex, with the polymer HOMO and HOMO-1 partially hybridizing with the dopant LUMO, and localizing on the left and right hand side of 3BTTT.

The degree of hybridization can be estimated from the electronic transfer integrals between the frontier levels, which we calculate by projecting the Kohn-Sham Hamiltonian of the complex onto the subspace of the HOMO-1, HOMO and LUMO levels of the two fragments, with DFT calculations performed in the gas phase. The projected Hamiltonian reads

$$\mathcal{H}_\pi = \begin{array}{c} \\ \text{L} \\ \text{H} \\ \text{H-1} \end{array} \begin{array}{c} \begin{array}{ccc} \text{L} & \text{H} & \text{H-1} \end{array} \\ \begin{pmatrix} -4.40 & 0.43 & -0.27 \\ 0.43 & -5.36 & 0.049 \\ -0.27 & 0.049 & -5.66 \end{pmatrix} \end{array}. \tag{4.1}$$

We indeed observe a large HOMO-LUMO transfer integral $t_{HL} = 0.43$ eV, explaining qualitatively the gap opening, as well as a sizable coupling between the HOMO and HOMO-1 $t_{HH\text{-}1} = -0.27$ eV which confirms the observed remixing of these levels. Thus, the photoemission gap in the π system, accounting for all environmental effects, results in being 1.08 eV, a value which prohibits spontaneous charge transfer at room temperature.

In the α structure, the gap is opened by only 0.11 eV in a seemingly symmetric fashion. Here, we do not observe an orbital remixing between the HOMO and HOMO-1 levels. Indeed, insofar as the dopant is located farther away from the host with respect to the π structure, featuring a smaller overlap with the polymer orbitals, we can expect a much lower hybridization.

This is in fact confirmed when looking at the electronic transfer integrals, namely

$$\mathcal{H}_\alpha = \begin{array}{c} \\ \text{L} \\ \text{H} \\ \text{H-1} \end{array} \begin{array}{c} \begin{array}{ccc} \text{L} & \text{H} & \text{H-1} \end{array} \\ \begin{pmatrix} -4.80 & 0.19 & -0.029 \\ 0.19 & -5.23 & -0.01 \\ -0.029 & -0.01 & -5.67 \end{pmatrix} \end{array}. \tag{4.2}$$

Together, the reduced hybridization resulting from the HOMO-LUMO transfer integral $t_{HL} = 0.19$ eV and the negligible coupling between the HOMO and HOMO-1, explain the moderate gap opening compared to the π system and the absence of remixing in the occupied levels. The resulting photoemission gap in the α system, including all of the environmental contributions, is found to be 0.18 eV. Such a small value indicates again that the α system may be more suitable for efficient host-dopant charge transfer.

We conclude this analysis by observing that the π and α systems feature very different

donor-acceptor gaps, resulting from a different electrostatic landscape and different host-dopant hybridization strengths. Both systems feature a moderate hybridization, with a slightly more pronounced gap opening in the π system. In the end, the dominating factor responsible for the difference between the donor-acceptor gap in either structure is the disparate electrostatic environment in which the dopant is located, as we discussed in the previous section.

4.4.2 Neutral Excitations

As of now, the emerging scenario for the charge transfer energetics in PBTTT doped with F4TCNQ is that of a substantial gap in the π system impeding spontaneous charge transfer and a moderate gap in the α system, which is however larger than room temperature thermal energy. But the process of charge transfer between host and dopant molecules actually involves the creation of an electron-hole pair. As we discussed in the Preface of this chapter, the binding energy of the formed CT exciton cannot be neglected as it may strongly stabilize its creation.

We therefore now compare the spectrum of the singlet optical excitations of the 3BTTT-F4TCNQ complex for the π and α systems. The two-particle electron-hole states are computed from the Bethe-Salpeter Equation (BSE), starting from the embedded *GW*e | DFTe quasiparticle levels. The product electron-hole basis was constructed by considering all *GW* states within an energy window of ± 10 eV around the gap, and the two-particle BSE Hamiltonian was solved within the Tamm-Dancoff approximation. The latter is considered to be a safe approximation in the limit of full charge-transfer excitations, which are expected to characterize the low-energy region of the spectrum in donor-acceptor complexes. The BSE calculations account for the static dielectric screening of the MM region which is introduced through the renormalization of the bare Coulomb potential within the QM region by means of the MM reaction field, as of use in the Dyson equation for *W* [132, 37]. The polarization of the embedding medium is expected to lead to a redshift of the optical excitation energies, since the surrounding polymer chains will tend to screen the electrostatic field arising from the excitonic states, resulting in a reduced screened Coulomb potential W_{eh} between the electron and the hole (see Sec. 3.3.8).

[132] Duchemin et al. (2018).

Fig. 4.11 presents the two lowest optical excitations for the π and α structures. The hole-averaged electron density (blue) and electron-averaged hole density (red) are presented for both states, along with their energy. A visual inspection of the electron-hole density of the two lowest-energy excitations in the π structure reveals that those are indeed CT transitions from the HOMO and HOMO-1 of the 3BTTT to the LUMO of F4TCNQ. This assignment is confirmed by the detailed BSE exciton compositions summarized in Table 4.2: S_1 is mainly HOMO→LUMO transition while S_2 is mainly a HOMO-1→LUMO transition. Turning now to the excitation energies, the lowest excitation energy of the π system corresponding to a charge transfer between the donor and the acceptor is $S_1 = 0.51$ eV, a value that is prohibitively large for a spontaneous ionization at room temperature. Besides, the expectation value of the screened Coulomb potential for this exciton is $W_{\mathrm{eh}} = 0.86$ eV, reflecting the fact that the electron-hole pair is strongly bound. This is explained by the spatial proximity of the host and dopant which tends to increase their Coulomb interaction, and consequently leads to a large exciton binding energy.

Table 4.2: Analysis of the first BSE optical excitations of the π structure. H/L stand for HOMO and LUMO respectively.

Transition	E (eV)	f_{osc}	W_{eh} (eV)	Composition
S_1	0.51	0.06	0.86	H → L (90%), H - 2 → L(8%)
S_2	0.75	0.44	1.11	H - 1 → L (94%), H - 2 → L (4%)
S_3	1.09	0.11	0.76	H - 2 → L (81%), H → L (8%)

Turning now to the α structure, we observe that the lowest optical excitations are of CT nature as well. The composition of these excitations, summarized in Table 4.3, reveal that the lowest optical excitation is mainly formed of a HOMO-1→LUMO

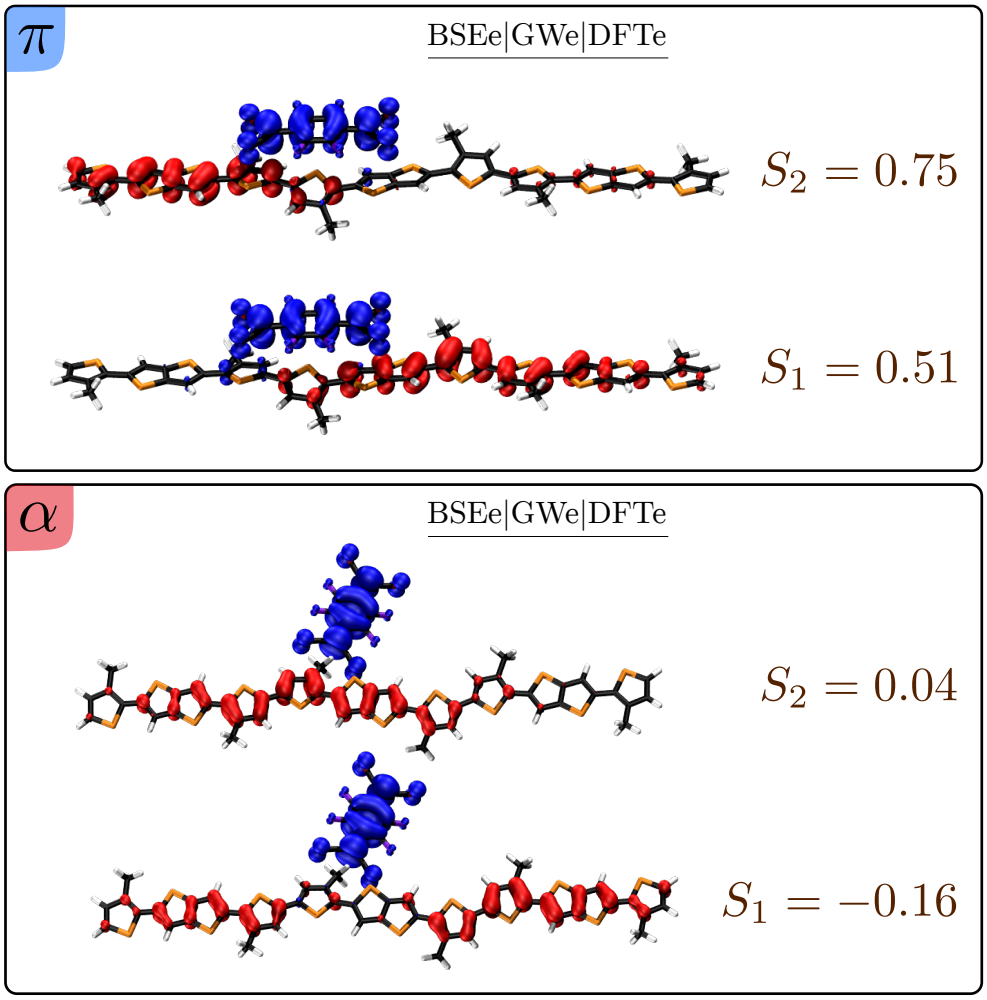

Figure 4.11: Rendering of the electron-averaged hole density (red) and hole-averaged electron density (blue) corresponding to the two lowest-energy optical excitations from embedded BSE calculations for the π-structure (top) and α-structure. All the excitations feature a Charge-Transfer character from the HOMO and HOMO-1 levels to the LUMO level.

transition, while the second is mainly a HOMO→LUMO transition.

Transition	E (eV)	f_{osc}	W_{eh} (eV)	Composition
S_1	-0.16	-0.17	0.79	H − 1 → L (96%), H → L (2%)
S_2	0.04	0.08	1.03	H → L (96%), H − 1 → L (3%)
S_3	0.33	4.10^{-3}	0.80	H − 2 → L (96%)

Table 4.3: Analysis of the first BSE optical excitations of the α structure. H/L stand for HOMO and LUMO respectively.

Let us finally discuss the most striking outcome of our calculations: the first optical excitation S_1 of the α structure is *negative*! Before discussing the interpretation of this ghastly result, we observe that the expectation value of screened Coulomb potential, determining the dominant contribution to the exciton binding energy for CT excitons, is $W_{eh} = 0.79$ eV, which is comparable to that of the lowest optical excitation in the π structure. We explain this similarity with two competing phenomena. The reduced hybridization in the ground state leads to a more localized hole and thus a higher W_{eh} , which is seen to compensate the reduction of W_{eh} arising from the increased distance between the host and the dopant. Recalling that the photoemission gap of the α system was found to be as low as 0.18 eV, such a large binding energy is responsible of this dramatic reduction of S_1.

Such a negative optical excitation energy should be interpreted in light of the fact that

BSE optical excitations are constructed from a many-body perturbation theory starting from an initial ground state obtained from DFT. By their very nature, the neutral quasiparticle excitations of the system are higher in energy than the ground state. Besides, applying the Tamm-Dancoff approximation has the only effect of casting the BSE Hamiltonian into a real symmetric eigenvalue problem [5], thereby imposing real eigenvalues that are not necessarily positive. It follows that the negative excitation energy might reflect an instability of the ground state obtained from spin-restricted Kohn-Sham theory, which has been used as a starting point for the perturbation theory. Irrespectively on the adequacy of the present scheme, we can still conclude that the α system dramatically favors full CT as compared to the π system, even though in the former the polymer and the dopant are located farther away. The topic of the next section will be to investigate the nature of the ground state of the embedded complexes beyond the spin-restricted DFT approach described herein.

5. We remind that the Tamm-Dancoff approximation is obtained by neglecting the resonant and anti-resonant parts of the BSE Hamiltonian, see Sec. 3.3.

4.5 Ground State Charge Transfer from Spin-Unrestricted DFT

In view of the rather elusive outcomes of many-body perturbation theory applied to the α structure, resulting in a negative lowest optical excitation energy, we now address the ground state of the system from spin-unrestricted DFT.

Despite great success, the treatment of spin in non-relativistic DFT features a number of shortcomings. Spin-restricted calculations within Restricted Kohn-Sham (RKS) theory imposing double occupancy of molecular orbitals are typically well-suited to describe closed-shell systems with an even number of electrons. However, notable exceptions exists such as in the description of open-shell systems or of the ground state CT in donor-acceptor complexes.

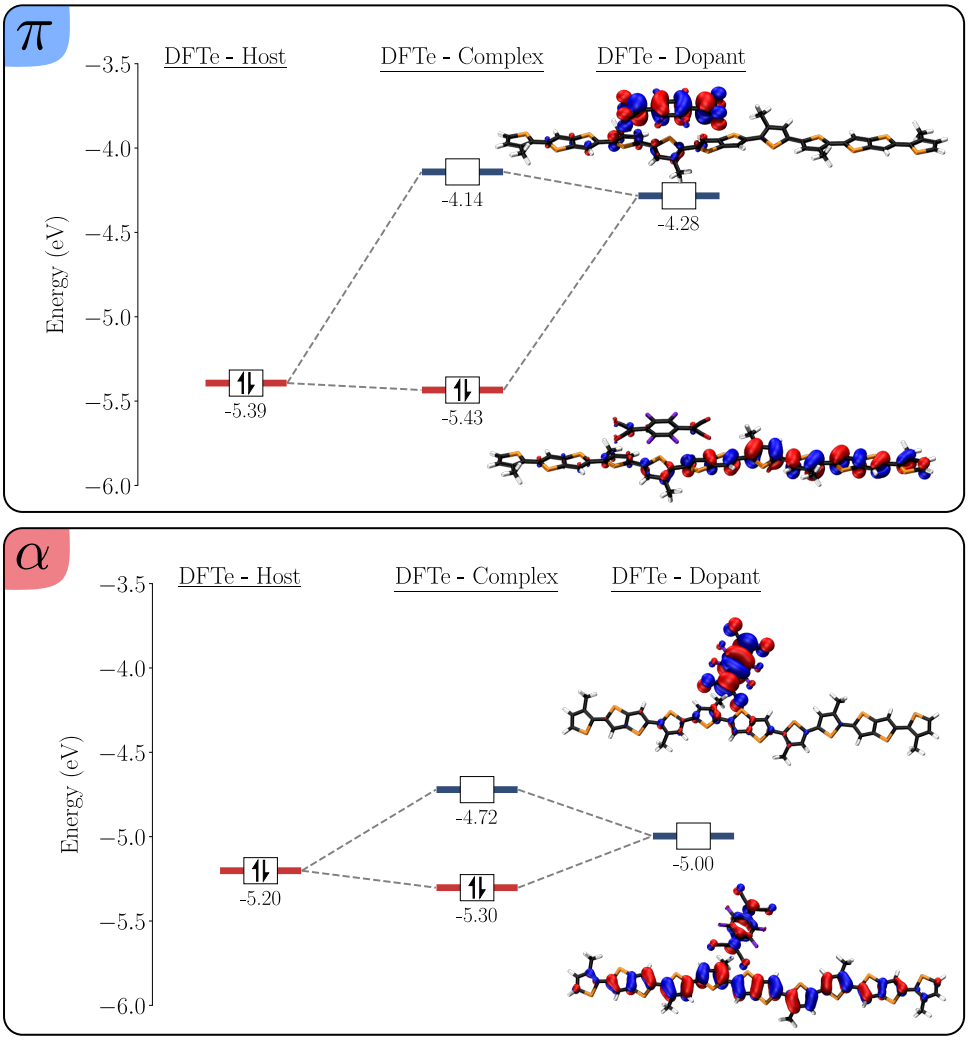

Figure 4.12: Energy levels hybridization between the HOMO of the isolated 3BTTT and the LUMO of the isolated F4TCNQ in the spin-restricted RKS approach for the π structure (top) and the α structure (bottom). Every molecular orbital is either doubly occupied or empty, impeding the formation of CT states.

Turning to the systems we are investigating, the host-dopant complex is closed-shell, and the HOMO is mostly localized on the host while the LUMO is localized on the dopant. Consequently, a charge transfer in the ground state is only possible in the RKS approach if the host and the dopant largely hybridize so as to have an occupied level partially extending over the dopant. However, the single Slater determinant obtained along RKS theory has intrinsic difficulties in describing states characterized by a large

or integer electron transfer. Indeed, in this case the transfer of one electron from the donor to the acceptor would be achieved only in the case of supramolecular orbitals formed by an equal weight of frontier levels of the donor and of the acceptor. This would mean that two half-electrons (both the spin up and down) are transferred from the donor to the acceptor. We emphasize that this picture is qualitatively different from a ground state in which a single electron is transferred to the acceptor molecule.

We quantify this assertion by showing in Fig. 4.12 the frontier energy level hybridization in the host-dopant complex between the isolated 3BTTT and the isolated F4TCNQ. These data are obtained from DFT/MM calculations at the PBEh(α_x^*)/6-311G* level, within the RKS approach. As noted previously, the HOMO of the complex is formed by a combination of the HOMO and HOMO-1 of the isolated host, while the LUMO of the complex corresponds to that of the isolated host. As such, the hybridization between the host and the dopant results in a ground state with moderate CT character in both systems, characterized by a net charge on the dopant Q_{dop} of -0.24 and -0.22 for the π and the α system respectively.

6. All subsequent DFT calculations are performed at the PBEh(α_x^*)/6-311G* level.

In order to alleviate the spurious effects of the RKS approach, we now characterize the CT nature of the ground state of the 3BTTT-F4TCNQ complex by means of spin-unrestricted DFT calculations [6]. The Unrestricted Kohn-Sham (UKS) approach refers to the independent treatment of each spin channel within DFT, namely, spin-up and spin-down electrons can occupy different spatial regions. This allows for a more appropriate description of open-shell and CT systems. This comes however at the price of the ground state not being anymore an eigenstate of the squared spin operator $\hat{\mathbf{S}}^2$, a problem called spin contamination [7] (see Sec. 3.2.6).

7. We note that the measure of spin contamination in DFT has only a limited relevance insofar as there is little theoretical justification to compute $\hat{\mathbf{S}}^2$ from the Kohn-Sham orbitals.

Table 4.4: Nature of the ground state of the α system as obtained from different spin treatments in DFT. The total QM/MM energy is measured from the reference RKS-MM calculation α_1, when applicable. The CT nature of the ground state is measured by the total charge on the dopant Q_{dop}.

Method	Spin state	Initial Guess	ΔU (eV)	Q_{dop} (e)
α_1: RKS-MM	Singlet	Atomic Densities	0.0	-0.24
α_2: UKS-MM	Singlet	Atomic Densities	~ 0	-0.23
α_3: UKS	Triplet	Atomic Densities	N.A.	-0.95
α_4: UKS	Singlet	α_3	N.A.	0.02
α_5: UKS-MM	Singlet	α_3	-0.58	-0.96

We will ascertain the nature of the ground state from the net charge on the dopant molecule Q_{dop}, as calculated from the Mulliken partial atomic charges [111], which is used as a measure of the degree of charge transfer in the system. Further, the stability of the ground state is assessed from the total ground-state energy of the QM/MM system U, given by the formula of Eq. 3.186 derived in Sec. 3.5.

[111] Mulliken (1955).

We begin by discussing the α system. Table 4.4 compares the embedded UKS calculation, named α_2, with the reference calculation, named α_1, which corresponds to the embedded RKS that has been discussed previously. The reference α_1 has a total energy variation $\Delta U = 0$ by definition and a net charge on the dopant $Q_{dop} = -0.22\,e$, reflecting a moderate charge transfer in the ground state. The outcome of the same calculation within the UKS approach results in the same ground state within numerical accuracy. Indeed, the total energy variation is only $\Delta U = 1.2\,meV$ and the net charge on the dopant $Q_{dop} = -0.22\,e$.

However, we must note that the DFT calculations proceed through a self-consistent

field (SCF) procedure in which the starting Hamiltonian is formed from an initial guess of the density, e.g. a superposition of atomic densities. Such a choice may well impact the result in the case of local minima in the spin-orbital Hilbert space, and we hence compare the CT nature and relative stability of DFT calculations with different starting points.

The α_3 gas-phase calculation denotes a full-CT initial guess, which is achieved by enforcing a triplet spin state[8]. It features a charge on the dopant Q_{dop} =−0.95 e, which slightly deviates from an integer charge because of intermolecular hybridization. Curiously, when a spin-singlet UKS calculation is started from the α_3 initial guess (α_4 calculation), the system evolves towards an essentially neutral dopant state. Conversely, upon including the interactions with the MM environment (α_5 calculation), the system instead maintains its full-CT character in the ground state, characterized by Q_{dop} =−0.95 e, and is importantly stabilized with respect to the reference spin-restricted ground state by a total energy variation of 0.58 eV. This energy is well above numerical accuracy and translates into a true stabilization of the system upon charge transfer. The ground state features however some spin contamination, characterized by $\langle \hat{\mathbf{S}}^2 \rangle = 1.16\hbar^2$.

8. The total energies of gas-phase DFT calculations are not comparable to those of embedded DFT calculations, hence the value of ΔU is not available.

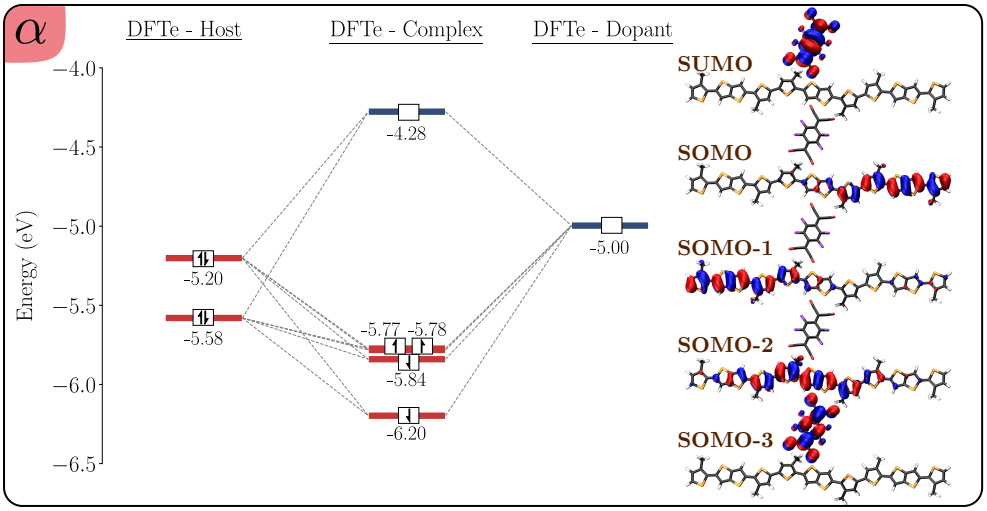

Figure 4.13: Energy levels hybridization between the SOMO of the isolated 3BTTT and the SUMO of the isolated F4TCNQ in the spin-unrestricted UKS approach for the α system. We witness the advent of a full CT ground-state characterized by a SOMO-3 exclusively located on the dopant. Note that the SOMO and SOMO-1 levels are quasi-degenerate spin-up levels.

The specific nature of this charge transfer ground state can be determined by considering the hybridization between the spin-unrestricted electronic levels of the isolated host and dopant, as illustrated in Fig. 4.13. Because considering either 3BTTT or F4TCNQ in the QM region impedes quantum-mechanical intermolecular interactions between the two species, the spin-unrestricted electronic levels of the isolated molecules correspond exactly to those obtained within the spin-restricted approach.

We observe an unequivocal separation between molecular orbitals pertaining to the host or the dopant, which reflects a lack of orbital hybridization. Of particular interest is the SOMO-3 level, which is fully localized on F4TCNQ, distinctly expressing the full CT scenario quantified by a near-integer net charge on the dopant. This analysis confirms that the full-CT ground state is the most stable solution found for the α system.

We now apply this same procedure to the π system. Table 4.5 presents the net charge on the dopant Q_{dop} and the total QM/MM energy relative to the reference RKS embedded calculation π_1 for different DFT starting points. The π_1 calculation

features a charge on the dopant $Q_{dop} = -0.24e$, reflecting here as well a partial CT in the ground state. Again, the outcome of the same calculation within the UKS approach results practically in the same ground state, as seen by a negligible total QM/MM energy difference ΔU well below 1 meV and a charge on the dopant that is equal to the reference one.

Table 4.5: Nature of the ground state of the π system as obtained from different spin treatments in DFT. The total QM/MM energy is measured from the reference RKS calculation π_1, when applicable. The CT nature of the ground state is measured by the total charge on the dopant Q_{dop}.

Method	Spin state	Initial Guess	ΔU (eV)	Q_{dop} (e)
π_1: RKS-MM	Singlet	Atomic Densities	0.0	-0.27
π_2: UKS-MM	Singlet	Atomic Densities	~ 0	-0.23
π_3: UKS	Triplet	Atomic Densities	N.A.	-0.92
π_4: UKS	Singlet	π_3	N.A.	-0.25
π_5: UKS-MM	Singlet	π_3	-0.07	-0.70

The π_3 calculation is that of the full-CT initial guess, which features by construction a large charge on the dopant $Q_{dop} = -0.90e$. This initial guess is then fed to the π_4 and π_5 singlet calculations. We observe that the absence of the environment in the π_4 calculation results in a moderate CT ground state with $Q_{dop} = -0.22e$, whereas it is only when the environment is included that a different ground state is obtained. Indeed, the π_5 calculation results in a large, but not integer, charge transfer with a net charge on the dopant $Q_{dop} = -0.68e$ and is only slightly lower in energy than the reference π_1, with $\Delta U = 0.07$ eV. This difference is probably within the accuracy of the method. The ground state is characterized by a spin contamination $\langle \hat{S}^2 \rangle = 0.78\hbar^2$. In general, the lower propensity for full CT in the π system can be ascribed to the less favorable electrostatic landscape in which the dopant is located with respect to the α system and to the larger intermolecular couplings, stabilizing hybridized states.

The sizable molecular orbital hybridization in the π system can be seen in Fig. 4.14. Indeed, the large mixing between the host's SOMO and the dopant's SUMO leads to a partial localization on both 3BTTT and F4TCNQ in the SOMO and SOMO-2. These results clearly explain the partial CT that was quantified by a net charge on the dopant $Q_{dop} = -0.68e$ in terms of a strong host-dopant hybridization.

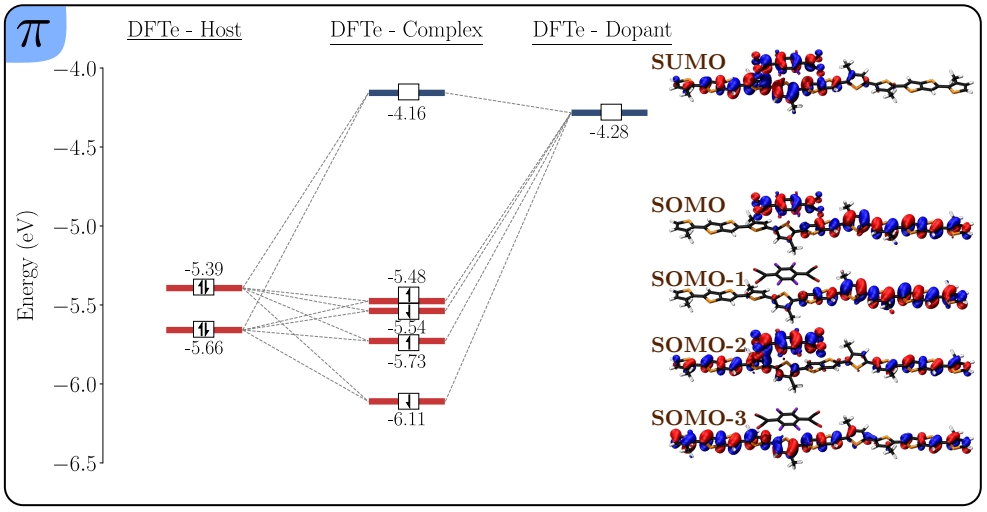

Figure 4.14: Energy levels hybridization between the SOMO of the isolated 3BTTT and the SUMO of the isolated F4TCNQ in the spin-unrestricted UKS approach for the π system. We observe a sizable hybridization between the host and dopant molecular orbitals, leading to a partial CT in the ground state originating from the SOMO and SOMO-2 levels being partially located on the dopant.

The results presented so far report qualitatively different ground states of the α and π systems that are both characterized by a finite spin contamination. Eigenvalues of $\hat{\mathbf{S}}^2$ in the spin-unrestricted ground state close to unity for both systems signal a similar proportion of singlet and triplet electronic configurations, which we recall has a charge transfer nature. Singlet and triplet configurations are generally nearly degenerate in pure intermolecular CT states, because of the very small intermolecular overlap. A sizable mixing is therefore expected to occur in spin-unrestricted calculations. Such a spin contamination is however not expected to have an appreciable impact on the charge density, which is in both cases characterized by a hole on the donor and an electron on the acceptor, nor or the ground state energy, the two spin configurations being nearly isoenergetic.

4.6 SUMMARY

IN THIS CHAPTER, we have discussed the relationships between nanostructure and electronic properties in the technologically relevant organic polymer PBTTT doped with the strong electron acceptor F4TCNQ. By applying diverse multiscale theoretical tools such as Density-Functional Theory and Many-Body Perturbation Theory combined with Micro-Electrostatics, we addressed host-dopant charge transfer energetics when the dopant is stacked on top of the π-conjugated chains of the polymer or in between the Alkyl chains.

The electrostatic landscape of the PBTTT crystal facilitates electron transfer when the dopant is in the Alkyl chains, a feature that is reflected by the reduced photoemission gap in the α system as compared to the π system. Consequently, in the latter the lowest charge-transfer excitation energy was found to be 0.51 eV, a value that is too large for spontaneous charge transfer at room temperature. Conversely, a negative optical excitation energy in the α system signaled that the perturbation theory was probably applied to the wrong ground state, suggesting a possible charge-transfer ground state [9].

By lifting the restriction of doubly occupied spin-orbitals in ground state calculations, we demonstrated that the true ground state of the α system is a full charge-transfer one, while that of the π system features only a fractional charge transfer. This is explained by the interplay between the less favorable electrostatic landscape and the larger intermolecular hybridization taking place in this structure. These findings explain the larger conductivity boost observed experimentally when dopants are placed within the Alkyl region of PBTTT, since in this case the charge transfer is complete, thus resulting in a higher carrier concentration.

9. In the present version of the Fiesta code, *GW* and BSE calculations cannot be performed on top a spin-unrestricted ground state. This could be a route of development to achieve in the near future.

Chapter 5

Doping of Organic Semiconductors by Polarization Catastrophe

5 Doping of Organic Semiconductors by Polarization Catastrophe

One becomes wise only in measures, as he goes through his own insanity.

Alejandro Jodorowsky

Contents

A part of this study has been published in Advanced Materials [151]. It also has been awarded the best short presentation award at the European Material Research Society (EMRS) Spring meeting 2021, Symposium C: Doping and charge transport processes in organic and hybrid materials for energy applications.

[151] Comin et al. (2021).

IN THE PREVIOUS CHAPTER we have focused on the first step in the doping process, namely the Charge Transfer (CT) between the dopant impurity and the host semiconductor. By specifically addressing the dopant ionization energetics in Organic Semiconductor (OSC) in the infinite dilution limit, we highlighted the important role of the exciton binding energy in the stabilization of charge-transfer states.

The CT between the host and dopant is governed by the energy $E_{CT}(r)$ for creating an electron-hole pair at finite distance r in the material

$$E_{CT} = U^* - U^0 = \Gamma - E_b(r), \tag{5.1}$$

where U^* is the total energy of the system with the added electron-hole pair and U^0 is the ground-state total energy. This energy E_{CT} is also expressed as the difference between the transport gap $\Gamma = IP - EA$, i.e. the energy required to create an electron-hole pair at infinite distance, and the exciton binding energy $E_b(r) > 0$. Here, the role of E_b is clear: a strong binding energy favors host-dopant charge transfer, even possibly resulting in a CT ground state when E_{CT} is negative.

However, the charge on the host must be released in order to participate to conduction, which constitutes the second step of the doping process. The energy barrier it must overcome is precisely the exciton binding energy, which for a full-CT exciton is given by the screened Coulomb potential $W_{eh} = 1/(\varepsilon r)$. Photoemission measurements [16] and theoretical calculations [37, 38, 39] set this binding energy in the 400-700 meV range in the low-doping regime, a value that is too large to permit a significant release of free carriers at room temperature.

[16] Gaul et al. (2018).

[37] Li et al. (2017).

[38] Li et al. (2019).

[39] Privitera et al. (2020).

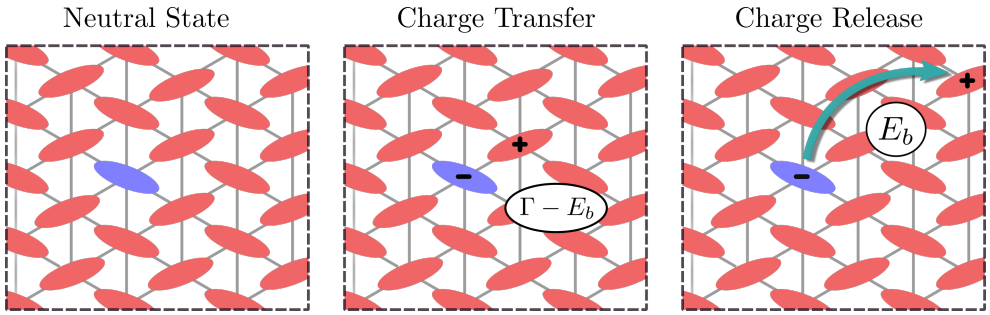

Figure 5.1: Illustration of the elementary steps of the doping process in organic semiconductors. An acceptor dopant molecule (violet) is inserted in a host organic crystal (magenta), mimicking the herringbone structure of Pentacene. The energy barrier for charge transfer is that of creating the electron-hole pair $\Gamma - E_b$. The energy barrier for releasing the doping-induced charge is the exciton binding energy E_b.

A large exciton binding energy is therefore double-edged: it is a *blessing* in the sense that it favors the spontaneous charge transfer and a *curse* because it forms localized bound excitons that cannot contribute to transport. This is at odds with the observation of an important conductivity increase upon doping OSCs, evidencing the fact that doping-induced charges are indeed released. Then, how can we reconcile the observed charge release with the formation of strongly bound excitons upon doping ?

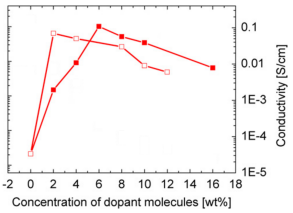

Figure 5.2: Conductivity of Pentacene doped with F4TCNQ (closed rectangles) and F6TCNNQ (open rectangles) as a function of the doping load. Adapted from [29].

This question calls for a rationale for the charge release phenomena taking place in doped OSCs. Experimental data for molecular OSCs and polymers show that a boost in conductivity is achieved at doping loads of about $5-10$ % (see Fig. 5.2), that is, at concentrations that are orders of magnitude larger than in inorganic semiconductors [29, 137, 32]. Several studies have shown that the conductivity follows a thermally activated behavior [33, 34, 35, 36], namely

$$\sigma(T) = \sigma_\infty \exp\left(-\frac{E_a}{k_B T}\right), \tag{5.2}$$

where σ_∞ is the conductivity at $T = \infty$, k_B is Boltzmann's constant and E_a is the activation energy, which has been related to the Coulomb interaction between an ionized dopant and the charge injected in the semiconductor [36]. Moreover, experiments carried on both p and n-doped OSCs have shown a universal tendency for a conspicuous and superlinear reduction of the activation energy with dopant concentration [36, 34, 35, 40]. This set of evidences, together with the high loads needed to boost conductivity, points to a collective depinning mechanism for charge carriers taking place upon increasing doping concentration.

The first step towards elucidating this mechanism is to notice that the low-lying CT excitations of host-dopant complexes may result in a sizable increase of their polarizability, as discussed in Sec. 5.1. Besides, at the doping loads commensurate with the conductivity boost in OSCs, dopants are only a few lattice constants apart and their interactions cannot be neglected. The presence of highly polarizable and interacting host-dopant complexes could thus ensue a considerable enhancement of the dielectric constant of the doped OSC, which in turn would effectively reduce the Coulomb energy barrier required to free the doping-induced charges.

In order to investigate this hypothesis, we study in this chapter the effects of collective screening phenomena on the release of doping-induced charges beyond the infinite dilution limit. Following a multiscale approach, we will assess the dielectric properties of doped OSCs at finite doping concentrations. To this aim, we first evaluate *ab initio* the polarizability associated with low-energy CT degrees of freedom in host-dopant complexes. By taking the prototypical case of a F4TCNQ dopant in Pentacene, our QM/MM calculations show that the CT polarizability can exceed 500

[137] Mendez et al. (2015).

[32] Pingel and Neher (2013).

[33] Maennig et al. (2001).

[34] Shen et al. (2003).

[35] Olthof et al. (2012).

[36] Schwarze et al. (2019).

[40] Zhang and Kahn (2018).

$Å^3$, i.e. 10 times that of a single host molecule.

This information is then used to build a microscopic model for doped OSCs at finite concentration that allows to obtain the dielectric properties of doped OSCs as a function of the doping load. These calculations reveal a large enhancement of the relative permittivity upon doping, with an order of magnitude increase at 8% concentration, which implies a drastic suppression of Coulomb energy barriers for the release of free charge carriers. The origin of this enhancement is rooted in the collective response of highly-polarizable host-dopant complexes in a disordered system, whose susceptibility increases upon doping and diverges at the approach of the dielectric catastrophe.

5.1 Charge-Transfer Polarizability of Host-Dopant Complexes

Whenever low-lying CT states appear upon doping OSCs, as we have discussed in Chapter 4, we might expect a strong contribution to the polarizability. This stems from the standard Sum-Over-States (SOS) formulation of the polarizability, and forms the basis upon which this study relies. Our analysis thus starts with the assessment of the charge-transfer polarizability of host-dopant complexes in different systems. We aim to describe both the magnitude of the polarizability enhancement in a specific prototypical system, as well as the generality of this phenomenon. As such, we first study the polarizability of F4TCNQ-doped Pentacene from first principles calculations, and then move to a more general and system-independent evaluation of this process on the basis of model Hamiltonians.

5.1.1 Definitions of Polarizability

The polarizability tensor $\underline{\alpha}$ is defined in terms of derivatives of the total dipole μ or the total energy of the system U with respect to the applied field, namely

$$\alpha_{ij} = \frac{\partial \mu_i}{\partial F_j^{ext}}\bigg|_0 = -\frac{\partial^2 U}{\partial F_i^{ext}\partial F_j^{ext}}\bigg|_0, \tag{5.3}$$

where $\mathbf{F}^{ext}$ is a uniform and static external field. The application of such a field to a quantum system described by the Hamiltonian $\hat{H}_0$ results in a perturbed Hamiltonian

$$\hat{H} = \hat{H}_0 - \boldsymbol{\mu}\cdot\mathbf{F}^{ext}. \tag{5.4}$$

Upon applying second-order perturbation theory one obtains the ground-state energy of the perturbed Hamiltonian as

$$U = E_0 - \langle\phi_0|\boldsymbol{\mu}\cdot\mathbf{F}^{ext}|\phi_0\rangle + \sum_{i,j=x,y,z} F_i^{ext}F_j^{ext}\sum_{n>0}\frac{\langle\phi_0|\hat{\mu_i}|\phi_n\rangle\,\langle\phi_n|\hat{\mu_j}|\phi_0\rangle}{E_0 - E_n}, \tag{5.5}$$

where E_n and ϕ_n are respectively the many-body energies and eigenstates of the unperturbed Hamiltonian $\hat{H}_0$. The identification with Eq. (5.3) provides the sum-over-states expression for the static polarizability

$$\alpha_{ij} = 2\sum_{n>0}\frac{\langle\phi_0|\mu_i|\phi_n\rangle\,\langle\phi_n|\mu_j|\phi_0\rangle}{E_n - E_0}. \tag{5.6}$$

The polarizability of Eq. (5.6) is determined by two quantities: the transition dipole moment amplitudes and the corresponding transition energies. Therefore, we can expect low-energy and optically-allowed CT excitations to strongly contribute to the polarizability of a CT complex.

In order to account for the environmental polarization effects, which can possibly stabilize CT states, we will adopt a Quantum/Classical (QM/MM) approach for the calculation of the polarizability. The SOS expression above does not apply to Density-Functional Theory (DFT) because it is a ground-state formalism: the Kohn-Sham (KS) eigenstates pertain to a fictitious non-interacting system and are conceptually different from the many-body states involved in Eq. (5.6).

Instead, the Coupled-Perturbed Self-Consistent Field (CP-SCF) equations [1] provide

1. The CP-SCF scheme is a variation on standard Density Functional Perturbation Theory (DFPT) for hybrid functionals, where the variation of the self-consistent field is non-local. It is the approach implemented in quantum chemistry codes to study linear and nonlinear responses to perturbations.

an analytical way to compute the polarizability within the DFT framework [152]. While the CP-SCF scheme permits the inclusion of an external field due to a fixed electrostatic environment, such a procedure misses the important screening effects that arise from a truly polarizable environment.

On the other hand, Eqs. (5.3) provides two different routes for the calculation of the polarizability in terms of the total dipole or the total energy which can be adopted in the context of embedded DFT. By computing such derivatives from finite differences between separate DFT/ME calculations for various intensities of the external field, the reaction of the polarizable environment specific to each field value is accounted for. In particular, computing the polarizability as the first derivative of the total dipole is preferable as it is more accurate, numerically stable and efficient.

5.1.2 Polarizability of F4TCNQ-doped Pentacene

We now assess the CT states contribution to the linear polarizability of a host-dopant complex in Pentacene doped with F4TCNQ acceptor with our accurate and original hybrid QM/MM scheme. The system is partitioned into an inner QM region, described at the DFT level, and an outer MM environment treated at the classical level with a polarizable Micro-Electrostatic (ME) model of atomistic resolution [2], as depicted in Fig. 5.3.

2. DFT and ME calculations have been performed with the ORCA 4.2 [54] and the MESCal code [107], respectively.

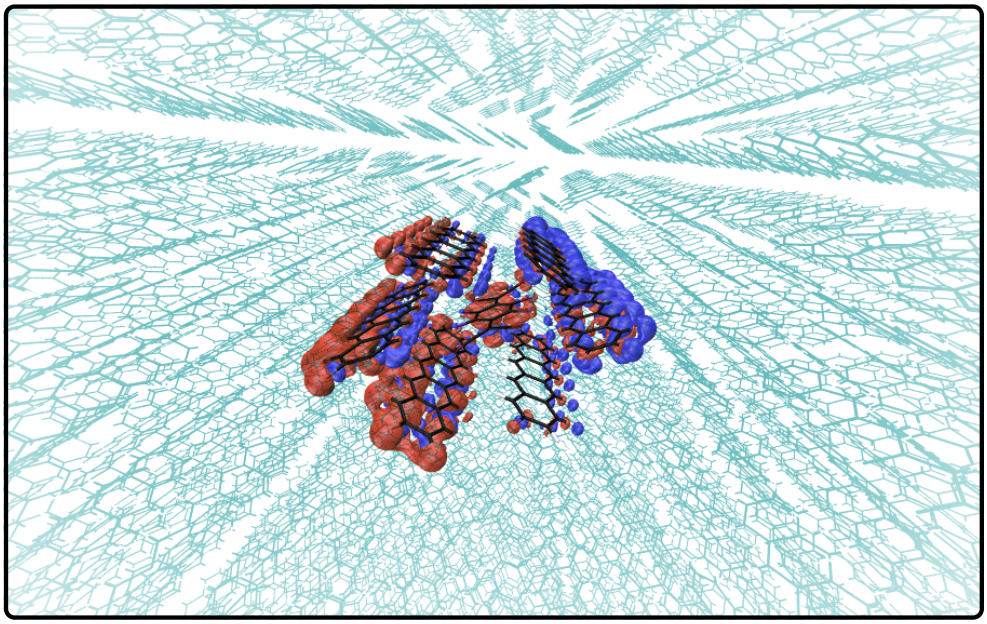

Figure 5.3: Illustration of a QM/MM calculation of a F4TCNQ-Pentacene complex (QM region) in the host Pentacene crystal (MM, cyan wireframe representation). The red/blue surface depicts the electron density induced by an electric field, $\Delta n(F^{ext}) = n(F^{ext}) - n(0)$, with isocontour drawn at negative/positive value. The large polarizability of the complex is determined by such a displacement of the electron density within the QM region from the Pentacenes on one side of the dopant to those on the other side.

The QM region consists of the central F4TCNQ dopant replacing a Pentacene molecule at the same center of mass and orientation, with its first shell of six Pentacene neighbors. We considered the triclinic Pentacene crystal structure obtained from X-ray diffraction data for the vapor-grown polymorph [133] whose structural parameters are given in Table 5.1. An accurate description of the relative energies of the frontier molecular orbitals, i.e the Pentacene HOMO and the F4TCNQ LUMO, is essential to correctly describe the intermolecular orbital hybridization in the solid-state. This is indeed a crucial factor controlling the ground-state and excitonic properties associated with intermolecular CT degrees of freedom. Considering the importance of the above factors and the strong variability of energy levels with DFT functionals, our calculations employed the tuned PBEh(a_x) hybrid functional in conjunction with the cc-pVTZ basis set. The amount of exact exchange a_x was tuned to reproduce the gap

[133] Siegrist et al. (2001).

obtained from accurate embedded many-body G*W* calculations [37], resulting in a value $a_x^* = 0.4$, namely 40% of exact exchange. The geometry of the QM region has been relaxed at the PBEh(a_x^*)-D3/cc-pVDZ level, accounting for dispersion (Van der Waals) and electrostatic interactions with the MM region.

The surrounding crystalline environment forming the MM region is composed of the Pentacene crystal up to a radius of $R = 50$ Å. It is described at the ME level [107], providing the proper electrostatic embedding and dielectric screening to the QM region. The atomic charges and molecular polarizability tensor of the Pentacene molecule, needed to parameterize the ME model, have been computed at the PBEh(a_x^*)/cc-pVDZ level with the Electrostatic Potential Fitting (ESP) [113] and CP-SCF schemes respectively.

a	6.265 Å
b	7.786 Å
c	14.511 Å
α	76.65°
β	87.50°
γ	84.61°

Table 5.1: Parameters of the triclinic crystal structure of Pentacene [133].

[113] Besler et al. (1990).

Then, the CT polarizability is computed within the DFT/ME framework as a finite-field derivative of the dipole moment of the QM region with respect to the applied field, taken from separate DFT/ME calculations at various external field strengths. Each DFT/ME calculation follows an iterative scheme which consists in coupled DFT and Micro-Electrostatics (ME) calculations: each subsystem is relaxed in the field of the other until the total energy is converged, which usually takes $\sim 5-6$ QM/MM iterations. This allows to obtain the electron density of the QM region in the self-consistent field of permanent and induced multipoles in the MM environment.

The polarizability is computed in the ab crystallographic plane of the Pentacene crystal, where the intermolecular overlap is sizable and thus where the applied field can induce a significant charge transfer. This results in an in-plane polarizability computed as

$$\underline{\boldsymbol{\alpha}} = \frac{\partial \boldsymbol{\mu}}{\partial \mathbf{F}^{\text{ext}}} \simeq \frac{1}{2F^{\text{ext}}}\left[\boldsymbol{\mu}\left(\mathbf{F}^{\text{ext}}\right) - \boldsymbol{\mu}\left(-\mathbf{F}^{\text{ext}}\right)\right], \tag{5.7}$$

where $\boldsymbol{\mu}$ is the induced dipole of the QM region, i.e. the differential dipole with respect to the permanent dipole at $F^{\text{ext}} = 0$. The external field magnitude must be chosen as to be in the linear response regime while ensuring a negligible numerical noise, which was attained at $F = 10^{-4}$ a.u. along either x or y axes.

System	Method	α (Å^3)
Pentacene	DFT gas	47
F4TCNQ	DFT gas	33
F4TCNQ-pentacene	DFT gas	418
F4TCNQ-pentacene	DFT/ME	804
F4TCNQ-pentacene, CT	DFT/ME	534
F4TCNQ-pentacene, CT	CT model	357

Table 5.2: Polarizability of Pentacene and F4TCNQ molecules and their complex, calculated at the DFT level in the gas phase and in the solid state (QM/MM embedding). Gas phase polarizabilities have been computed with the CP-SCF scheme, which has been used to validate gas phase calculations with finite differences. The last two lines report the polarizability associated to intermolecular CT degrees of freedom, calculated at QM/MM level and with a generalized Mulliken model for intermolecular CT (vide infra).

The results in Table 5.2 reveal that the polarizability of a host-dopant complex in the gas-phase (418 Å^3) is larger than the sum of the molecular polarizabilities (315 Å^3), testifying a contribution from intermolecular CT interactions. This polarizability enhancement originates from the field-induced electron transfer between the Pentacene molecules located in different positions with respect to the dopant, see Fig. 5.3. Physically, this is a CT contribution to the polarizability of the complex.

Moreover, the polarizability of the complex roughly doubles when accounting for

the MM environment, an observation that can be explained in terms of the screening of the dipolar field generated by the induced charge density (Δn, shown in Fig. 5.3), or equivalently as a result of the stabilization of CT excitations by the polarizable environment. This polarizability enhancement by the environment is consistent with the SOS formula of Eq. (5.6), since the polarization energy reduces the energy of the CT states with respect to the ground-state, as compared to the situation where the complex is in the vacuum.

The full QM/MM polarizability of the host-dopant complex, $\alpha_{cpx} = 804$ Å^3, accounts for both an intramolecular and an intermolecular CT contribution. We obtain the CT contribution to α_{CT} as the difference between the full polarizability and the one computed upon neglecting the possibility for host-dopant CT. To such an aim, we design two specific DFT/ME polarizability calculations, in which we include in the QM region either the F4TCNQ dopant, or the 6 neighboring Pentacene molecules. The omitted molecules are however included in the MM region. The CT polarizability is finally computed as

$$\alpha_{CT} = \alpha_{cpx} - (\alpha_{F4TCNQ} + \alpha_{6PEN}). \tag{5.8}$$

This results in a leading CT contribution of $\alpha_{CT} = 534$ Å^3, which is above 10 times that of a single Pentacene molecule, revealing the very large contribution to the polarizability of doped OSCs sourced from low-lying CT excitations.

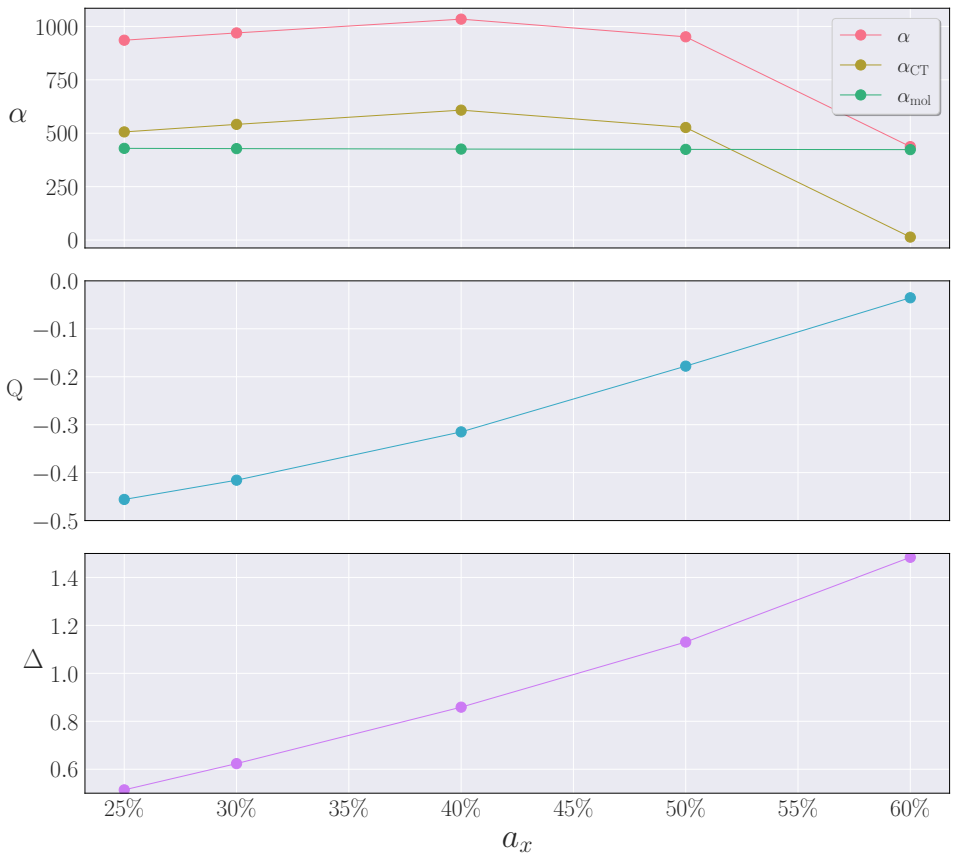

Figure 5.4: Dependence of the polarizability α (Å^3), charge Q on the dopant (e) and gap Δ (eV) of the Pentacene-F4TCNQ complex on the amount of exact exchange a_x in the PBEh DFT functional. The calculations are performed with a PCM embedding with $\varepsilon = 3$ and the cc-pVTZ basis set. The figure reports the Mulliken atomic charges and the in-plane polarizability in the Pentacene crystal ab plane.

We conclude this section by discussing the robustness of the results reported in Table 5.2 against the basis set and functional employed. In Table. 5.3, the value of the gas-phase polarizability of the 6-Pentacene + 1-F4TCNQ complex is reported with respect to the amount of exact exchange in the PBEh(a_x) functional, for a double- (cc-pVDZ) and triple-zeta (cc-pVTZ) basis sets. The polarizability is to a very good

approximation stable in a broad range of a_x around the *GW*-tuned ($a_x^* = 0.4$) value employed in all calculations. Moreover, we observe that the effect of the basis is negligible, ensuring that our results, i.e. the polarizability, are close to the complete basis set limit.

a_x (%)	cc-pVDZ	cc-pVTZ
25	425 (401)	444 (417)
40	429 (401)	449 (418)
50	423 (396)	441 (412)

Table 5.3: Dependence of the gas-phase polarizability (Å^3) of the Pentacene-F4TCNQ complex on the amount of exact exchange in the PBEh DFT functional and the cardinality of the basis set. The table reports the in-plane polarizability $(\alpha_{xx} + \alpha_{yy})/2$, the values between parentheses are the isotropic values.

We further report in Fig. 5.4 the polarizability α, Kohn-Sham gap Δ and charge on the dopant Q of the Pentacene-F4TCNQ complex as a function of the amount of exact exchange a_x using a Polarizable Continuum Model (PCM) embedding with $\varepsilon = 3$. The latter has been employed in these complementary calculations for describing the environmental screening in place of the atomistic ME model. The intramolecular polarizability α_{mol} in this figure is obtained as the sum of the polarizabilities of the individual molecules, computed in PCM with PBEh(a_x) functionals.

These data show that the HOMO-LUMO gap of the complex strongly depends on the functional, with an approximately linear increase with a_x. A fractional negative charge develops on the dopant upon shrinking the gap, as a result of the hybridization between the dopant's LUMO and the HOMOs of Pentacene molecules. The CT polarizability, $\alpha_{CT} = \alpha_{cpx} - \alpha_{mol}$, remains again essentially stable over a wide range of a_x value around the *GW*-tuned value $a_x^* = 0.4$. α_{CT} drops by two orders of magnitude at large a_x, i.e. when the gap is unphysically large and the intermolecular ground-state CT is negligible.

5.1.3 Polarizability of host-dopant complexes

We can now ask ourselves the following question: Is the remarkably large value of the CT polarizability that we found for the F4TCNQ-Pentacene complex only a specific case of the system, or is it a general characteristic of doped OSCs ? In this section, we address this question in order to go beyond the paradigmatic case of F4TCNQ-Pentacene. We will consider in particular the generalized Mulliken model [3] as a complementary approach to assess the polarizability associated to intermolecular CT degrees of freedom. Such a model also allows to play with the parameters that affect the polarizability, namely the donor's Highest Occupied Molecular Orbital (HOMO), the acceptor's Lowest Unoccupied Molecular Orbital (LUMO) and the electron-hole binding energies, allowing to explore the generality of the host-dopant complex polarizability enhancement in various physical situations. We will see that this effect is indeed general, so much that it will provide a solid basis for the subsequent investigations of the dielectric properties of doped OSCs.

3. This model has been briefly introduced in Sec. 4.1.

A. Mulliken Model Definition

The Mulliken model provides a general framework to describe intermolecular CT degrees of freedom based on a Valence-Bond picture accounting for a minimal set of frontier orbitals (namely the HOMO and LUMO of the donor and acceptor species) and the explicit account of the excitonic electron-hole interaction. This model has been introduced by Mulliken to describe optical excitations in molecular donor-acceptor complexes in solution [153, 154]. This model and its generalization have been widely applied to describe the CT physics in molecular systems [98], recently also in the context of molecular doping by Li et al [131].

[153] Mulliken (1952).

[154] Mulliken and Person (1969).

[98] Macdonald et al. (2020).

[131] Li et al. (2018).

In this model, the dopant molecule (DOP) in the host organic semiconductor (OSC) is described in terms of a Hamiltonian written on the basis of low-energy singlet electronic configurations, corresponding to the neutral state $|0\rangle$, and CT states, $|i\rangle$, in which the DOP molecule accepts (donates) an electron from (to) the OSC molecule i. The electronic Hamiltonian reads:

$$H = \sum_{i\in OSC} \epsilon_i^{CT} |i\rangle\langle i| + \sum_{i\in OSC} t_i^{CT}\big(|0\rangle\langle i| + |i\rangle\langle 0|\big) + \sum_{\langle i,j\rangle\in OSC} t_{ij}\big(|i\rangle\langle j| + |j\rangle\langle i|\big) \quad (5.9)$$

where t_i^{CT} and t_{ij} are the host-dopant and host-host charge-transfer couplings, respectively. The energy of the basis CT states is

$$\epsilon_i^{CT} = \Gamma - V_i, \quad (5.10)$$

where Γ is the host-dopant gap and V_i is the host-dopant distance dependent screened Coulomb binding energy. The model applies to both p-doping, where $\Gamma = EA_{DOP} - IP_{OSC}$, and n-doping, where $\Gamma = IP_{DOP} - EA_{OSC}$; with IP and EA being the Ionization Potential and the Electron Affinity. Charge transfer integrals couple the frontier molecular orbitals involved in charge hopping processes relevant to p or n-type doping. In this model, the dipole moment operator in the Mulliken approximation $\boldsymbol{\mu} = e\sum_i(\mathbf{r}_i - \mathbf{r}_0)$, where e is the elementary charge, $\mathbf{r}_0$ and $\mathbf{r}_i$ are respectively the positions of the dopant and host sites, allows the calculation of optical properties and the electrical susceptibility.

As discussed in Sec. 4.1, this model has been accurately parameterized for F4TCNQ-doped Pentacene, for which the parameters are presented in Fig. 5.5, and validated against embedded MBPT calculations, where it presented an excellent agreement

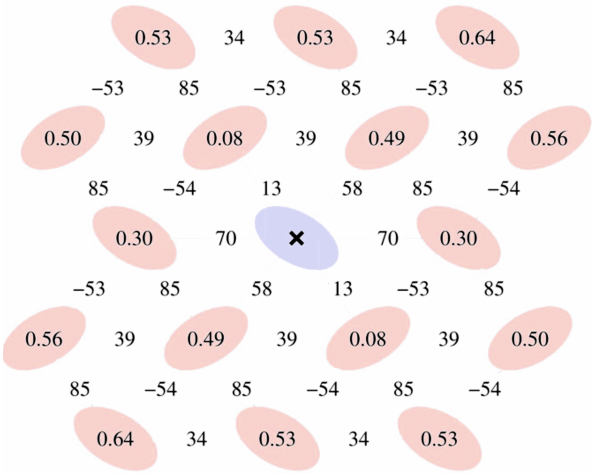

Figure 5.5: Energies of diabatic CT states used to parameterize the Mulliken Hamiltonian of Eq. (5.9). The site energies correspond to a configuration with an electron localized on the central F4TCNQ and holes in the HOMO of each Pentacene site, and the energy of the neutral state is set to zero. Intermolecular couplings are annotated on the respective bonds. Adapted from [37].

with the ground-state and excitonic properties [131]. The model can be inexpensively diagonalized for the set of parameters corresponding to F4TCNQ-doped Pentacene, granting access to the CT polarizability of host-dopant complexes. The latter is obtained from the SOS expression of Eq. (5.6) applied to Hamiltonian (5.9). The result is reported in Table 5.2, i.e. a CT polarizability of 357 Å^3 which compares reasonably well with that arising from embedded DFT/ME calculations.

B. *Application to F6TCNNQ-doped NPD*

The good agreement of this Mulliken model with reference calculations for the excitonic properties and the CT polarizability for the Pentacene-F4TCNQ complex gives us full confidence on the ability of Hamiltonian (5.9) to faithfully describe the low-energy physics of host-dopant complexes. Therefore, we explore as a further example the host-dopant complexes extracted from a single realistic morphology obtained from atomistic Molecular Dynamics (MD) simulations mimicking the vapor deposition process of doped organic layers, which has been simulated and characterized in a previous study [38]. Specifically, we consider an amorphous sample of N,N′-di(1-naphthyl)-N,N′-diphenyl-(1,1′-biphenyl)-4,4′-diamine (NPD), p-doped by the strong electron acceptor 2,2′-(1,3,4,5,7,8-Hexafluoro-2,6-naphthalenediylidene)bis-propanedinitrile (F6TCNNQ) at a concentration of 2% mol., as shown in Fig. 5.6.

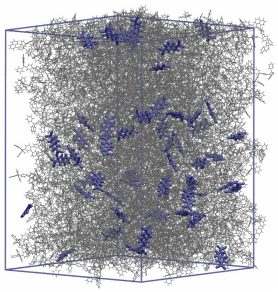

Figure 5.6: Snapshot of the Molecular Dynamics (MD) simulation of an amorphous sample of NPD doped with F6TCNNQ at 2% doping used in this work. Taken from [38].

Hamiltonian (5.9) has been parameterized and diagonalized for all the host-dopant complexes formed by each F6TCNNQ dopant in the sample and its first shell of NPD neighbors, selected as those with a nearest intermolecular atom-atom distance below 5.5 Å. The values of IP_{OSC} and EA_{DOP} were computed for all the molecules in the sample by combining accurate *ab initio* techniques (DFT and G*W*) and classical Micro-Electrostatic calculations, fully accounting for energetic disorder of conformational and environmental nature [131]. Intermolecular CT couplings have been evaluated at DFT level with the dimer projection method [4] [155], with the PBE0 functional and def2-SVP basis set.

The electronic properties gathered over the complexes in the NPD-F6TCNNQ sample are obtained from the diagonalization of Hamiltonian (5.9) and are shown in Fig. 5.7. In particular, the statistics of polarizabilities feature large values exceeding 500 Å^3, confirming the high polarizability associated to host-dopant CT degrees of freedom. These calculations based on the vertical gap, namely calculated with nuclear positions frozen at the geometry of the neutral system, place this system in the fractional CT

4. The dimer projection method allows to obtain the intermolecular hopping matrix elements by taking the interacting dimer Hamiltonian projected on the basis of the molecular orbitals of the isolated fragments.

[155] Valeev et al. (2006).

regime, as indicated by the charge on the dopant Q_{dop} (central panel in Fig. 5.7), with an average polarizability of 577 Å^3 resulting from an extremely broad distribution. Upon considering the adiabatic gap, i.e. taken at the relaxed geometries for molecular ions [5], the system moves towards the full ionization of dopants and smaller gaps, determining a further enhancement of polarizability with an average polarizability of 1267 Å^3. Thus, the CT polarizability of host-dopant complexes is found here as well to be one order of magnitude larger than the typical polarizability of a single host molecule in F6TCNNQ-doped NPD.

5. Here the adiabatic gap is approximated by subtracting the polaron binding energy $\lambda = 0.28$ eV to the vertical gap.

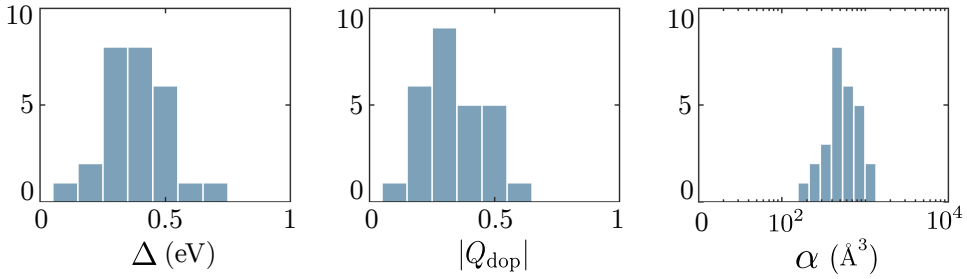

Figure 5.7: Distributions of the optical gap, dopant charge, and polarizability calculated for the host-dopant complexes extracted from a MD-simulated NPD-F6TCNNQ amorphous sample. These results were obtained from the diagonalization of the Mulliken Hamiltonian parameterized with atomistic calculations.

C. *General Insights*

Having determined similar enhancements of the CT polarizability in two different systems, we now go beyond the system-specific picture and explore the Mulliken Hamiltonian of Eq. (5.9) on very general grounds. By spanning a set of microscopic parameters relevant to doped OSCs, this carefully validated model can provide a complete picture of the ground and excited-state properties as well as of the polarizability of host-dopant complexes.

We consider a dopant and its first shell of 12 neighbors in a FCC lattice arrangement at 6 Å distance, in a medium where Coulomb interactions are screened by a dielectric constant of $\varepsilon = 3.5$. All these parameter values are representative of the broad family of molecular OSCs. Then, we discuss the properties of the model as a function of the most important parameters, namely the transfer integral $t_i^{CT} = t_{ij} = t$, the energy $E_{CT} = \Gamma - V$, corresponding to the optical gap at vanishing t, and the standard deviation of the on-site Gaussian energetic disorder σ.

The electronic properties of this model such as the energy gap Δ between the ground and the first excited state, the ground-state charge Q_{dop} on the dopant and the polarizability α are presented in Fig. 5.8. The results are obtained as an average over 1000 disorder realizations.

Two limiting regimes and a crossover region can be identified for the lowest-energy excitation (gap) and the charge on the dopant Q_{dop}. The dopant is largely neutral ($Q_{dop} \approx 0$) for a CT state energy $E_{CT} \gg 0$, where the gap grows linearly with E_{CT}. In the opposite regime, the dopant is fully ionized (integer CT, $|Q_{dop}| \approx 1$) and the gap becomes independent on E_{CT}. Indeed, in this limit the ground state belongs to the quasi-degenerate subspace of CT basis states, with the gap vanishing for disorder-free ($\sigma = 0$) non-overlapping molecules ($t = 0$). The intermediate regime, $E_{CT} \approx 0$, sometimes referred to as "orbital hybridization" or "CT complex" scenario, is characterized by a fractional charge on the dopant.

Fig. 5.8 shows that an order of magnitude amplification of the polarizability occurs when the system presents either a fractional or full CT in the ground-state. This

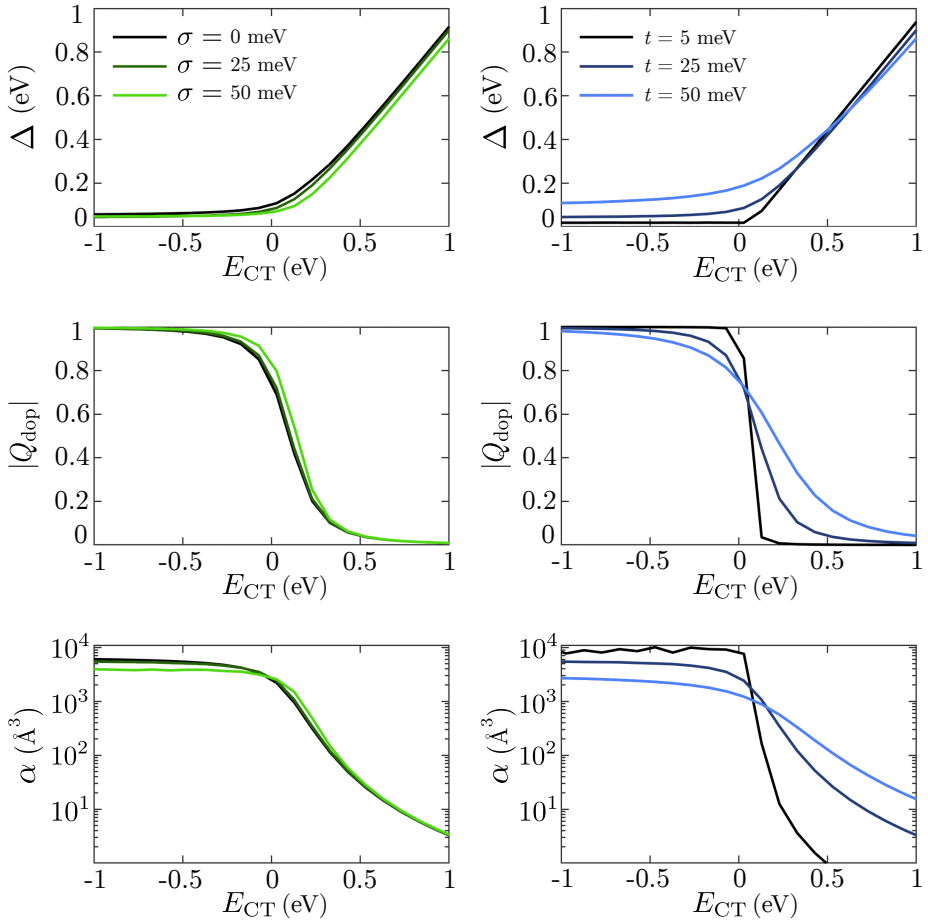

Figure 5.8: Optical gap (top panels), dopant ground-state charge (middle) and polarizability (bottom) calculated with the Mulliken Hamiltonian as a function of the model's parameters, which span the range of interest for doped OSCs.

is the result of low-energy dipole-allowed excitations that greatly contributes to α, according to the sum-over-state Eq. (5.6). The polarizability for integer CT complexes ($q \to 1$, $E_{CT} \ll 0$) is primarily amplified by small host-host CT integrals and by low values of disorder. In the fractional CT regime ($q \lesssim 0.5$, $E_{CT} \gtrsim 0$) large-magnitude host-dopant couplings determine an increase of the polarizability.

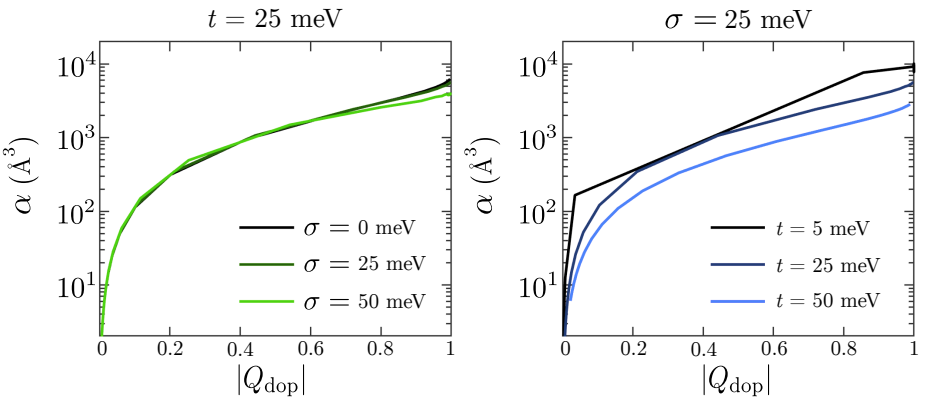

Figure 5.9: Relationship between the polarizability and the charge on the dopant in the Mulliken model, showing that large polarizability enhancements are predicted for host-dopant complexes with both fractional and integer CT in the ground state.

Fig. 5.9 provides an alternative representation of the same data, explicitly showing that the large enhancement in the polarizability of the complex for both a fractional or full CT in the ground state. This also underlines the robustness of this important result against the hopping strength (delocalization) and disorder.

This analysis allows us to conclude that host-dopant complexes featuring both a fractional and an integer CT in the ground state are intrinsically highly polarizable. Indeed, in these two scenarios that have been documented in experiments [137], the excitation spectrum of host-dopant complexes is expected to be characterized by the presence of dipole-allowed CT transitions at very low energy that largely contribute to the polarizability. Recent experiments on Pentacene doped with F4TCNQ and F6TCNNQ showed indeed the occurrence of electronic transitions at 0.6 eV and 0.25 eV, respectively [146]. The predicted polarizability enhancement associated to host-dopant complexes appears to be a general phenomenon that we expect to occur in a large variety of doped organic systems.

[146] Theurer et al. (2021).

5.2 Model for the Dielectric Properties of Doped OSCs

We have demonstrated that host-dopant complexes in organic semiconductors are in general highly polarizable, with a typical polarizability that is one order of magnitude higher than that of a single host molecule. This section will treat the details of the microscopic model that will be used to relate this polarizability increase to the dielectric properties of doped OSCs, accounting for disorder at various doping loads. We will first discuss how to obtain the dielectric properties of a system within the Micro-Electrostatic framework, which will allow us to calculate the spectrum of the collective polarization modes of a doped OSC. We then define the tools that will be used to perform a spectral analysis of such modes, which will be later employed to characterize the doping-induced dielectric screening phenomena in OSCs.

5.2.1 Macroscopic dielectric response

The material's response to an applied electric field determines whether it is a metal or an insulator. The microscopic quantity that characterizes such a response is the polarizability, while the associated macroscopic quantity is the electrical permittivity ε (historically named *dielectric constant*). ε describes the long-range screening of the Coulomb interactions within the solid, and it is finite for an insulator while it is expected to diverge at the insulator-to-metal transition, for a continuous transition in the static limit. It can therefore be used to probe the insulator-to-metal transition from the insulator side. The divergence of the dielectric constant was termed the *dielectric catastrophe* [6] by Mott [45], and was first reported for doped Silicon [46].

6. Also known as the *Drude divergence*.

[45] Mott and Littleton (1938).

[46] Rosenbaum et al. (1983).

A standard relation between the dielectric constant and the molecular polarizability, which is valid for homogeneous and isotropic systems such as gases and liquids, is the Clausius-Mossotti Equation (CME) [7]

$$\boxed{\varepsilon = \frac{1 + \frac{2}{3\epsilon_0}\alpha n}{1 - \frac{1}{3\epsilon_0}\alpha n},} \tag{5.11}$$

where n is the number density. In crystals with cubic space group, the CME holds exactly since the microscopic fields of induced dipoles vanish at the lattice points by symmetry. This relation assumes in particular that the system polarizes uniformly in response to an external field and that the contribution of the permanent dipoles of molecules is negligible, either because the molecules are non-polar or because molecular motion can be neglected at the time scale characterizing the experiment.

7. In this equation, the polarizability is expressed in units of volume, and the equation can be restated in Hartree atomic units noting that $\epsilon_0^{-1} = 4\pi$.

The CME is intrinsically inadequate to describe inhomogeneous media, such as molecular solids with dopant impurities. Several corrections to CME have been proposed for doped inorganic semiconductors, where the contribution of the local fields acting on dopant impurities resulted in an enhancement of the dielectric constant [156, 157]. These approaches were based on local-field corrections in a continuum host, which finds a justification in the low impurity densities typical of doped inorganic semiconductors [8].

[156] Castellan and Seitz (1951).

[157] Dhar and Marshak (1985).

The Micro-Electrostatics (ME) approach, presented in detail in Sec. 3.4, allows to lift most of the approximations inherent to the CME by fully accounting for the

8. We shall compare this enhancement of the dielectric constant upon doping to what is obtained from ME later on.

complex structure of molecular solids that are characterized by anisotropic molecular polarizabilities and low-symmetry lattices [116, 158, 107]. While based on the neglect of the molecular overlap, this approximation does not compromise significantly the accuracy of the ME framework when compared to other methods and experiments [158].

[116] Soos et al. (2001).

[158] D'Avino et al. (2016).

Moreover, the advantage of this approach consists in the possibility to describe with numerically exact calculations inhomogeneous systems of thousands of molecular sites, possibly accounting for periodic boundary conditions: as a result, the present method allows to span the interval of doping loads of relevance for experiments. We note that while the dielectric properties and the related dipole-field sums have been extensively investigated in high-symmetry lattices of equivalent sites [159, 160, 161, 162, 163, 164, 165], to the best of our knowledge, the effect of disorder and inhomogeneity remains completely unexplored to date.

[159] Mueller (1935).

[160] Clark (1958).

[161] Coldwell-Horsfall and Maradudin (1960).

[162] Colpa (1971).

[163] Purvis and Taylor (1982).

[164] Vanzo et al. (2015).

[165] Pines (2018).

In the ME approach, the polarization of the system is described in terms of induced dipoles, obtained by solving the linear equations (see Sec. 3.4)

$$\underline{\mathbf{T}} \cdot \boldsymbol{\mu} = \mathbf{S}, \tag{5.12}$$

where $\underline{\mathbf{T}}$ is the Hessian matrix of the system, accounting for the harmonic energy required to form induced dipoles and mutual dipolar interactions, and $\mathbf{S}$ is the source vector comprising external fields and those arising from the fixed charges present in the system. In particular, we will study in the next section the response of a system to a uniform external field and to a point charge. Solving Eq. (5.12) for the induced dipoles grants access to the macroscopic polarization $\mathbf{P}$ of the sample in response to the source fields, i.e. the dipole per unit volume

$$\mathbf{P} = \frac{1}{V} \sum_i \boldsymbol{\mu}_i, \tag{5.13}$$

where i indexes atomic lattice sites and V is the unit cell volume. Considering a uniform applied field $\mathbf{E}$ as the only source, the electrical susceptibility $\underline{\chi}$ is related to the polarization $\mathbf{P}$ by

$$\mathbf{P} = \varepsilon_0 \underline{\chi} \cdot \mathbf{F} = \varepsilon_0 \underline{\zeta} \cdot \mathbf{E}, \tag{5.14}$$

where $\mathbf{F}$ is the total macroscopic field and $\underline{\zeta}$ is defined as the system's response to the external field. The matrix $\underline{\zeta}$ can readily be computed from Eq. (5.14). For a uniformly polarized dielectric in the shape of an ellipsoid, or a limiting case such as a thin film, $\mathbf{F}$ and $\mathbf{E}$ are constant within the material and related by

$$\mathbf{F} = \mathbf{E} - \frac{\mathbf{P} \cdot \underline{\boldsymbol{\Delta}}}{\varepsilon_0} \tag{5.15}$$

where $\boldsymbol{\Delta}$ is the depolarization tensor which accounts for the macroscopic shape of the system [9] [19]. Then, $\underline{\zeta}$ can be used to compute the electric susceptibility $\underline{\chi}$ as follows

$$\boxed{\underline{\chi} = \left(1 - \boldsymbol{\Delta} \cdot \underline{\zeta}\right)^{-1} \cdot \underline{\zeta},} \tag{5.16}$$

and the relative permittivity is then directly obtained as $\underline{\varepsilon} = \underline{\mathbf{1}} + \underline{\chi}$.

9. The depolarization tensor always verifies $\mathrm{Tr}\,\underline{\boldsymbol{\Delta}} = 1$. For a spherical system,

$$\underline{\boldsymbol{\Delta}} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

For a planar system in the xy plane,

$$\underline{\boldsymbol{\Delta}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

[19] Kittel (1976).

It is instructive to evaluate Eq. (5.16) for a cubic lattice. We consider a system with N sites in a volume V, each carrying the same isotropic polarizability α. The isotropy of the system leads to vanishing off-diagonal elements in $\underline{\varepsilon}$ and $\underline{\zeta}$, leading to the definition of the scalar quantities $\varepsilon = \mathrm{Tr}\,\underline{\varepsilon}/3$ and $\zeta = \mathrm{Tr}\,\underline{\zeta}/3$, while the scalar

depolarization factor is always $\Delta = \mathrm{Tr}\,\underline{\mathbf{\Delta}}/3 = 1/3$. Assuming a neutral system, the total field is $\mathbf{F} = \mathbf{E} + \mathbf{F}^{\mu}$, where $\mathbf{F}^{\mu}$ is the microscopic field due to the induced dipoles. However, in the cubic lattice the dipole fields at each lattice point vanish by symmetry, thus $\mathbf{F}^{\mu} = 0$. The polarization is then obtained straightforwardly from Eq. (5.14) as

$$\mathbf{P} = \frac{1}{V}\sum_i \boldsymbol{\mu}_i = \frac{1}{V}\sum_i \alpha_i \mathbf{F}_i = \frac{N}{V}\alpha\mathbf{E} = \varepsilon_0 \zeta \mathbf{E}, \tag{5.17}$$

leading to $\zeta = \alpha n/\varepsilon_0$. As expected, by inserting this result into Eq. (5.16) the Clausius-Mossotti relation is exactly recovered

$$\varepsilon = 1 + \left(1 - \zeta/3\right)^{-1}\zeta = \frac{1 + \frac{2}{3}\zeta}{1 - \frac{1}{3}\zeta}. \tag{5.18}$$

Thus, we see that in the case of cubic systems the response to an external field ζ coincides with the polarizability density $\alpha n/\varepsilon_0$. In non-cubic systems, disordered lattices or extended molecules, ζ must be calculated numerically from a self-consistent calculation of mutual dipolar interactions.

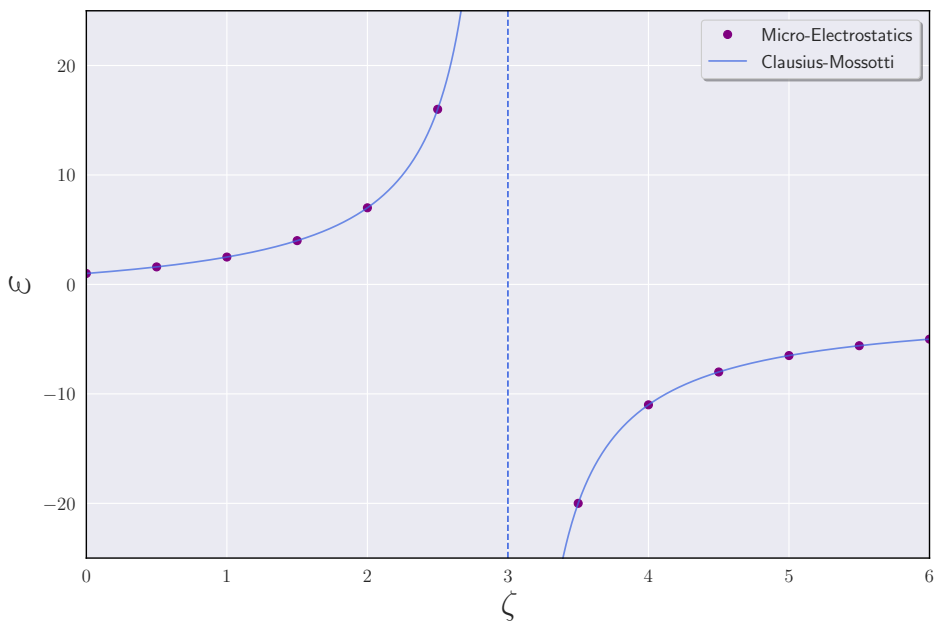

Figure 5.10: Static relative permittivity ε calculated analytically from the CME and numerically from ME as a function of $\zeta = 4\pi\alpha n$. The system considered in the ME calculations consist of a FCC lattice with 4000 sites and a lattice constant $a = 5\,\text{Å}$, accounting for interactions with periodic replica up to a radius of $R = 700\,\text{Å}$.

The above relation also provides a mean of validation for numerical ME calculations of the dielectric constant. Fig. 5.10 presents the relative permittivity of a spherical system with a FCC crystal structure as a function of αn, as calculated analytically from the CME and numerically from ME. The unit cell was replicated assuming Periodic Boundary Conditions (PBC) up to a large radius in order to converge the electrostatic dipole sums. As anticipated, the CME holds exactly, and the ME calculations faithfully reproduce the analytical result for the homogeneous cubic system, giving us confidence on the applicability of this approach to inhomogeneous and disordered systems.

Fig. 5.10 shows the superlinear increase of the dielectric constant with ζ, translating the growing response of the system upon increasing its polarizability density. The dielectric constant eventually diverges at the critical value $\zeta = 3$, corresponding to the point where the depolarization field in Eq. (5.16) exactly compensates the external field, i.e. the system perfectly screens the applied field, signaling the transition into a

metallic state. This phenomenon, called the *dielectric catastrophe*, marks the reach to a metallic state from the insulating side, and has been first discussed by Herzfeld [44] as a criterion for the occurrence of an insulator-to-metal transition. The negative values of ε after the divergence determine the establishment of a spontaneous polarization corresponding to unstable polarization modes associated with negative eigenvalues of the Hessian, as we will see in detail in the next section.

5.2.2 Spectral analysis of the response

The analysis of the linear system of Eq. (5.12) governing the polarization of the material will be particularly helpful in the study of the enhancement of the dielectric constant and its possible divergence upon doping. Indeed, the spectrum of the Hessian matrix provides information on the susceptibility of the collective polarization normal modes. This is a valuable information, especially when one is interested in systems presenting amplified responses due to the approach of an instability, signaled by vanishing (soft modes) or negative (unstable modes) eigenvalues. Therefore, we now relate $\underline{\zeta}$ to the spectrum of the Hessian matrix $\underline{\mathbf{T}}$ (spectral decomposition).

We begin by noting that Eq. (5.14) is a vectorial equation for the unknown matrix $\underline{\zeta}$ which must be evaluated three times in order to obtain all the components of the response. This is achieved by applying three non-collinear external fields to the same neutral system. Without loss of generality, we assume the three external fields to be along the Cartesian axes with magnitude E, i.e. $E^{\alpha\gamma} = E\delta_{\alpha\gamma}$. Here, the index γ denotes the distinct linear systems associated to each applied field, namely $\underline{\mathbf{T}}\boldsymbol{\mu}^{\gamma} = \mathbf{E}^{\gamma}$, or equivalently

$$\sum_{j\beta} T_{ij}^{\alpha\beta}\mu_j^{\beta\gamma} = \delta_{\alpha\gamma}. \tag{5.19}$$

In this notation, $\mu_i^{\beta\gamma}$ is the β component of induced dipole at site i in reaction to the field applied along γ [10]. The induced dipoles then reduce to

10. Following the notation introduced in Section 3.4, Greek indices label Cartesian axes and Roman indices label the lattice sites.

$$\mu_i^{\alpha\beta} = E\sum_j \left[T^{-1}\right]_{ij}^{\alpha\beta}. \tag{5.20}$$

On the one hand, we now have the expression of the polarization associated to each applied field

$$P^{\alpha\beta} = \frac{1}{V}\sum_i \mu_i^{\alpha\beta} = \frac{E}{V}\sum_{ij}\left[T^{-1}\right]_{ij}^{\alpha\beta}. \tag{5.21}$$

On the other hand, recasting Eq. (5.14) into a single matrix equation leads to

$$P^{\alpha\beta} = \varepsilon_0\sum_\gamma \zeta^{\alpha\gamma}E^{\gamma\beta} = \varepsilon_0 E\zeta^{\alpha\beta}. \tag{5.22}$$

Combining the above equations results in an expression of $\underline{\zeta}$ solely in terms of the Hessian matrix

$$\boxed{\zeta^{\alpha\beta} = \frac{1}{\varepsilon_0 V}\sum_{ij}\left[T^{-1}\right]_{ij}^{\alpha\beta}}. \tag{5.23}$$

In accordance with linear response theory, this expression highlights that $\underline{\zeta}$ is independent of the external field, reflecting the fact that it is an intrinsic property of the system. Eq. (5.23) is used to compute $\underline{\zeta}$ when the Hessian is directly inverted, thus avoiding the need of performing three separate calculations for calculating the dielectric response of the material.

The spectral decomposition of the response $\underline{\zeta}$ can now be related to the Hessian's spectrum. Indeed, since $\underline{\mathbf{T}}$ is symmetric, it admits a unique factorization

$$\underline{\mathbf{T}} = \underline{\mathbf{O}} \cdot \underline{\mathbf{\Lambda}} \cdot \underline{\mathbf{O}}^{\mathsf{T}}, \tag{5.24}$$

where $\underline{\mathbf{O}}$ is the orthogonal matrix whose columns are the eigenvectors and $\underline{\mathbf{\Lambda}}$ is the diagonal matrix of eigenvalues. In terms of global indices [11], this yields

$$T_{ij} = \sum_k O_{ik}\, \lambda_k\, O_{jk} \tag{5.25}$$

where now $i, j, k \in [1, \ldots, 3N]$ span the whole basis composed 3 Cartesian components for each atomic site. Then, Eq. (5.23) is recast as

$$\zeta^{\alpha\beta} = \frac{1}{\varepsilon_0 V} \sum_{ijk} \delta_i^\alpha\, O_{ik}\, \lambda_k^{-1}\, O_{jk}\, \delta_j^\beta, \tag{5.26}$$

where δ_i^α is the projection matrix which selects the global indices corresponding to the α Cartesian component. The above expression can be used to resolve $\zeta^{\alpha\beta}$ in terms of its spectral contributions, i.e. that of each eigenmode, by writing

$$\zeta^{\alpha\beta} = \frac{1}{\varepsilon_0 V} \sum_k \zeta_k^{\alpha\beta}. \tag{5.27}$$

The spectral weights $\zeta_k^{\alpha\beta}$, corresponding to the contribution of eigenmode k to $\zeta^{\alpha\beta}$, are thus defined as

$$\zeta_k^{\alpha\beta} = \lambda_k^{-1} \sum_{ij} \delta_i^\alpha\, O_{ik}\, O_{jk}\, \delta_j^\beta = \lambda_k^{-1}\, O_k^\alpha\, O_k^\beta, \tag{5.28}$$

where we defined the projection of eigenmode k on the Cartesian axis α as

$$O_k^\alpha = \sum_i \delta_i^\alpha O_{ik}. \tag{5.29}$$

In other words, the spectral weight $\zeta_k^{\alpha\beta}$ in Eq. (5.28) is the product of the inverse of the k^{th} eigenvalue with the projections of the k^{th} eigenvector along axes α and β.

In order to gain physical insights on the spectral weights of $\underline{\zeta}$, we define the effective mode polarizability α_k (in Å^3) and the effective mode density n_k (in Å^{-3}) as

$$\alpha_k = \frac{1}{4\pi\varepsilon_0 \lambda_k} \tag{5.30a}$$

$$n_k^{\alpha\beta} = \frac{1}{V} O_k^\alpha O_k^\beta. \tag{5.30b}$$

The effective polarizability α_k is inversely proportional to the Hessian's eigenvalues, while the effective mode density $n_k^{\alpha\beta}$ is proportional to the product of the eigenvector's projections along Cartesian axes α and β defined in Eq. (5.29). We can define an isotropic effective mode density as

$$n_k = \operatorname{Tr} \frac{n_k^{\alpha\beta}}{3} = \frac{1}{3V}\left(\left(O_k^x\right)^2 + \left(O_k^y\right)^2 + \left(O_k^z\right)^2 \right). \tag{5.31}$$

Thus, n_k is proportional to the square of the algebraic sum of the eigenvector's weights along each Cartesian axis. As such, it is an intricate measure of both the delocalization and the net polarization of the mode, see Fig. 5.11.

With the help of such definitions, we arrive at an insightful and compact expression for the spectral decomposition of $\zeta = \operatorname{Tr} \underline{\zeta}/3$ [12]

11. Under this new notation, the Roman indices span the whole basis in which the matrices are written, which comprises 3 basis elements per atomic site. As such, Eq. (5.23) is re-expressed as

$$\zeta^{\alpha\beta} = \frac{1}{\varepsilon_0 V} \sum_{ij} \delta_i^\alpha \left[\mathbf{T}^{-1}\right]_{ij} \delta_j^\beta,$$

where $\delta_i^\alpha = 1$ if global index i corresponds Cartesian direction α, and $\delta_i^\alpha = 0$ otherwise.

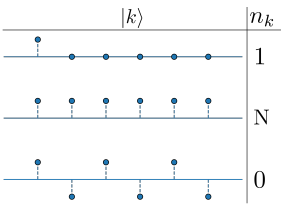

Figure 5.11: Effective mode density in three limiting cases. A fully localized eigenvector (top) has $n_k = 1$ and a fully delocalized eigenvector with a net polarization (middle) has $n_k = N$ whereas a fully delocalized eigenvector with no net polarization (bottom) has instead $n_k = 0$.

$$\boxed{\zeta = 4\pi \sum_k \alpha_k \, n_k.} \tag{5.32}$$

In the thermodynamic limit, obtained by taking $N \to \infty$, $V \to \infty$ while keeping the density N/V constant, Eq. (5.32) can be rewritten as an integral

$$\boxed{\zeta = \int d\lambda \, \zeta(\lambda) = 4\pi \int d\lambda \, f(\lambda)\alpha(\lambda)n(\lambda),} \tag{5.33}$$

where $f(\lambda)$ is the spectral density of polarization states. $\alpha(\lambda)$ and $n(\lambda)$ are defined as

$$\alpha(\lambda) = \frac{1}{4\pi\varepsilon_0\lambda} \tag{5.34a}$$

$$n(\lambda) = \sum_k n_k \delta(\lambda - \lambda_k). \tag{5.34b}$$

Eq. (5.32), or equivalently Eq. (5.33), express the response to the external field ζ to the case of inhomogeneous systems as a sum over its spectral weights. Each mode contributes with the product of its effective polarizability α_k and its effective mode density n_k in a way that is analogous to the polarizability density that controls the CME in Eq. (5.11). Large effective polarizabilities are expected to arise from soft modes since $\alpha_k \sim 1/\lambda_k$, but only those determining a net polarization $n_k \neq 0$ will effectively contribute to the response.

We conclude this section by mentioning that all the above features, namely the full diagonalization or inversion of the Hessian and the subsequent spectral decomposition of $\underline{\zeta}$, have been implemented in *the original software Tequila that was developed by myself during this Thesis* [119]. The code is written in C++, and is parallelized for shared memory architectures with OpenMP.

[119] Comin (2022).

5.3 Doping-Induced Dielectric Catastrophe

Our investigation on the doping-induced dielectric catastrophe now moves forward by relating the local polarizability enhancement of charge-transfer complexes to the dielectric properties of the doped material with the microscopic model described in the previous section. Our model is set up with typical parameters for doped Pentacene, obtained from Sec. 5.1, that are broadly representative of the entire class of molecular OSCs.

We begin our discussion with the analysis of the polarization normal modes, which will provide important insights on the collective dielectric response to an external field. In particular, we will evaluate the density of polarization modes, revealing the central role played by the the collective polarization modes associated with dopant complexes. We then assess the enhancement of the bulk dielectric constant of doped OSCs as a function of the doping load, revealing the advent of a doping-induced dielectric catastrophe at dopant concentrations of 5-10%, equivalent to what needed experimentally to trigger an increase of the conductivity by orders of magnitude. Finally, we will address the charge dissociation free energy profiles, namely the potential barrier that carriers must overcome to escape from their parent ionized dopant, showing that the inhomogeneous electrostatic profile further favors charge release.

5.3.1 Polarization Normal Modes

Following the methodology developed in the previous section, we now evaluate the dielectric properties of doped OSCs, specifically addressing the case of F4TCNQ-doped Pentacene. To such an aim, we set up a Micro-Electrostatic model on a $10 \times 10 \times 10$ Face-Centered Cubic (FCC) lattice, whose sites represent either a host molecule or a host-dopant complex. The lattice constant is set to $a = 12.26$ Å, which corresponds to a dielectric constant of $\varepsilon = 3.5$ for the pristine host, according to the Clausius-Mossotti Equation of Eq. (5.11).

The polarizability of host sites is set to $\alpha_{host} = 50$ Å^3, while dopant sites are assigned a polarizability of $\alpha_{CT} = 500$ Å^3, consistent with estimates obtained in Sec. 5.1 for F4TCNQ-doped Pentacene. The latter are randomly distributed across the lattice, with the only constraint of avoiding impurity clustering by disallowing dopants to be placed as first or second neighbors of each other [13] The doping load ρ is explicitly controlled by the amount of impurity sites in the system. Since dopants actually bear the polarizability of a host-dopant CT complex, the interaction between its polarizability and that of the first host neighbors has been neglected. This corresponds to set a cutoff between host-dopant interactions $n_{cut} = 1$, where n_{cut} is the number of nearest-neighbor shells up to which the host-dopant interactions are ignored. We will discuss the effect of this parameter in the following section.

13. This constrained random uniform distribution effectively limits the accessible doping values to z %, where $z = 12$ is the coordination number of the FCC lattice.

Moreover, our calculations account for periodic boundary conditions. Specifically, the limit of an infinite bulk system is achieved by periodically replicating the simulation cell up to a spherical cutoff of 700 Å, ensuring the convergence of dipole field sums. Finally, the disorder has been sampled by considering 1000 realizations of dopant positions.

A. *The Density of Polarization States*

As a first insight on the effect of the increasing dopant density ρ, we compute the spectrum of the Hessian matrix of the extended system [14] by diagonalizing it exactly. The Hessian eigenvalues λ quantify the stiffness of the collective polarization modes of the system of interacting polarizabilities, i.e. their inverse, λ^{-1}, is proportional to the polarizability contribution associated to each normal mode. In particular, vanishing positive Hessian eigenvalues determine a finite polarization even in the presence of a tiny perturbation, while modes characterized by $\lambda < 0$ are unstable, namely they develop a spontaneous polarization even in the absence of an electric field.

14. The Hessian was defined previously in Eq. (3.151) of Sec. 3.4 and Eq. (5.12) of Sec. 5.2.

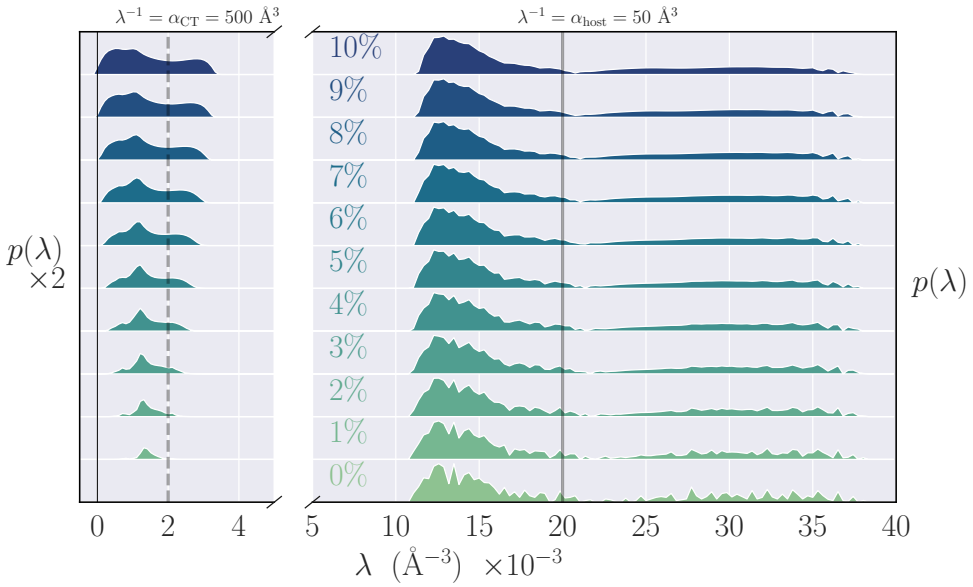

Figure 5.12: Density of polarization states λ, measuring the stiffness of the collective polarization modes in a doped OSC, as a function of the impurity concentration ρ. The vertical lines mark the inverse polarizabilities of non-interacting host molecules and host-dopant complexes. The y-scale axis in the left panel has been magnified for better visualization.

The Hessian's Density of States (DoS), namely the distribution of λ values at different doping loads is shown in Fig. 5.12, where the vertical lines mark α_{host}^{-1} and α_{CT}^{-1}, i.e. the stiffness of non-interacting host and dopant sites, respectively. A sizable eigenvalue dispersion due to inter-site interactions is obtained already in the pristine system, ρ=0%. This distribution is bimodal, with a peak at $13 \cdot 10^{-3}$ Å^{-3}, and a much broader feature centered around $30 \cdot 10^{-3}$ Å^{-3}. These spectral regions correspond respectively to the transverse and longitudinal polarization mode branches, as will be made clear in the Appendix of this Chapter (Sec. 5.5).

Upon increasing doping, the shape of the distribution barely changes in the region corresponding to the host sites ($\lambda > 10 \cdot 10^{-3}$ Å^{-3}). On the other hand, new peaks resulting from highly polarizable host-dopant complexes appear around α_{CT}^{-1}, indicated by the dashed vertical line in Fig. 5.12.

These peaks grow in intensity and significantly broaden with increasing concentration as a result of the interaction between complexes. Most interestingly, the low-λ tail of the distribution approaches zero upon increasing ρ, with 1.6% of the modes associated to host-dopant complexes becoming unstable ($\lambda < 0$) at 8% doping. The softening of the collective polarization modes with doping signals an enhancement of the susceptibility as the system moves toward a *dielectric catastrophe*.

B. *Spectral Decomposition of the Response*

We are now in the position to evaluate the contribution of the different polarization normal modes to the response of the system, and in particular those of the host and dopant polarization bands. We derived in Eq. (5.32) and Eq. (5.33) the decomposition of the response to the external field ζ as a sum, or integral, of spectral weights $\zeta(\lambda)$ originating from each polarization mode. In particular, each spectral weight is proportional to the product of the effective polarizability $\alpha(\lambda) = 1/(4\pi\varepsilon_0\lambda)$ and effective density $n(\lambda)$ of each mode, in a way that is reminiscent of the site polarizability and the number density entering the Clausius-Mossotti equation in Eq. (5.11). By analyzing the eigendecomposition of the Hessian at various doping loads, we present in Fig. 5.13 the spectral distribution of $\alpha(\lambda)$, $n(\lambda)$ and $\zeta(\lambda)$ as a function of the Hessian's eigenvalues λ.

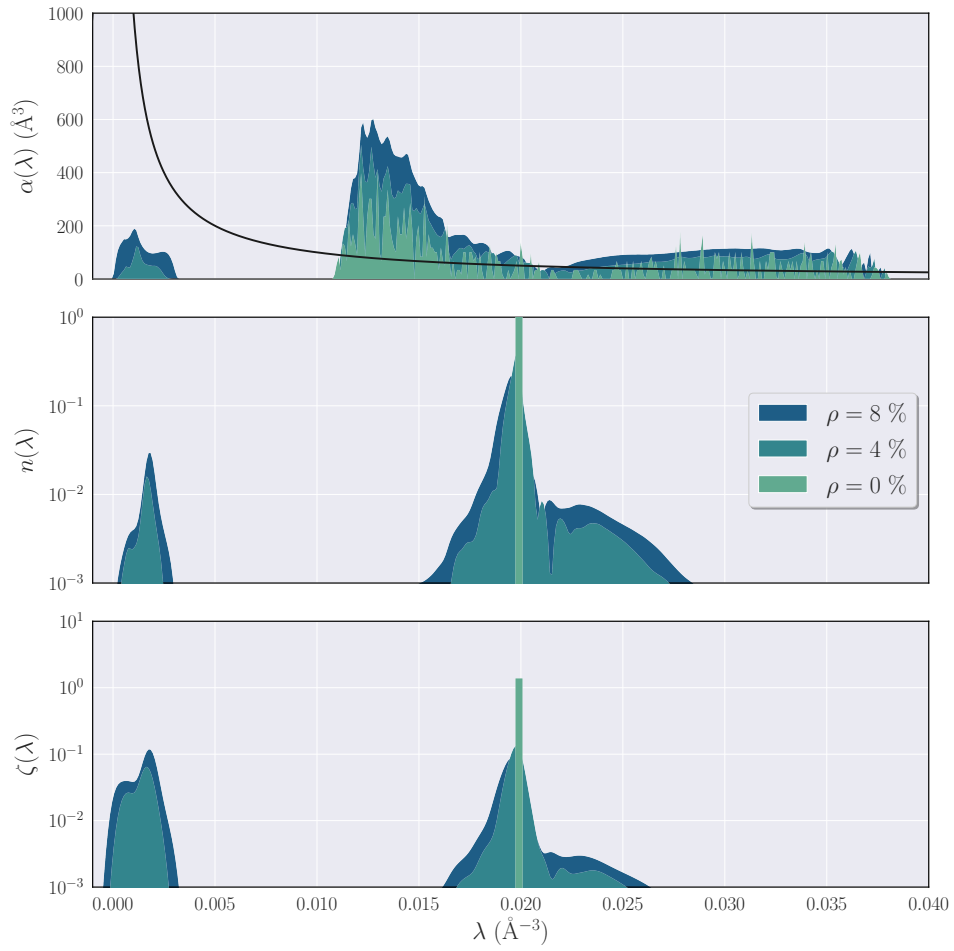

Figure 5.13: Spectral decomposition of the response function ζ in terms of the Hessian's eigenvalues for doping loads $\rho = 0, 4, 8\%$. (Top panel) The density of states $p(\lambda)$ is shown at an arbitrary scale along with the effective polarizability of each state $\alpha(\lambda)$ (black curve). (Middle panel) The distribution of the effective densities $n(\lambda)$ is shown on a semi-logarithmic scale and is multiplied by the volume V such that $n(\lambda) = 1$ corresponds to a fully delocalized mode. (Lower panel) The distribution of the spectral weights $\zeta(\lambda)$ is shown on a semi-logarithmic scale, and is defined such that $\int d\lambda\, \zeta(\lambda) = \zeta$.

As mentioned above, the effective mode polarizability is the inverse of the corresponding Hessian eigenvalue, as shown in Fig. 5.13 (top panel). This implies that polarization modes with lower eigenvalues are associated with larger effective polarizabilities. At finite doping loads, we witness the emergence of soft polarization modes in the spectral region associated with dopants ($\lambda \leqslant 5 \cdot 10^{-3}$ Å^{-3}). As the doping load increases, this band broadens and polarization modes approaching, and eventually crossing, $\lambda \to 0$ appear. This results in quasi-divergent effective mode polarizabilities, signaling a possibly strong contribution of the dopant band to the response.

The system's response is however weighted by the density of such polarization states (top panel of Fig. 5.13) and the effective density of each mode $n(\lambda)$ (middle panel of Fig. 5.13), which is related to the net polarization associated with each mode.

In the pristine system ($\rho = 0\%$), the effective mode densities $n(\lambda)$ are distributed as δ-like peaks for the three degenerate fully delocalized modes with $n(\lambda) = 1$ located at $\lambda_0 = \alpha_{host}^{-1}$. Indeed, in the ordered system the only modes featuring a net polarization are the three modes corresponding to equal induced dipoles at all lattice sites along each of the 3 Cartesian directions. Because these modes are in-phase combinations of the dipoles in each unit cell, we can already state that they are located at the Γ point in reciprocal space. Consequently, the response of the pristine system is determined solely by the polarization modes at Γ, whose effective polarizability coincides precisely with the site polarizability of the system, namely $\alpha_{host} = 50$ Å^3.

As a result, the spectral distribution of $\zeta(\lambda)$ is also composed of a single peak at λ_0, and the response of the pristine system is simply given by

$$\zeta = \sum_{\lambda \in \Gamma} \zeta_\lambda = 4\pi\alpha_{host} \sum_{\lambda \in \Gamma} n_\lambda = 4\pi\alpha_{host} n. \tag{5.35}$$

As expected from the CME for the homogeneous system, the response is entirely governed by the host polarizability α_{host} and the number density n.

At finite doping loads, the presence of randomly distributed and highly polarizable species breaks the translational symmetry of the lattice. The effective mode densities $n(\lambda)$ undergo a sizable broadening around the position of the δ peak of the pristine system, with the development of a new peak in the dopant spectral region. As a result many modes contribute to the response of doped samples, whose spectral distributions are now formed of two broad peaks located in the host and dopant spectral regions.

The resulting spectral weights of the response $\zeta(\lambda)$, shown in the bottom panel of Fig. 5.13, account for the combined effect of the density of states, the effective polarizability and effective density. We observe that while in the pristine system the response is totally dictated by the sole contribution of the three modes at $\lambda = \alpha_{host}^{-1}$ Å^{-3}, the response at finite doping loads stems from both the host and dopant spectral regions in a similar way.

An alternative way to quantify the net contribution of each spectral region to the total response ζ comes from the cumulative distribution function of $\zeta(\lambda)$, shown in Fig. 5.14 for increasing doping loads. We see a rising contribution to ζ from the dopant spectral region as the doping increases, eventually exceeding that of the host at $\rho = 8\%$. Conversely, the part of the host spectral region goes from a single δ-like peak to a broadened distribution that results in an overall contribution to the response that is practically constant with increasing doping. The impact of such an enhanced response with doping on the bulk dielectric constant will be discussed in detail in the next section.

C. *Analysis of the Polarization Normal Modes*

A question that might arise from this analysis is whether the bands so far associated with host and dopant sites are actually only composed by induced dipoles of the respective chemical species, or if a substantial mode mixing occurs because of host-

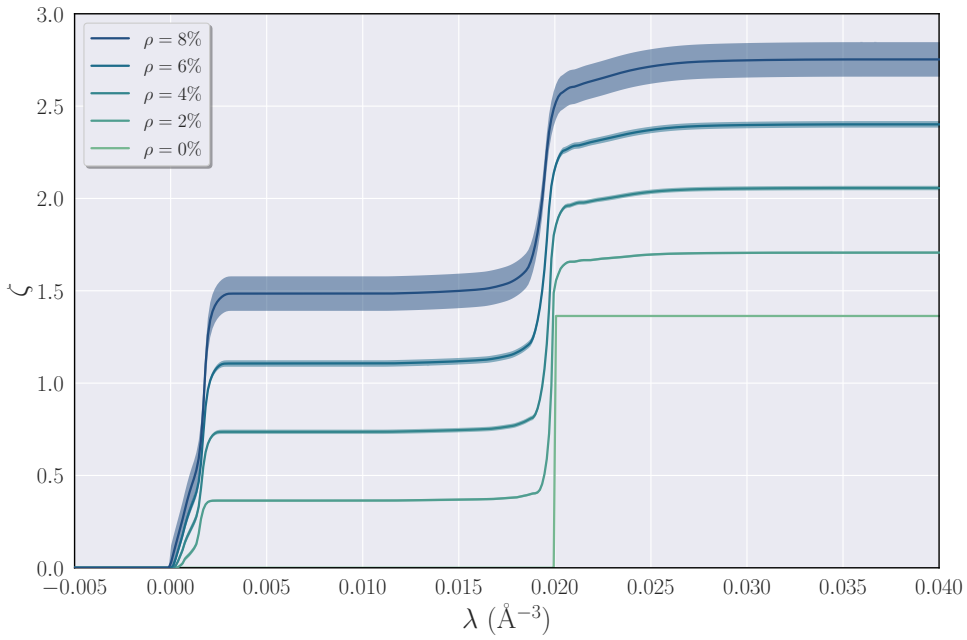

Figure 5.14: Cumulative distribution function of the spectral response function $\zeta(\lambda)$ for doping loads $\rho = 0, 2, 4, 6, 8$ %. The shaded areas show the standard deviation around the distribution function with respect to the dopant position realizations.

dopant interactions. To address this question, we have computed the contribution of host and dopant sites to each normal mode. Specifically, the host and dopant weights $W_h(\lambda), W_d(\lambda)$ of each mode λ are obtained as

$$W_h(\lambda) = \sum_{i \in \text{host}} |O_{i\lambda}|^2 \tag{5.36a}$$

$$W_d(\lambda) = \sum_{i \in \text{dop}} |O_{i\lambda}|^2, \tag{5.36b}$$

where $O_{i\lambda}$ is the i^{th} component of the polarization mode with eigenvalue λ. Because of the eigenvector's normalization, the sum of the host and dopant weights of each mode yields $W_h(\lambda) + W_d(\lambda) = 1$, and the weights can be interpreted as the portion of hosts (dopants) involved in each mode.

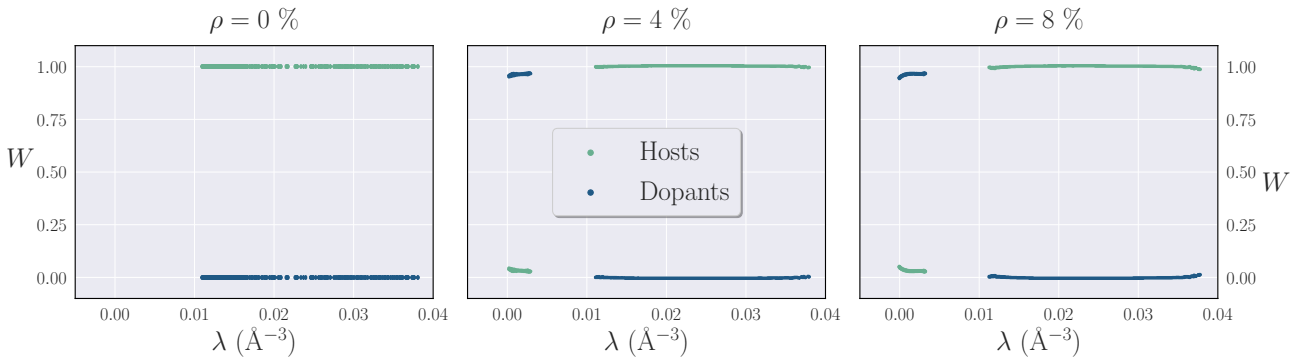

Figure 5.15: Spectral distribution of the weights of each polarization mode onto the host and dopant sites, for doping loads $\rho = 0\%$ (left), 4% (middle) and 8% (right). The spectral regions associated to host's and dopant's modes involve only the respective sites, with a small mixing between the modes.

The spectral distribution of such weights at doping loads $\rho = 4, 8\%$ is shown in Fig. 5.15 along with the pristine system for which all modes obviously involve only host sites. In doped systems we observe a small, but not negligible, hybridization between host and dopants modes that is particularly pronounced near $\lambda = 0$. Dipolar interactions between host and dopant sites are hence not able to substantially mix the

contribution of the two species, characterized by very different polarizabilities. This allows us to conclude that the two separated bands of collective polarization modes are characterized by a well-defined host and dopant character.

To go further with the present analysis, we now discuss the degree of delocalization of the host and dopant polarization modes, in order to get insights into the individual or collective response of highly-polarizable dopant sites and quantify whether the disorder introduced with doping is able to cause a localization of the host modes. The degree of delocalization is measured by the mode's Participation Ratio (PR) $P_r(\lambda)$, which is defined as

$$P_r(\lambda) = \frac{1}{N}\left[\sum_{i=1}^{N} p_i(\lambda)^2\right]^{-1} \tag{5.37a}$$

$$p_i(\lambda) = \sum_{\alpha}^{x,y,z} |O_{i\lambda}^{\alpha}|^2 \tag{5.37b}$$

where $p_i(\lambda)$ is the site population, $O_{i\lambda}^{\alpha}$ is the α Cartesian component at site i of the polarization mode with eigenvalue λ, and N is the number of sites in our simulations. The PR quantifies the number of sites contributing to a given eigenmode, hence measuring the degree of delocalization. A fully localized eigenmode has $P_r = 1/N$ while a fully delocalized mode having $P_r = 1$. This allows to interpret the PR as the fraction sites over which each mode is delocalized.

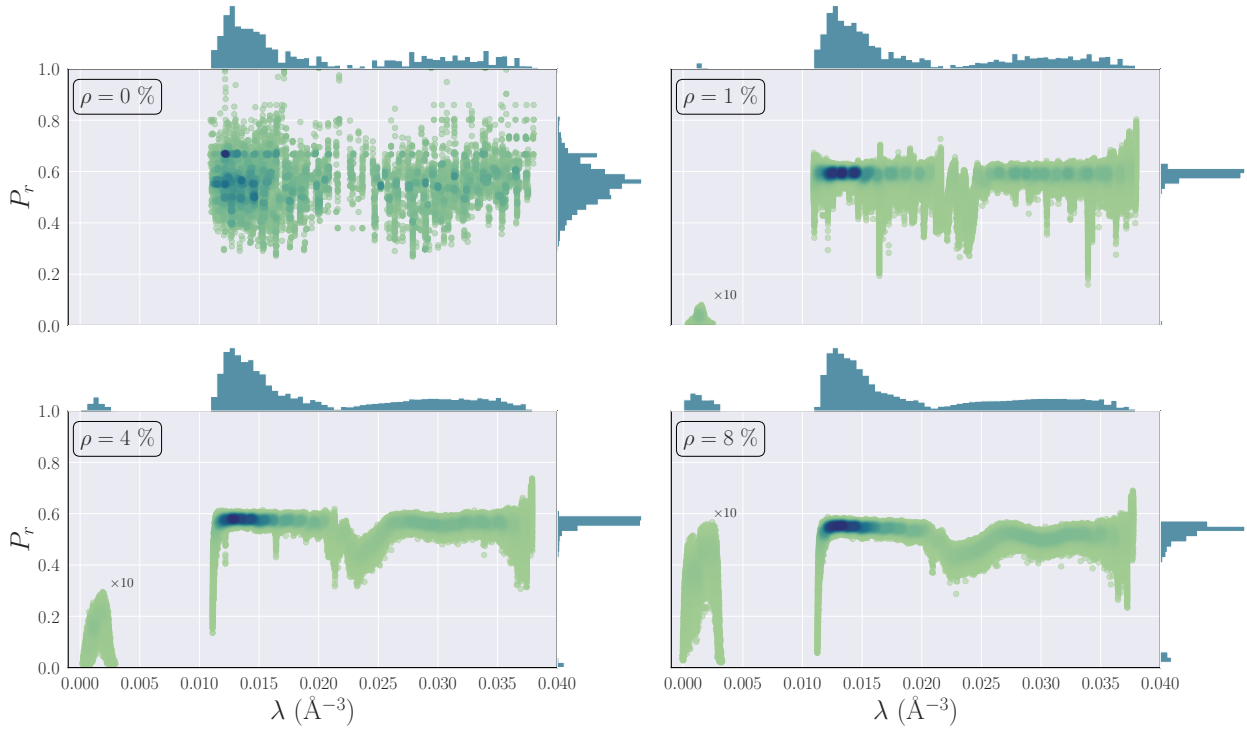

Figure 5.16: Spectral distribution of the Participation Ratio of each polarization mode $P_r(\lambda)$ for doping loads $\rho = 0\%$ (top left), 1% (top right), 4% (lower left) and 8% (lower right). Individual distributions of the eigenvalues (DoS) and of the Participation Ratio are shown on the corresponding axes. The values of the PR in the dopant spectral region are scaled by a factor 10 in the main plot, and the color map indicates the point density.

The spectral distribution of the Participation Ratio is presented as a function of λ in Fig. 5.16 for doping loads $\rho = 0, 1, 4, 8\%$. In the pristine system, the Participation Ratios are distributed similar to a Gaussian, except for a significant peak at 66.6%. This

implies an important spread in the PRs ranging from 25% to 100% of the system's size, which is approximately constant along the spectrum. The fully delocalized modes at $\lambda_0 = \alpha_{\text{host}}^{-1}$ correspond to the modes at Γ mentioned above.

At finite doping loads, we notice the emergence of a peak in the Participation Ratio associated with dopant polarization modes at $\lambda \leqslant 5 \cdot 10^{-3}$ Å^{-3}. In particular, this peak is formed by a localized edges and a delocalized center, a feature that is characteristic of the eigenvalue spectrum of disordered systems. Indeed, the states at band extrema are characterized by a smaller dispersion, hence being those that are more prone to undergo a disorder-induced localization. Moreover, the delocalization of the dopant polarization modes appears to increase with doping, an effect that can be directly attributed to the increasing number of dopant sites over which the collective normal modes can delocalize.

The chemical disorder induced by the dopant impurities strongly affects the localization of the host polarization modes, leading to a narrower distribution of the PR that is peaked at the center of the pristine system's distribution. As for the dopant band, a sizable localization at the host band edges is seen. We further notice a dip in the distribution of the PR at the central region of the host band (between 0.020 and 0.025 Å^{-3}) corresponding to the minimum of the DoS.

In summary, our study of the polarization normal modes composition allows us to state that i) normal modes pertaining to the host and dopant spectral region involve only the respective sites and ii) such modes are effectively delocalized over dopant and host sites away from the dopant and host band edges.

5.3.2 Bulk dielectric properties of doped OSCs

A. Enhancement of the dielectric constant upon doping

We now turn our attention to the bulk dielectric constant as a function of the doping load. We begin our analysis by reporting in Fig. 5.17 the distribution of the values of ζ among the dopant position realizations as a function of the doping load ρ, as obtained from the inverse Hessian in Eq. (5.26). Since the average response of the system over the dopants' positional disorder realizations is isotropic, in the following we will only report the scalar value of the tensors Tr $\underline{\zeta}/3$ and Tr $\underline{\varepsilon}/3$.

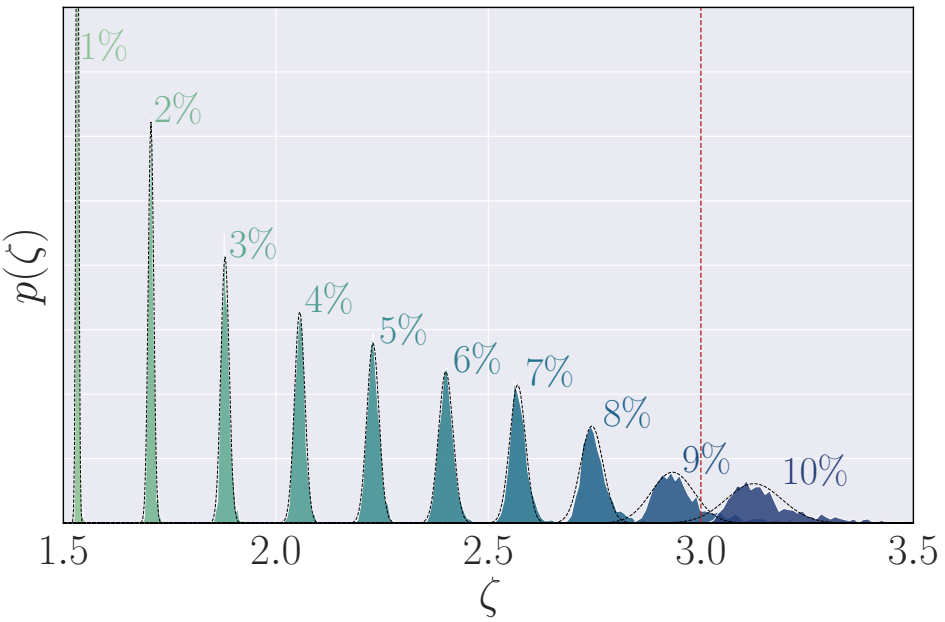

Figure 5.17: Distributions of the values of ζ for doping loads from 0% to 10% for 1000 dopant disorder realizations. The black dashed curves are Gaussian fits of the distributions, and the red dashed line at $\zeta = 3$ corresponds to the divergence the dielectric constant.

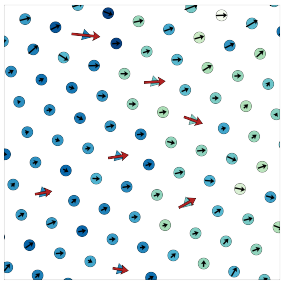

Figure 5.18: Illustration of the induced dipoles in the (111) plane of a doped OSC at $\rho = 8\%$ doping in an homogeneous external field. Host (dopant) molecules are shown as circles (triangles), and the color map indicates the screened Coulomb potential at each site. The black (red) arrows depict the inducted dipoles at host (dopant) sites, not in scale.

The interaction-driven softening of the polarization modes that is seen in Fig. 5.12 results in an enhanced susceptibility to perturbations, and this translates into an average increase of $\zeta(\rho)$ as a function of the doping load in Fig. 5.17. These distributions are symmetric and nicely approximated by Gaussians (black dotted lines) up to $\rho = 8\%$, and develop some skewness at larger doping, when the dielectric becomes more and more unstable. According to Eq. (5.11), the dielectric constant diverges at $\zeta = 3$, and this threshold is passed between 9% and 10% doping load, translating the advent of the dielectric catastrophe.

Indeed, by employing Eq. (5.16), we retrieve the dielectric constant ε as a function of the impurity concentration, as reported in Fig. 5.19. Our calculations predict a divergent behavior of the dielectric constant, leading to a remarkable non-linear amplification of ε upon approaching the singularity from $\zeta < 3$. This amounts to a four (ten) fold increase of ε at 6% (8%) doping with respect to the pristine OSC, which can have a dramatic effect on the release of Coulombically bound charges, as will be discussed later. We emphasize that such a striking increase of ε occurs at the typical concentrations leading to a boost in the measured electrical conductivity [11, 13, 30].

The trespassing of the divergence results in negative values of ε, as in the CME, which signals the advent of the dielectric catastrophe upon doping. Such negative values follow mathematically from the negative effective polarizabilities of unstable modes, reflecting the fact that the system is not anymore at a stable equilibrium point.

[11] Walzer et al. (2007).

[13] Salzmann et al. (2016).

[30] Jacobs and Moulé (2017).

Negative susceptibilities corresponds to unphysical solutions in the present linear response model, which allows to compute only the response of a stable system. It tackles the insulator-to-metal transition from the insulator side by assuming bound charges, and thus it cannot describe conducting states appearing after the dielectric catastrophe.

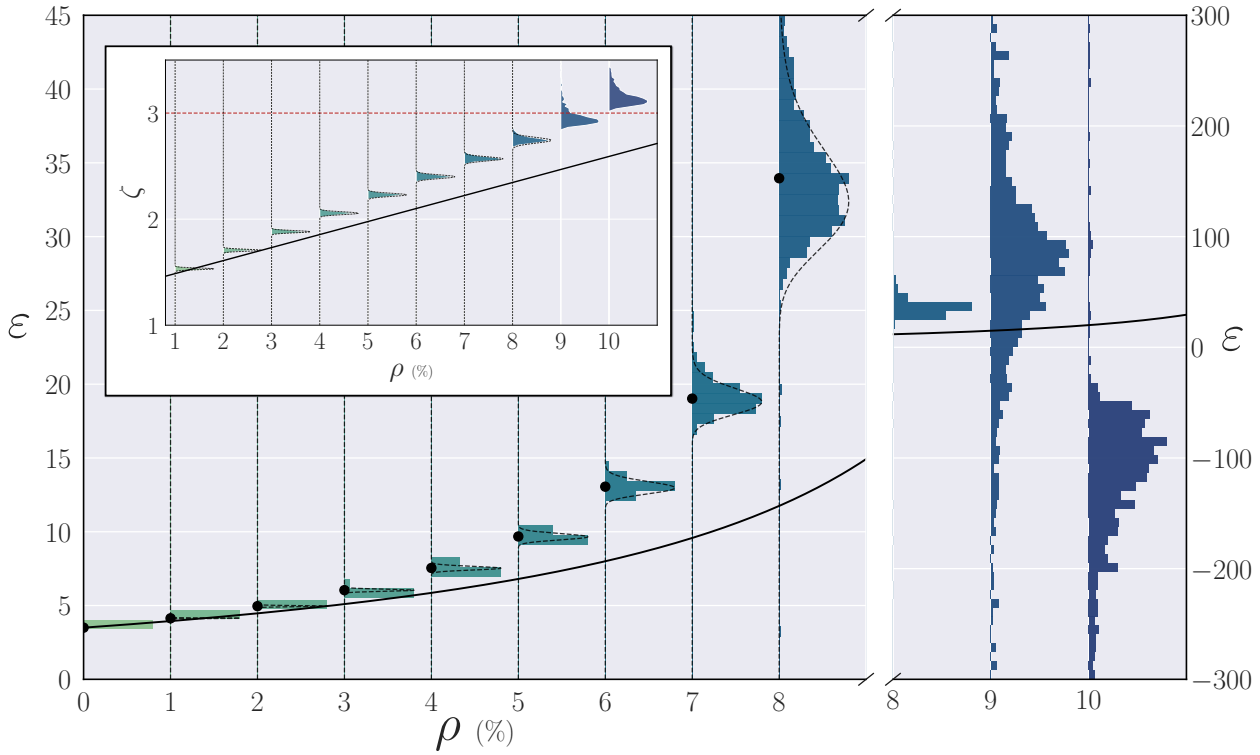

Figure 5.19: Enhancement of the bulk dielectric constant ε of a doped organic semiconductor (F4TCNQ-doped Pentacene) with the impurity concentration ρ. The main panel shows the distribution (histograms) and median (dots) of ε obtained upon sampling over dopants positional disorder. The full line corresponds to the dielectric constant obtained from the CME with an effective site polarizability $\alpha_{\text{eff}}(\rho)$. The inset shows the dependence of the external-field susceptibility ζ on doping, displaying numerical results for the inhomogeneous lattice (distributions and their Gaussian fits as dashed lines) and $\zeta_{\text{eff}}(\rho)$ for an effective homogeneous medium (full line). The Gaussian fits are then analytically transformed into distributions for ε, shown as dashed lines in the main panel. The right-hand panel illustrates the system behavior at the dielectric catastrophe. Note the different scale on the ordinate.

The distribution of ε shifts to higher values, broadens and becomes more and more skewed upon approaching the dielectric catastrophe by increasing ρ, which is reminiscent of the corresponding changes seen on the distribution of ζ. The Gaussian fits of the distributions of ζ allow to draw analytical curves for $\varepsilon(\zeta(\rho))$ (dashed lines in the main panel of Fig. 5.19). For a given value of ρ, we describe ζ as a Gaussian distributed random variable, i.e. $p(\zeta) = f_{\mu,\sigma}(\zeta)$. Then, the distribution of $\varepsilon = \frac{1+a\zeta}{1-b\zeta}$, where $a = \frac{2}{3}$ and $b = \frac{1}{3}$, is given by [15]

$$p(\varepsilon) = \frac{a+b}{(a+b\varepsilon)^2} f_{\mu,\sigma}\left(\frac{\varepsilon-1}{a+b\varepsilon}\right), \tag{5.38}$$

where $f_{\mu,\sigma}(x)$ is the probability density function of the Gaussian distribution of mean μ and standard deviation σ. We observe that the analytical distributions of $\varepsilon(\rho)$, shown as dashed lines in Fig. 5.19, capture to a good approximation the true distributions for $\rho \leqslant 8\%$. For larger doping loads, the distributions of $\zeta(\rho)$ are not Gaussian anymore, which forbids the definition of the analytical distributions of $\varepsilon(\rho)$. Table 5.4 reports the statistical moments of the distributions of $\varepsilon(\rho)$, highlighting the increase of the average (μ_ε), standard deviation (σ_ε) and skewness (γ_ε) of the distributions with the impurity concentration.

15. Indeed, the quotient distribution of ε is given by

$$\begin{aligned} p(\varepsilon) &= \frac{d}{d\varepsilon} P(E \leqslant \varepsilon) \\ &= \frac{d}{d\varepsilon} P\left(\frac{1+a\zeta}{1-b\zeta} \leqslant \varepsilon\right) \\ &= \frac{d}{d\varepsilon} P\left(\zeta \leqslant \frac{\varepsilon-1}{a+b\varepsilon}\right) \\ &= \frac{d}{d\varepsilon} F_{\mu,\sigma}\left(\frac{\varepsilon-1}{a+b\varepsilon}\right) \\ &= \frac{a+b}{(a+b\varepsilon)^2} f_{\mu,\sigma}\left(\frac{\varepsilon-1}{a+b\varepsilon}\right), \end{aligned}$$

where $f_{\mu,\sigma}$ and $F_{\mu,\sigma}$ denote the Gaussian PDF and CDF, respectively. We assumed here $1 - bY > 0$; the opposite case $1 - bY < 0$ proceeds similarly and results in $-p(\varepsilon)$.

The picture that emerges from our results is that two phenomena cooperate in determining the boost of the dielectric constant with ρ. The first one is the straightforward effect of increasing the concentration of highly polarizable host-dopant complexes. This effect can be captured by considering an effective medium with average site polarizability $\alpha_{\text{eff}}(\rho) = (1-\rho)\alpha_{\text{host}} + \rho\alpha_{\text{CT}}$. The corresponding value ζ_{eff} is shown as a solid line in the inset of Fig. 5.19. We observe that both the average of $\zeta(\rho)$ and $\zeta_{\text{eff}}(\rho)$ feature a linear trend, with the cooperative dipolar interactions taking place in the disordered lattice resulting in a faster increase and a broadening of the distributions.

Table 5.4: Statistical moments of the distributions of $\varepsilon(\rho)$ as a function of the impurity concentration ρ. μ_ε is the sample mean, σ_ε the standard deviation, and γ_ε the skewness, which quantifies the deviation to a normal distribution (i.e. $\gamma_\varepsilon = 0$ for a normal distribution). $p(\varepsilon < 0)$ is the fraction of negative ε values. Note the difference between the slow increase of μ_ε, σ_ε and γ_ε before $\rho = 7\%$ and the corresponding non-linear increase beyond that value, signaling a change of trend that is reminiscent of phase transitions.

ρ (%)	μ_ε	σ_ε	γ_ε	$p(\varepsilon < 0)$ (%)
1	4.14	0.01	0.79	0.00
2	4.96	0.04	0.50	0.00
3	6.05	0.07	0.63	0.00
4	7.55	0.12	0.30	0.00
5	9.70	0.22	0.45	0.00
6	13.09	0.47	0.71	0.00
7	19.26	1.35	3.11	0.00
8	29.01	255.67	-30.27	1.60
9	281.41	$4.75.10^3$	24.68	20.20
10	-607.85	$1.77.10^4$	-30.91	59.70

The evaluation of the relative permittivity of this effective medium by the Clausius-Mossotti Equation, shown as a solid line in Figure 5.19, only accounts for 26% of the doping-induced enhancement of $\varepsilon(\rho = 8\%)$. This marks a crucial qualitative difference between a homogeneous dielectric, which obeys the Clausius-Mossotti equation, and an intrinsically chemically inhomogeneous disordered medium, such as a doped OSC. The remaining, dominant, contribution originates from microscopic dipole fields that are strongly polarizing on average at highly-polarizable dopant sites, thus enhancing the external field (see Fig. 5.8).

Specifically, this enhancement and eventual divergence of the dielectric constant upon doping arises from the formation of a dopant band, that results in highly polarizable modes. Such modes determine a depolarization field that can exceed in magnitude the applied field itself, hence completely counteracting it ($\zeta = 3$ in Eq. (5.16)). In the case of a homogeneous medium described by the CME, the dielectric catastrophe is reached when the smaller Hessian eigenvalue vanishes. This is not anymore the case in the inhomogeneous and disordered medium that we modeled, where the criterion for the advent of the dielectric catastrophe ($\varepsilon \to \infty$) involves both the effective mode polarizability and the effective mode density, as was discussed previously.

B. *Influence of the interaction cutoff*

We now turn to the effects of an important parameter in our modeling, namely the interaction cutoff between dopant and host sites n_{cut}. In our model, the CT polarizability of host-dopant complexes was attributed to the impurity sites while the molecular polarizability of the was assigned to the host sites. Since the charge-transfer

polarizability is associated with host-dopant complexes, and not with individual dopants, the interaction of dopant sites with neighboring host sites has been neglected, setting the interaction cutoff to $n_{cut} = 1$. While this is the most natural choice, also consistent with *ab initio* calculations in Sec. 5.1 considering the dopant and its first shell of Pentacene neighbors, this choice remain somewhat arbitrary, hence motivating the following analysis.

We hence explore the effect of increasing the interaction cutoff to the second and third neighbor shells by reporting the average of the dielectric constant $\langle \varepsilon(\rho) \rangle$ with respect to the doping load for different values of n_{cut}. The data in Fig. 5.20 show that upon increasing values of n_{cut}, the average dielectric constant from explicit numerical calculations approaches the curve for the effective homogeneous medium, given by the CME.

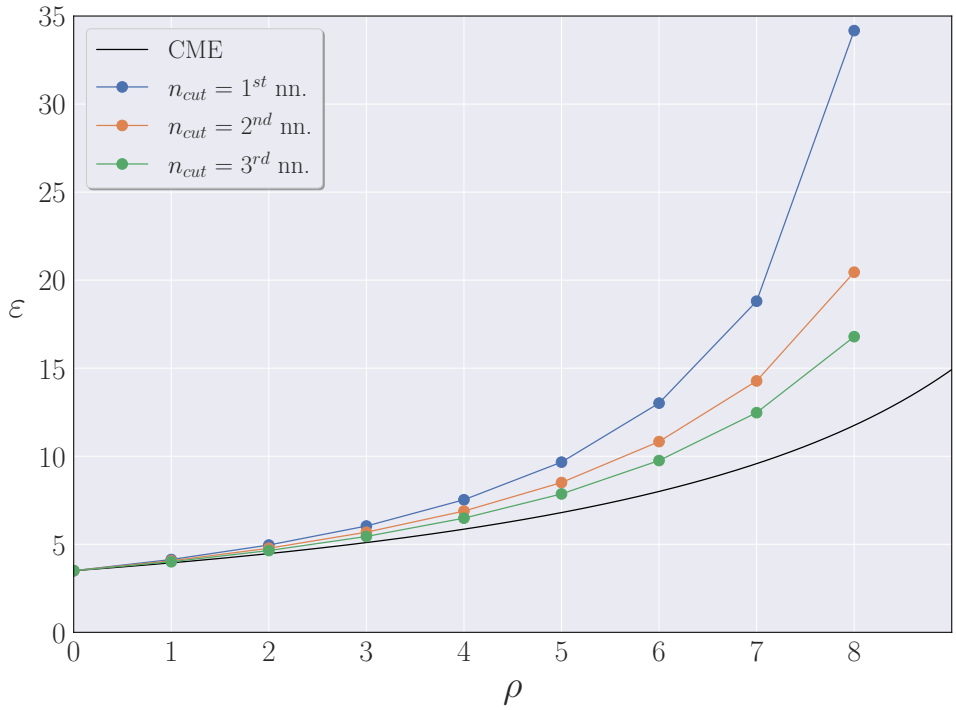

Figure 5.20: Median bulk dielectric constant ε as a function of the impurity concentration ρ for different values of the interaction cutoff n_{cut}. The data is obtained from an average over 500 realizations of dopant disorder, with the same parameters as in Fig. 5.19. The curves are compared to the dielectric constant of the effective homogeneous medium, as given by the CME.

To go further in our analysis, we present in Fig. 5.21 the distributions of the microscopic dipolar fields F_μ at lattice sites for bulk samples polarized by an external field $\mathbf{F}^{ext} = 0.1$ V/Å, for different doping loads ρ and interaction cutoffs n_{cut}. In the case of a uniformly polarized isotropic system, such as the pristine lattice, these fields cancel out exactly by symmetry in the cubic system, which is precisely the reason why the Clausius-Mossotti relation holds. This is qualitatively different from an inhomogeneous system where this symmetry is valid only in a statistical sense, i.e. upon averaging over different sites and different random distributions of dopants over the lattice. And indeed, the average of the dipolar fields $\langle F_\mu \rangle$ is zero, but the average performed over dopant sites $\langle F_\mu \rangle_{dop}$ is instead significantly larger (+37%) than the external field, and this value is notably independent of the doping load ρ. This, together with the larger polarizability of dopant sites, determines the stronger response in our calculations for inhomogeneous systems with respect to an effective homogeneous medium.

We observe in Fig. 5.21 that the shift of $\langle F_\mu \rangle_{dop}$ is reduced when the host-dopant in-

teractions are decreased. Therefore, we ascribe the increase of $\langle F_\mu \rangle_{dop}$, and ultimately the enhancement of the dielectric constant with respect to the homogeneous effective medium, to the interactions between dopants and their neighboring hosts. Indeed, in the absence of host-dopant interactions ($n_{cut} \to \infty$) the two subsystems would respond independently and their polarization would simply sum up. This would lead, on average, to a response that would be equal to that of an homogeneous medium with average effective polarizability α_{eff}, i.e. $\langle \varepsilon(\rho) \rangle \to \varepsilon_{eff}(\rho)$ and $\langle F_\mu \rangle_{dop} \to 0$. Thus the host-dopant interactions are precisely what differentiates our ME model from the CME applied to α_{eff}.

In conclusion, the interaction between the polarizabilities of dopant complexes and neighboring host molecules is seen to be crucial in the description of the dielectric catastrophe in doped OSCs. It is this interaction that, along with the effect of disorder, implies a significant boost of the relative permittivity of the material beyond that predicted by the corresponding effective homogeneous medium.

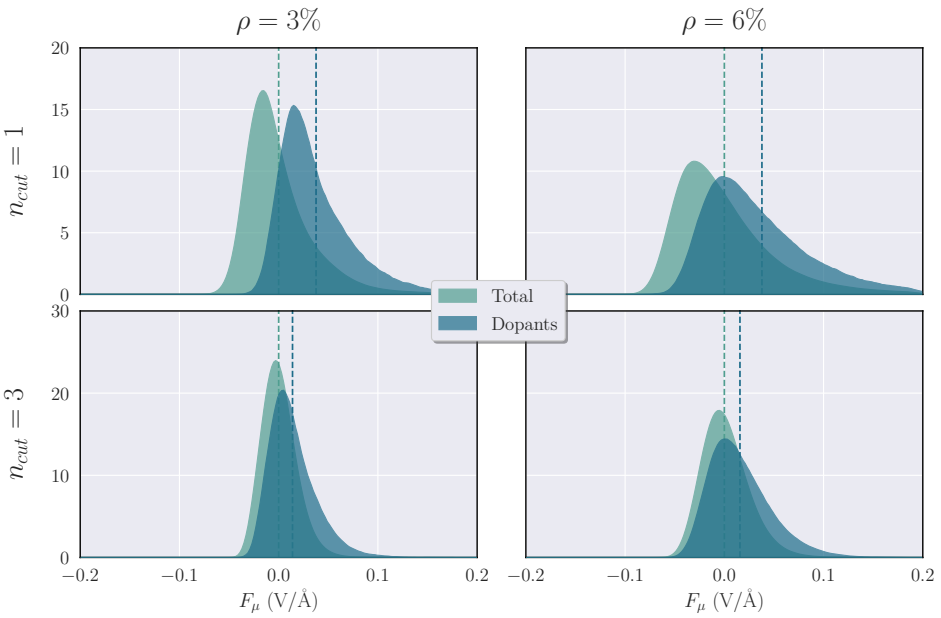

Figure 5.21: Microscopic fields at lattice sites, $\mathbf{F}_i^\mu$, for bulk samples polarized by a an external field $\mathbf{F}^{ext} = 0.1$ V/Å. The figure reports the component of the field parallel to $\mathbf{F}^{ext}$ for systems at 3% (left column) and 6% (right column) doping, and for an interaction cutoff of $n_{cut} = 1$ (top row) and $n_{cut} = 3$ (bottom row). The distribution of microscopic fields are broad and asymmetric as a result of the dipolar interactions in the disordered lattice. The field components perpendicular to $\mathbf{F}^{ext}$ average to zero.

C. *Influence of positional disorder*

The host-dopant interactions and chemical disorder are two major effects in the boost of the dielectric constant in the doped OSC. The results presented above have been obtained for a FCC lattice, which represents an idealization for the structure of doped organic semiconductors. To go further in our analysis and make one more step towards a less idealized description of disordered doped materials, we now investigate the effect of relaxing the assumption of an ordered lattice.

Specifically, we consider a disordered FCC lattice, in which the Cartesian components of the site positions of the ordinary FCC are supplemented with a random Gaussian displacement with zero mean and standard deviation σ. In the following we discuss results obtained for $\sigma = 0.45$ Å, which was estimated from the width of the first-neighbor peak of the radial distribution function of the NPD-F6TCNNQ amorphous sample in Fig. 5.6. We stress that such a site-position disorder should not be confused with the random distribution of dopants over lattice sites; the two sources of disorder are both at play in these calculations.

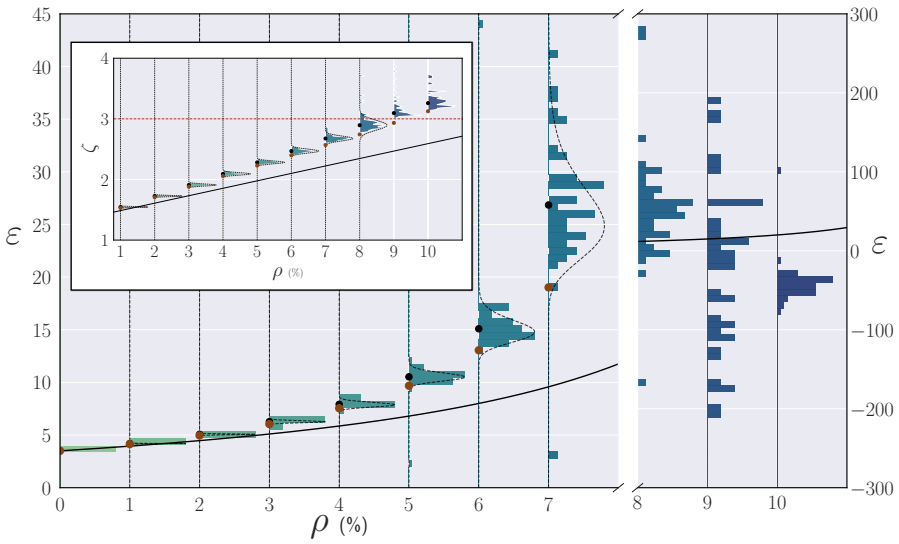

Figure 5.22: Distribution (histograms) and medians (black dots) of the bulk dielectric constant of F4TCNQ-doped pentacene with respect to doping load (vertical histograms) for the disordered lattice. The brown dots show the median of the ordered lattice calculations of Fig. 5.19 in the manuscript. Analytical fit from the distributions of ζ (dashed lines). Inset: distribution of ζ wrt. doping load and least-squares fit (dashed lines). Results obtained for 50 disorder realizations, medians are shown as a robust estimator of the mean at low sampling.

Our results, shown in Fig. 5.22, reveal a qualitative agreement between the doping-induced enhancement of the dielectric constant in the ordered and in the disordered lattice. The dielectric catastrophe, signaled by the occurrence of negative ε values, occurs however at lower doping density in the disordered lattice. This suggests that the effect of the site-position disorder and that of dopants distribution actually sum up in amplifying the dielectric response, as can be seen by the non-negligible enhancement in the average values of ε with respect those of the ordered lattice. We stress that the two sources of disorder considered here do not only result in a broadening of the distributions of ε, but more importantly in an increased average of such distributions.

5.3.3 Dissociation energy profiles

We have shown that the bulk dielectric constant of doped OSCs undergoes a remarkable increase, and eventually diverges, upon increasing the impurity concentration. However, the microscopic charge dissociation energetics in a chemically inhomogeneous and disordered OSC may differ greatly from those due to the macroscopic, bulk dielectric screening. In a continuous medium of bulk dielectric constant $\langle\varepsilon\rangle$, a charge sitting at a distance r from its parent dopant feels a screened Coulomb potential

$$W(\mathbf{r}) = -\frac{1}{\langle\varepsilon\rangle\, r}. \tag{5.39}$$

Conversely, in a disordered and inhomogeneous medium such a charge is subject to a microscopic screened Coulomb potential (which arises from Eq. (3.98) in the static limit $\omega \to 0$)

$$W(\mathbf{r},\mathbf{r}') = \int d\mathbf{r}''\, \mathcal{E}^{-1}(\mathbf{r},\mathbf{r}'')V(\mathbf{r}'',\mathbf{r}'), \tag{5.40}$$

where V is the bare Coulomb potential and $\varepsilon(\mathbf{r},\mathbf{r}'')$ is the non-local microscopic dielectric function. Judging by the highly dispersed distributions of the bulk dielectric constant at the doping loads of interest, we expect these two scenarios to be quite different.

Therefore, in order to complete our description of the charge release in doped OSCs, we discuss in this section the energetics of charge dissociation in the presence of the enhanced dielectric screening sourced from highly-polarizable host-dopant complexes. To such an aim, we calculate the collective response to a point charge on a dopant site, which grants access to the dissociation energy profiles as a function of the electron-hole distance r.

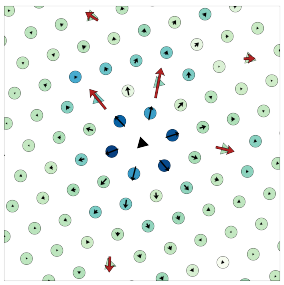

Figure 5.23: Illustration of the induced dipoles in the (111) plane of a doped OSC at $\rho = 8\%$ doping in the field of a central dopant charge. Host (dopant) molecules are shown as circles (triangles), and the color map indicates the screened Coulomb potential at each site. The black (red) arrows depict the inducted dipoles at host (dopant) sites, not in scale.

We accomplish this by calculating the screened Coulomb potential felt at lattice sites from open-boundary Micro-Electrostatic calculations on spherical samples centered around a positively charged dopant site at $\mathbf{r} = 0$, as depicted in Fig. 5.23. In the absence of screening, each site is subject to the bare Coulomb potential

$$V_i = V(\mathbf{r}_i) = \frac{1}{r_i}, \tag{5.41}$$

where r_i is the distance of site i to the source ion at the origin. The screened Coulomb potential felt at each lattice site resulting from the mutual interactions in the system is given by

$$W_i = W(\mathbf{r}_i) = V_i^0 + \sum_{j\neq i} \frac{\boldsymbol{\mu}_j \cdot \mathbf{r}_{ij}}{r_{ij}^3}, \tag{5.42}$$

where $\boldsymbol{\mu}_j$ is the induced dipole at site j, obtained from a self-consistent ME calculation, and $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ the separation between sites i and j. For each random distribution of dopant sites over the lattice, we compute the screened Coulomb potential for three spherical systems with radii $R = 60, 80, 100$ Å in order to attain the infinite system limit. Indeed, since the Coulomb potential must decay asymptotically as R^{-1}, we extrapolate each site's potential within the smallest system ($R < 60$ Å) with respect to R^{-1} to the limit $R = \infty$. Then, the screened Coulomb potential $W(r)$ as a function of the electron-hole radius r is obtained by averaging the extrapolated potentials on lattice sites over spherical shells of increasing radii (up to $R = 60$ Å), namely

$$W(r) = \frac{1}{N(r)} \sum_{r<r_i\leqslant r+\delta r} \widetilde{W}_i, \tag{5.43}$$

where $\widetilde{W}_i$ is the extrapolated potential at site i, and $N(r)$ is the number of sites between r and $r + \delta r$. In practice, we considered the spherical shells that lie exactly between two FCC neighbor shells [16]. We note that the distributions that are shown below are all characterized by the median and median absolute deviation [17].

Fig. 5.24 reports the screened Coulomb potential $W(r)$ as a function of r for different doping loads. The dots and error bars show the mean and standard deviations for the different coordination shells sampled over dopant disorder, while the full lines show the screened Coulomb potential calculated with the bulk value of ε. In the infinite dilution limit $\rho = 0\%$, except for small oscillations at short distance, the energy profile has the form of a Coulomb potential, screened by the bulk $\varepsilon = 3.5$, similar to analogous atomistic results [37]. We attribute such oscillations to the effects of the lattice. The energy barrier to free charges sitting on nearest-neighbor molecules amounts to 0.5 eV, consistent with first-principles calculations [37, 16] and conductivity data at low doping [36].

The situation radically changes at finite dopants concentration. The screening provided by host-dopant complexes determines a sizable suppression of the average charge-separation energy barrier, reducing to 0.3 and 0.15 eV at 4% and 8% doping, respectively. This large screening effect provides a decisive contribution to the charge release mechanism at zero temperature.

16. The sequence of neighbor shells distances of the FCC lattice with nearest neighbor distance a is given by

$$a\sqrt{n}$$

where n is an integer represented by the integer quadratic form $n = x^2 + y^2 + 2z^2$, $x, y, z \in \mathbb{N}$. This sequence can be found on the Online Encyclopedia of Integer Sequences as sequence A000401.

17. The median absolute deviation around the median is defined as the median of absolute deviations from the data's median, namely

$$M(|X_i - M(X)|)$$

where $M(X)$ is the median operator.

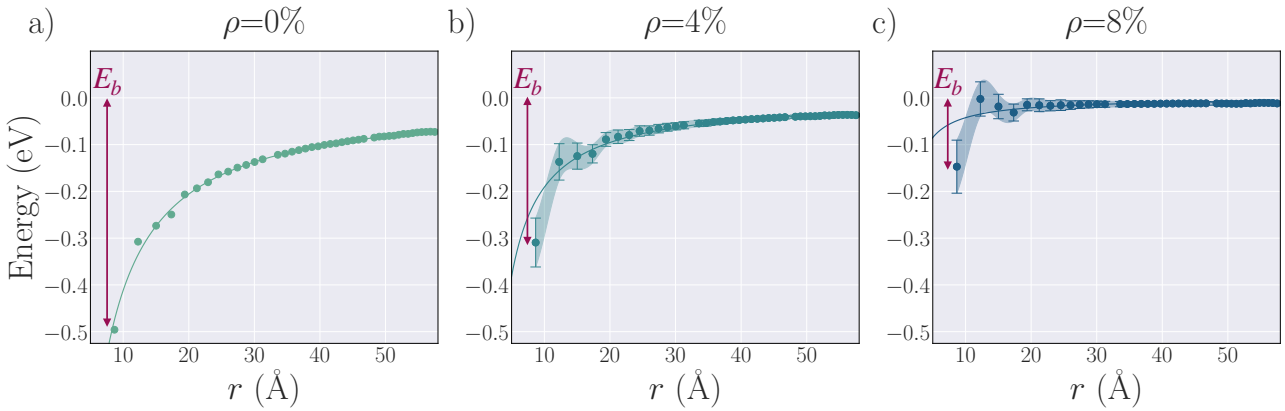

Figure 5.24: Energy profiles for charge separation from a source charge placed on a dopant site at $r = 0$, as a function of the electron-hole distance r for increasing doping loads. Dots and shaded areas are the medians and median absolute deviations of the spherical average of the site energies, while the full lines show the energy profile provided by the bulk dielectric constant ε.

However, the process of charge separation at finite temperature is governed by the free energy of releasing the charge, which accounts for the entropy variation upon carrier release. Therefore, we assess the free energy profiles, defined as $F(r) = E(r) - TS(r)$, where $E(r)$ are the energy profiles that we computed above and $S(r)$ is the entropic contribution. The latter is calculated in the continuum limit as the entropy arising from the radial density of microstates $\Omega(r)$ that carrier can access in the process of being released for conduction. This approach was proposed by Gregg in the context of organic photovoltaics [166]. Assuming a homogeneous medium and a Coulomb potential for the distance dependence of microstate energies, we obtain $\Omega(r) = 4\pi r^2$. The latter approximation concerns only the calculation of the entropy contribution to the free-energy, while the energy term is obtained from Micro-Electrostatic calculations on the lattice with randomly dispersed dopants. This results

[166] Gregg (2011).

in an entropy variation

$$\Delta S(r) = k_B \log \frac{\Omega(r)}{\Omega(r_0)} = k_B \log \left[(1-\rho) \left(\frac{r}{r_0} \right)^2 \right] \tag{5.44}$$

where k_B is the Boltzmann constant, $\Omega(r)$ is the number of microstates with e-h distance r, and r_0 is the nearest-neighbor distance. The term $(1-\rho)$ accounts for the reduced number of sites accessible to the charge carrier at finite doping.

Fig. 5.25 depicts the room-temperature ($T = 300$ K) free energy profiles for charge separation. We observe that the entropy variation always assists charge separation, leading to a further decrease of the free-energy barrier, which reduces to 0.17 and 0.13 eV at 4% and 8% doping, rationalizing the possibility for the thermal release of free carriers at the impurity concentrations that actually correspond to the conductivity boosts seen in experiments.

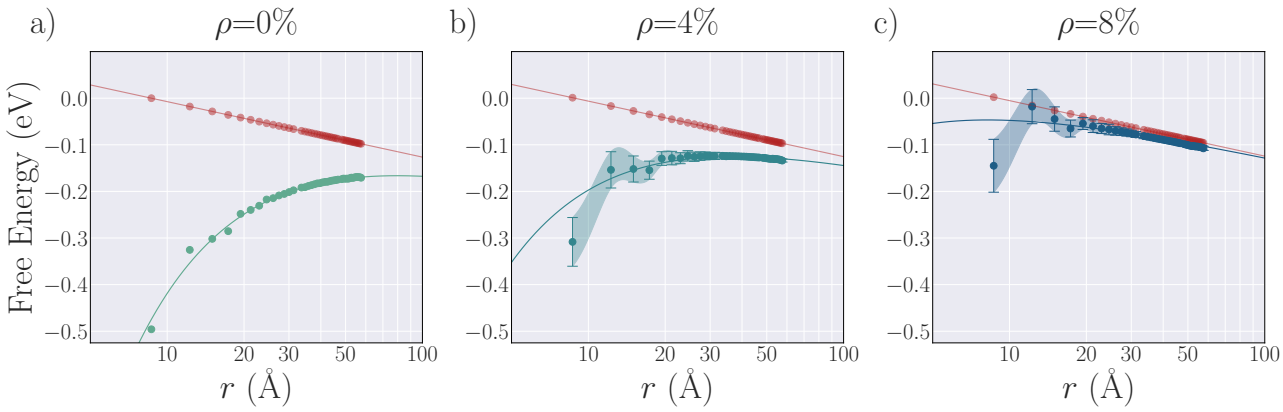

Figure 5.25: Room-temperature free energy profiles for charge separation from a source charge placed on a dopant site at $r = 0$, as a function of the electron-hole distance r for increasing doping loads (semi-logarithmic scale). Dots and shaded areas are the medians and median absolute deviations of the spherical average of the site free energies, while the full lines show the energy profile provided by the bulk dielectric constant ε. The entropic contributions are shown in red, dots correspond to numerical results, lines are analytic expressions for the continuum.

In addition to the free energy profiles for the charge release, our calculations provide insights into the effects of the inhomogeneity of the medium. Indeed, they determine a substantial spread in the energy profiles, with important local deviations from the screened Coulomb potential with bulk ε, which is recovered only at large r. The spatial inhomogeneity of the system can be best appreciated from the fluctuations of the local microscopic dielectric constant $\varepsilon(\mathbf{r})$, which is calculated as the ratio of the screened Coulomb potential with the bare one.

Fig. 5.26 presents the values of $\varepsilon(\mathbf{r})$ on lattice sites situated in the (111) plane of the FCC lattice. As expected, the pristine system shows an isotropic dielectric profile, but at larger doping loads the anisotropy is strengthened with values of the dielectric constant that can vary by an order of magnitude at $\rho = 8\%$.

Furthermore, we calculate the microscopic dielectric constant $\varepsilon(r)$ as a spherical average of the ratio of the screened and bare Coulomb potentials, namely

$$\varepsilon(r) = \frac{1}{N(r)} \sum_{r < r_i \leqslant r+\delta r} \frac{\widetilde{W}_i}{V_i}. \tag{5.45}$$

In the limit $r \to \infty$, $\varepsilon(r)$ tends to the bulk value that has been calculated previously. The median values and median absolute deviations of $\varepsilon(r)$ for increasing values of

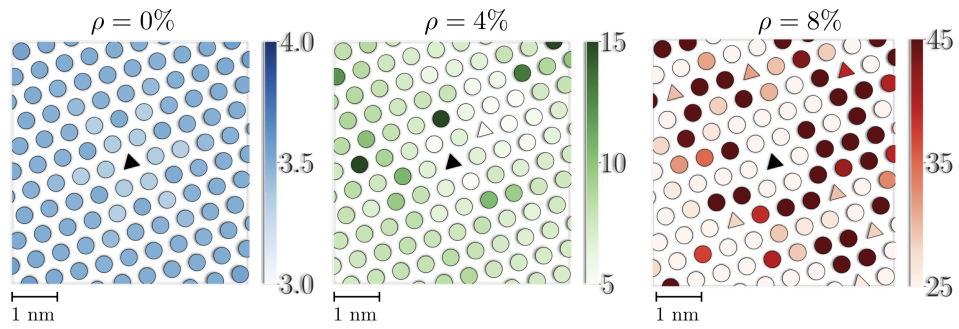

Figure 5.26: Illustration of the (111) plane of a system modeling a doped OSC at $\rho = 8\%$ doping load in the field of a source charge. Host and dopant molecules are shown as circles and triangles, respectively.

the doping load ρ are depicted in Fig. 5.27. We observe very strong local fluctuations in $\varepsilon(r)$, that can even result in local instabilities when $\varepsilon(r) < 0$. These fluctuations, which would be missed by reasoning only in terms of the bulk ε, may determine pathways for charge separation that are energetically more favorable than the average one.

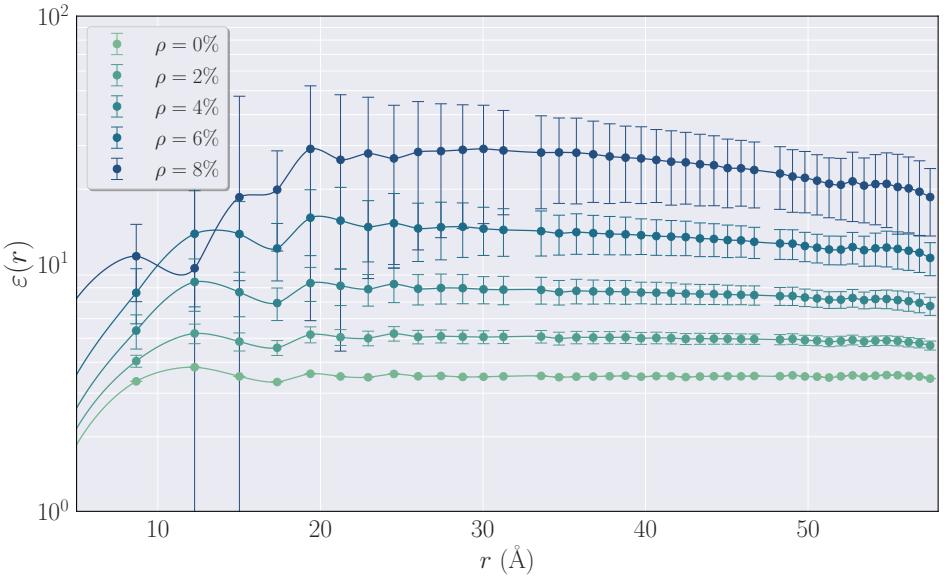

Figure 5.27: Non-local microscopic dielectric constant $\varepsilon(r)$ as a function of the electron-hole distance r for increasing doping loads (semi-logarithmic scale). The figure reports the medians and median absolute deviations of the spherical average of the site's values, while the full lines are interpolants serving as a guide to the eye.

5.4 SUMMARY

We have proposed a model description of collective screening phenomena in doped organic semiconductors and their implications on the release of Coulombically bound charges introduced upon doping. Our treatment, based on local polarizabilities of host molecules and highly polarizable host-dopant complexes, strictly holds upon approaching conducting states from the insulating side, namely for bound charges, and predicts a factor ten increase of the bulk dielectric constant at doping loads of 8% and very large fluctuations at the microscopic scale. This enhanced screening, together with entropic effects, drastically reduces the Coulomb barrier for charge separation, providing a mechanistic explanation for the impurity concentration dependence of the activation energies extracted from conductivity data [36]. The present analysis builds a fresh picture of the doping-induced insulator-to-conductor transition in terms of an incipient dielectric catastrophe, that is broadly consistent with available experimental data for organic semiconductors, including the 5-10% loads needed to boost conductivity [11, 13, 30].

Divergent dielectric responses are *de facto* expected in conducting states and the advent of the dielectric catastrophe upon doping is well documented both theoretically and experimentally in doped Silicon, where arguments similar to ours, linking the polarizability of hydrogen-like impurities to the dielectric constant ε, have been proposed [44, 45, 156, 157, 46, 167]. However, such a mechanism has not been discussed in organics, where conductivity enhancements upon doping are mostly described as due to the favorable effect of energetic disorder [17, 168, 169]. In OSCs the situation is in fact different, as we predict the advent of a dielectric catastrophe at doping loads of a few %, which are much larger than in Silicon, where is occurs at concentrations of the order of 0.01% [46]. Early studies demonstrated that the local impurity fields and the associated host polarization contribute to a large increase of the dielectric constant with doping [156, 157], a feature that emerges from our analysis of host-dopant interactions as well. However, the microscopic treatment proposed herein goes far beyond early works based on local-field corrections, as it fully captures host-dopant electrostatic interactions as well as the major role played by inhomogeneity and disorder at the molecular scale. Indeed, we have shown that soft impurity polarization bands involving mostly dopant units are formed at doping loads of a few %, and these modes strongly contribute to the dielectric properties. Moreover, the effect of inhomogeneity and disorder is to allow for more polarization modes to contribute to the dielectric properties, resulting in further enhancement of ε.

[167] Mott and Kaveh (1985).

[17] Tietze et al. (2018).

[168] Fediai et al. (2019).

[169] Koopmans et al. (2020).

Our findings also agree with the conductivity enhancements upon doping observed in earlier studies based on kinetic Monte Carlo simulations [17, 168, 169], yet providing additional microscopic insights on the charge release mechanism. In fact, our analysis indicates that the fundamental cause of the phenomenon lies in many-body screening phenomena, rather that in the direct effect of the energetic disorder sourced from the dipolar fields of host-dopant complexes, as it was previously assumed. Interestingly, the screening effects revealed here are in principle also captured in kinetic Monte Carlo simulations, where they could be specifically accessed by evaluating the space and time correlations of particles trajectories and of the electric polarization, which have not been reported to date. In the prevailing view, charge release is considered the result of the formation of energetically favorable pathways in the rough electrostatic field of many ionized host-dopant complexes [170, 17, 36].

[170] Mityashin et al. (2012).

In a strongly interacting environment such as the doped systems studied here, the formation of such energetically favorable pathways can only occur if the Coulomb binding of a probe electron-hole pair is counteracted by the electrical potentials of the other moving charges. The charge rearrangements that make this possible are, in essence, the collective screening processes studied here.

5.5 Appendix: Dispersion of the Polarization Normal Modes

In this Appendix, we present the continuation of the polarization normal modes analysis that has been carried out in Sec. 5.3.1. We have seen that in the inhomogeneous, doped system, many more modes contribute to the response as compared to the pristine case. The host and dopant polarization modes were distinguished by the fact that they involve mostly their respective sites. At this point, we focused solely on a real space analysis of such modes. Their study can be brought one step further by resolving the polarization normal modes dispersion in reciprocal space. Some of the results reported in this Appendix are still preliminary.

5.5.1 Unitary Transformation of the Hessian

In analogy with the standard treatment of phonons in solids, the dispersion of the polarization modes can be carried out by expressing their wavevector dependence in reciprocal space. Exploiting the symmetry of the lattice, this is achieved by performing the Fourier Transform (FT) of the Hessian matrix of Eq. (5.12) in order to associate each eigenmode to a polarization wave with a definite wavevector $\mathbf{q}$. Indeed, provided that the Hessian describes a homogeneous and periodic system, which is the case for the pristine system at $\rho = 0\%$, Bloch's theorem states that plane waves form a suitable eigenbasis. As a result the Hessian will be block-diagonalized by the unitary transformation corresponding to the plane wave basis, namely

$$U_{ij}^{\alpha\beta}(\mathbf{q}) = \frac{\delta_{\alpha\beta}}{\sqrt{N_c}} e^{-i\mathbf{r}_i\cdot\mathbf{q}_j}, \tag{5.46}$$

where $\mathbf{r}_i$ and $\mathbf{q}_j$ are the lattice positions and reciprocal vectors, and the normalization factor ensures the unitarity of the matrix. The sought-after form of the Hessian matrix is obtained by applying the aforementioned change of basis

$$\underline{T}(\mathbf{q}) = \underline{U}^{\dagger}(\mathbf{q}) \cdot \underline{T} \cdot \underline{U}(\mathbf{q}). \tag{5.47}$$

According to Bloch's theorem, the transformed Hessian $\underline{T}(\mathbf{q})$ will then be block-diagonal, where each 3×3 block describes the polarization modes with wavevector $\mathbf{q}$ pertaining to each of the 3 polarization bands. Upon diagonalization of each block, the full spectrum of the Hessian is recovered and each eigenstate corresponds to a polarization wave that can be uniquely labeled by its wavevector and band index.

The unitary transformation $\underline{U}(\mathbf{q})$ must be built from the reciprocal lattice defined as the 3D Fourier transform of the direct lattice, namely an FCC lattice in our case. Since we are considering a cubic supercell of N_c FCC conventional unit cells per dimension, our system contains $4N_c^3$ sites in real space. It therefore follows that the first Brillouin Zone (BZ) in reciprocal space will be sampled by the same number of independent reciprocal lattice points. Exploiting the fact that the FCC lattice is a cubic lattice with a basis, the complete set of $4N_c^3$ independent reciprocal lattice vectors can be constructed as

$$\mathbf{q}_{n,j} = \mathbf{K}_n + \mathbf{k}_j. \tag{5.48}$$

where $\mathbf{K}_n$ are the N_c^3 reciprocal vectors associated with the cubic supercell and $\mathbf{k}_j$ are the 4 reciprocal vectors associated with a single conventional cube. It can be shown

that they take the form

$$\mathbf{K}_n = \frac{2\pi}{a}\left(\frac{n_1}{N_c}\hat{\mathbf{x}} + \frac{n_2}{N_c}\hat{\mathbf{y}} + \frac{n_3}{N_c}\hat{\mathbf{z}}\right) \tag{5.49a}$$

$$\mathbf{k}_j = \frac{2\pi}{a}\left\{\mathbf{0}, \quad \hat{\mathbf{x}}, \quad \hat{\mathbf{y}}, \quad \hat{\mathbf{z}}\right\} \tag{5.49b}$$

where $n_i \in [0, N_c - 1]$ are integers and a is the lattice constant.

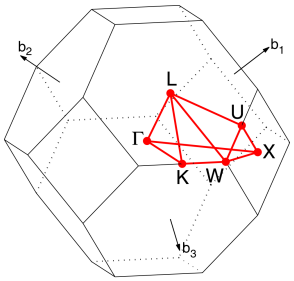

Figure 5.28: First Brillouin zone and high-symmetry points of the reciprocal FCC lattice. Taken from [171].

5.5.2 Reciprocal Space Analysis of the Polarization Normal Modes

As we have just discussed, transforming the Hessian into the basis of plane waves allows to resolve each polarization normal mode λ in terms of its wavevector $\mathbf{q}$ and band index $i = 1, 2, 3$, resulting in three distinct bands $\lambda_i(\mathbf{q})$. Since we consider a cubic simulation box of N_c^3 conventional FCC cells with $N_c = 10$, our system is composed of $4N_c^3$ sites and each band in the first BZ is sampled by the same number of $\mathbf{q}$ points. By unfolding a high-symmetry path in the first BZ, as depicted in Fig. 5.28, we present the dispersion of the polarization modes $\lambda_i(\mathbf{q})$ in Fig. 5.29.

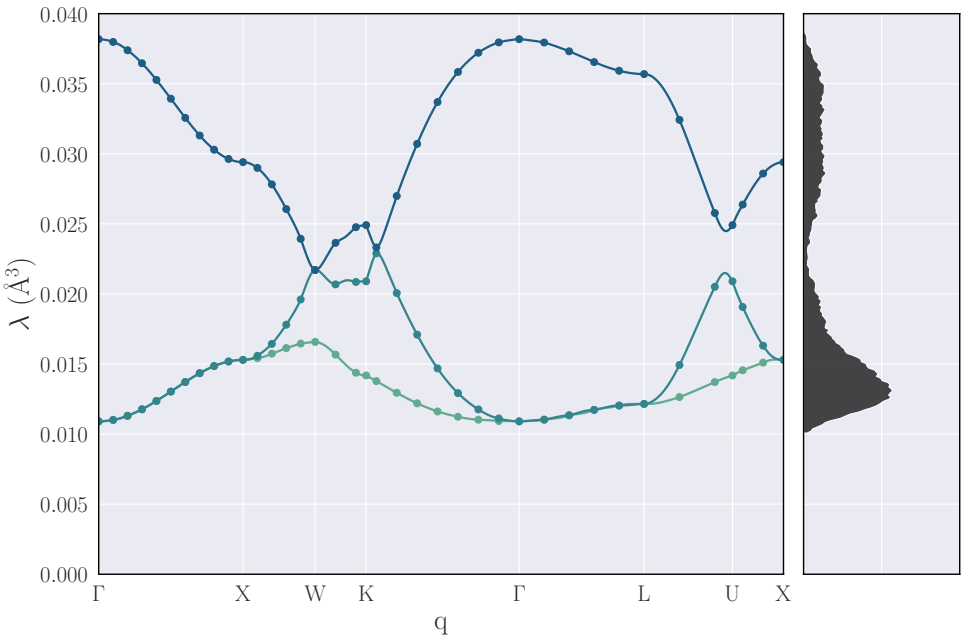

Figure 5.29: Polarization band dispersion along a high-symmetry path in the first BZ of the FCC lattice. The longitudinal band (blue) and the two transverse bands (turquoise and green) of the pristine system are obtained by interpolation in reciprocal space of the simulated values (dots). The right panel shows the corresponding density of polarization states.

The polarization modes of the pristine system pertain to three distinct branches, which can be assigned to longitudinal and transverse polarization bands from the relative alignment of the mode's polarization with the wavevector direction, which is respectively parallel or perpendicular. This results in two transverse polarization bands $\lambda^T(\mathbf{q})$ (turquoise and green) that are degenerate along the Δ ($\Gamma - X$) and Λ ($\Gamma - L$) high-symmetry paths, and one longitudinal polarization band $\lambda^L(\mathbf{q})$ (blue). The bimodal shape of the density of states is now seen to emerge from the longitudinal band, resulting in the broad peak in the DoS centered at $\lambda = 0.03$ Å^{-3}, and the two transverse bands, giving rise to the more pronounced peak centered at $\lambda = 0.013$ Å^{-3}. We further observe that the central sparse region of the DoS corresponds to a highly dispersive region at the BZ border, while the edges of the DoS correspond to the poorly dispersive region in reciprocal space around the Γ ($\mathbf{q} = 0$) extremal point, where the bands are almost flat. In particular, the softer modes, associated with the strongest polarizability, are the transverse polarization modes located at Γ.

It can be shown that because of the long-range nature of electrostatic interactions, longitudinal modes are inherently less polarizable than the softer transverse modes because the macroscopic fields act against the polarization in the former case while they support the polarization in the latter [172, 19]. Such a splitting between longitudinal and transverse modes, similar to that associated with polar modes in ionic crystals, explains why the longitudinal band $\lambda^{L}(\mathbf{q})$ is located in the upper region of the spectrum, while the transverse bands $\lambda^{T}(\mathbf{q})$ are the most polarizable.

[172] Ashcroft and Mermin (1976).

However, this splitting at Γ does not stem directly from the diagonalization of the Hessian. In fact, we found previously that the eigensubspace of the Hessian T_Γ pertaining to Γ, which completely determines the response ζ, is composed of three degenerate eigenmodes at $\lambda_0 = 1/\alpha_{\text{host}} = 0.02$ Å^{-3}, corresponding to orthogonal in-phase and fully delocalized polarization modes. The degeneracy of the eigenmodes at Γ is inconsistent with the dispersion of the polarization modes and makes both the band dispersion non-analytic and the Γ point not an extremum.

This incompatibility is rooted in the conditional convergence of dipole field sums, which are dependent on the order of summation. Equivalently, the Fourier Transform of the dipole-dipole interaction is proportional to $q^\alpha q^\beta/q^2$, where α, β denote Cartesian directions, making the limit $\mathbf{q} \to 0$ dependent on the direction. Because of the way we construct the Hessian matrix, the eigensubspace of the Hessian T_Γ corresponds to taking the limit $\mathbf{q} \to 0$ with increasing spherical shells in real space (bulk limit), whereas the transverse and longitudinal nature of the modes at Γ must be recovered by taking this limit first along the plane perpendicular to $\mathbf{q}$ and then along $\mathbf{q}$ (film limit) [18].

18. These two limits are associated with different macroscopic shapes, i.e. a sphere and a thin film, and thus to distinct depolarization tensors. The depolarization tensor is $\Delta_{ij} = \delta_{ij}/3$ for a sphere and $\Delta_{ij} = \delta_{iz}\delta_{jz}$ for a thin film.

Hence, the limiting eigenvalues $\lambda^{T}(\mathbf{q} \to 0)$ and $\lambda^{L}(\mathbf{q} \to 0)$ can be addressed by accounting for the macroscopic fields due to the non-vanishing polarization of transverse and longitudinal modes in the long-wavelength limit, which act as distinct electrostatic restoring forces. These limiting eigenvalues were obtained analytically for the Lorentz dipole lattice [173], which is an equivalent formulation of the Hessian we consider [19]. The Hessian's eigenvalues in the limit $\mathbf{q} \to 0$ of the longitudinal and transverse modes are then found to be

[173] Bagchi (1969).

$$\lambda(\mathbf{q} \to 0, L) = \frac{1}{4\pi\varepsilon_0\alpha_{\text{host}}} + \frac{2}{3\varepsilon_0}n \tag{5.50a}$$

$$\lambda(\mathbf{q} \to 0, T) = \frac{1}{4\pi\varepsilon_0\alpha_{\text{host}}} - \frac{1}{3\varepsilon_0}n \tag{5.50b}$$

where n is the number density of the system. We note that the bulk dielectric constant of the pristine system ($\varepsilon = 3.5$) is left unchanged by this correction.

19. The Lorentz dipole lattice is a generalization of the Wigner crystal in which each site bears a finite polarizability. The potential energy is given precisely by the Hessian of the pristine system we are considering, and the model further accounts for the kinetic energy. The translation between the two models can be made by noting the plasma frequency is $\omega_p^2 = 4\pi n$. The resulting equations of motion were solved in the limit $\mathbf{q} \to 0$ by Bagchi [173].

The analysis of the polarization modes dispersion of the doped systems is more involved, since the result of inhomogeneity is to break translational invariance, making plane waves no longer a suitable eigenbasis of the system. In this case, polarization normal modes are not uniquely labeled by a wavevector $\mathbf{q}$ and a band index. However, plane waves can still be employed as a complete basis over which the polarization modes can be characterized by their projection. We thus compute for 100 realizations of dopant positions the projection matrices

$$P_{ij} = \sum_k \left|\langle \mathbf{O}_i | \mathbf{q}_j, k\rangle\right|^2 \tag{5.51}$$

where $\mathbf{O}_i$ is the i^{th} eigenvector of the doped system, $\mathbf{q}_j$ is an allowed reciprocal vector in the first BZ and $|\mathbf{q}, k\rangle$ is the plane wave corresponding the k^{th} band. The spectral

resolution of these projections results in the weights of each mode of eigenvalue λ over the modes at reciprocal vector **q**, which is presented in Fig. 5.30 for a doping load ρ = 8%.

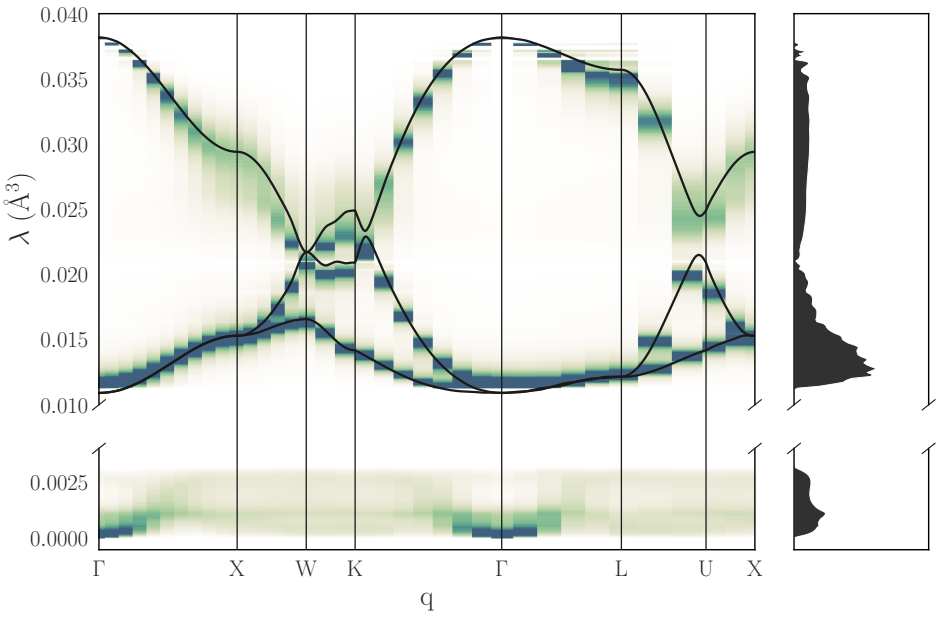

Figure 5.30: Projections of the eigenvectors of doped samples at ρ = 8% over the pristine polarization band dispersion along a high-symmetry path in the first Brillouin zone of the FCC lattice. The left panel shows the projections of the doped eigenvectors (density plot) and the three bands of the pristine system (black lines). The right panel shows the DoS of the doped samples, where the dopant region has been magnified for better visualization. The projections at Γ are obtained by extrapolation.

In the host spectral region $\lambda > 10^{-2}$ Å^{-3}, we observe a moderate broadening around pristine bands, which mostly retain their identity. Thus, the presence of a large number of randomly dispersed dopant impurities seems not to alter the host's band structure. In the dopant spectral region $\lambda < 10^{-2}$ Å^{-3}, we witness the formation of the dopant polarization bands which, despite being less defined, feature a dispersion that is reminiscent of that of the host. Also in this case we recognize a bimodal density of states arising from two lowest transverse bands a an upper longitudinal one. The dispersion of the soft modes associated with transverse bands is characterized by an appreciable dispersion around the extremal Γ point.

It appears that the fusion of the longitudinal band with the upper transverse band, at X in the host dispersion, happens midway between Γ and X in the dopant dispersion. We interpret this as the result of the fact that dopant bands are associated to longer inter-site distances, since dopant modes involve mostly dopant sites, resulting in an effectively renormalized Brillouin Zone. Indeed, the typical inter-dopant distance is larger than the nearest neighbor distance of the lattice by a factor $\rho^{-1/3}$, which at ρ = 8% gives a factor ≈ 2.3. This results in a reduced BZ with wavevectors that are typically 2.3 times smaller, which is in qualitative agreement with the position of the fusion of the dopant transverse bands.

In summary, we have performed a complete spectral analysis of the polarization normal modes arising upon doping. We showed that in the pristine system, only modes featuring a net polarization at Γ contribute to the response, while at finite doping loads the formation of a highly polarizable dopant band leads to a strong contribution to the response, arising from several polarization normal modes. The host and dopant bands are distinct and are seen to involve only the respective sites. This gives rise to a reciprocal-space dispersion of the dopant bands that is renormalized by the longer dopant-dopant distances.

Chapter 6

Conclusion

6 Conclusion

Prendre son temps est le meilleur
moyen de n'en pas perdre.
Nicolas Bouvier

Summary

THROUGHOUT THIS THESIS, WE STUDIED THE DOPING MECHANISMS in Organic Semiconductors both in the infinite dilution limit and at finite impurity concentration. We discussed the two related elementary steps of doping, namely the charge transfer and charge release processes. The variety and complexity of the interactions present in organic semiconductors drove us to adopt a multiscale approach. In particular, we combined state-of-the-art embedded Green's function Many-Body Perturbation Theory, classical Micro-Electrostatics and *ab initio*-parameterized Hamiltonian models to study the environmental effects on charge transfer and release, unveiling new physical phenomena and setting the basis for future research.

In the dilute limit (Chapter 4), we studied the charge transfer energetics between a single dopant molecule in a doped organic polymer and discussed the strong interplay between the nanostructure and the doping efficiency. In particular, we demonstrated that when the dopant is stacked on top of the conjugated backbone, the strong orbital hybridization between the polymer valence band and the acceptor level results only in a partial charge transfer in the ground state. Conversely, when the dopant is interlaced between the Alkyl side chains of the polymer, the favorable electrostatic landscape of the host material results in a full charge transfer in the ground state. Our calculations demonstrate that dopant molecules inserted into different regions of the host polymer may feature different Electron Affinities, being hence more or less prone to accept an electron. On the experimental side, this calls for a precise control of the preparation protocols of doped organic polymers in order to maximize the doping efficiency. Our study also confirmed that the Ionization Potential of the host and the Electron Affinity of the dopant are not the only parameter controlling the ionization energetics. Indeed, the excitonic electron-hole Coulomb binding and the host-dopant transfer integrals play a key role in the energetics of charge transfer in the low-doping regime, determining the nature of the ground state and the presence of low-lying CT excitations in organic host-dopant complexes.

At finite doping concentrations (Chapter 5), we evidenced the advent of a dielectric catastrophe in doped organic semiconductors at the doping loads corresponding to the conductivity boosts seen experimentally, namely 5-8%. We first demonstrated that host-dopant complexes featuring either fractional or full charge transfer are highly polarizable, with a polarizability that is typically one order of magnitude larger than than of a single host molecule. Then, we developed a novel approach to the modeling of the doped OSCs approaching conducting states from the insulating side, which allowed us to explore the microscopic details of the material's dielectric properties as a function of the impurity concentration. By studying the collective polarization

modes of randomly distributed dopants at increasing concentrations, we unveiled the formation of soft dopant polarization bands which strongly contributes to the system's response, resulting in a tenfold increase of the bulk dielectric constant at 8% doping load. The microscopic charge dissociation energy profiles lead to a remarkably low average free energy barrier that is comparable with room-temperature thermal energy, with large fluctuations due to the inhomogeneity of the doped material. The present analysis proposes on solid theoretical grounds that the microscopic mechanism by which free carriers are released in doped organic semiconductors is the collective screening of the Coulomb interactions between the carriers and the ionized dopants by means of highly polarizable host-dopant complexes. We are confident this work will stimulate experimental studies targeting the observation of low-energy electronic excitations responsible for the large polarizability of host-dopant complexes, as well as the direct measurement of the dielectric constant of doped organic semiconductors. A thorough understanding of the dielectric properties of doped organic semiconductors represents a promising gateway to the control of their electric properties.

Perspectives

Despite the theoretical advancements in the understanding of the doping mechanisms in organic semiconductors achieved in the present Thesis, many aspects still remain unexplored.

Our study on the charge transfer processes in doped polymers (Chapter 4) stimulated several pathways for methodological improvements. In particular, it highlighted the limits of spin-restricted DFT, and of the Many-Body Perturbation Theory (MBPT) built on top of it, in molecular complexes characterized by a large charge transfer in their ground state. Indeed, the constraint of a double occupancy of supramolecular orbitals represents a strong bias to the nature of the ground state, and a full CT has been obtained only by means of a spin-unrestricted DFT formalism. The generalization of the *GW* and BSE approaches to spin-unrestricted ground states would allow to compute the associated quasiparticle excitations, resulting in a more faithful description of charge-transfer complexes.

Moreover, research work is ongoing in the host laboratory to go beyond the static Δ^{COHSEX} embedding by considering explicitly the frequency dependence of the environmental dielectric function, within the framework of a fragment-based *GW* scheme (Ph.D. Thesis of David Amblard). This stands as an important goal since most embedding schemes, from PCMs to QM/MM atomistic approaches, assume the environment's dielectric constant to be frequency-independent in the optical range by taking the low-frequency limit; an approximation of which the validity has hardly been discussed so far. Preliminary results indicate that the static limit leads to overestimating the polarization (induction) energies by about 10% in organic systems.

Our investigations on the charge release mechanisms in doped OSCs (Chapter 5) also inspires several new research questions. A first interrogation regards the universality of the collective screening mechanism for charge release. We asserted the advent of the dielectric catastrophe to be very robust with respect to donor and acceptor relative energy levels on the basis of a Hamiltonian model, and explicitly validated our findings on two paradigmatic doped molecular systems. Nevertheless, we believe

that further investigations, possibly carried out by means of different theoretical tools, would benefit and consolidate the universality of our findings.

Furthermore, future work may address some open questions regarding the assumptions and approximations underpinning our Micro-Electrostatic modeling of doped OSCs that has been put aside in the first place. Firstly, at the finite concentrations we considered, dopants start to be very close to each other and potentially form clusters, an eventuality that we explicitly disregarded. The investigation of the electronic properties of dopant clusters, such as the effects of direct hybridization between dopants with the eventual formation of an impurity band in the presence of disorder, may be another approach to the study of the insulator-to-metal transition, and lies well beyond the present work. Secondly, the polarizability of host-dopant complexes was calculated for isolated complexes, namely in the infinite dilution limit, by embedding them in an environment that reproduces the pristine organic semiconductor. Recalculating it in the presence of a doped environment reproducing the bulk dielectric constant that we obtained from our model would yield a more reliable self-consistent evaluation.

Our model tackles the insulator-to-metal transition from the insulating side, which inherently restricts attainable concentrations to relatively low values, allowing to follow the phenomena determining the charge release, but preventing the study of conducting states. Indeed, typical experiments show that the conductivity of OSCs steadily increases with doping load and reaches a maximum at typical concentrations of 10%, then it saturates and eventually decreases. Beyond this limit, hosts and dopants coexist in similar proportions, thus altering the very nature of the material. Describing doped organic semiconductors near and beyond the transition requires lifting the assumption of bound localized charges, which can be achieved with itinerant electron models. Specifically, such models target the problem without making any assumption on the bound or free nature of the charge carriers by including charge-transfer hopping integrals between host and dopants as well as long-range Coulomb interactions between mobile carriers and dopants. Kinetic Monte Carlo (KMC) simulations provide a possible pathway to the description of the equilibrium charge dynamics at finite temperature in a classical framework. According to linear response theory, these simulations can grant access to the frequency-dependent dielectric function from the time auto-correlation of the sample polarization at different doping loads. Research in this direction is ongoing in collaboration with colleagues at the Karlsruhe Institute of Technology.

Quantum model Hamiltonians for itinerant electrons on a lattice represent a complementary direction of investigation that has been recently proposed in the host laboratory [41]. The basic ingredients entering the model are essentially the same as in KMC, the difference lying in the fully quantum treatment of the hopping terms, allowing to account for charge delocalization or fractional dopant ionization. Because of the exponential growth of the Hilbert space, the exact solution of the model becomes prohibitive beyond ~ 20 sites, a size which hardly can capture the complex features of doped organic semiconductors. Hence, ongoing efforts have been done focusing on the Hartree-Fock solution for spinless fermions, computing transport properties within the framework of the Transient Localization Theory. This allowed to show that at very high doping loads (quarter filling and beyond) the transport is essentially limited by disorder, rather than by Coulomb interactions, in very good agreement with experiments. This promising approach is however just at its begin-

[41] Jacobs et al. (2021).

nings, calling for in-depth future studies. The characterization of the transport and the dielectric properties as a function of the dopants concentration and temperature, possibly beyond the present mean-field approach, represent exciting developments of the work started in this Thesis.

Bibliography

[1] A. Einstein, "On a heuristic point of view concerning the production and transformation of light", *Annalen der Physik* **17** (1905).

[2] O. Bubnova, Z. U. Khan, A. Malti, S. Braun, M. Fahlman, M. Berggren, and X. Crispin, "Optimization of the thermoelectric figure of merit in the conducting polymer poly (3, 4-ethylenedioxythiophene)", *Nature materials* **10** (2011).

[3] V. A. Dediu, L. E. Hueso, I. Bergenti, and C. Taliani, "Spin routes in organic semiconductors", *Nature materials* **8** (2009).

[4] S. R. Forrest, "The path to ubiquitous and low-cost organic electronic appliances on plastic", *Nature* **428** (2004).

[5] F. Kagawa, S. Horiuchi, M. Tokunaga, J. Fujioka, and Y. Tokura, "Ferroelectricity in a one-dimensional organic quantum magnet", *Nature Physics* **6** (2010).

[6] G. D'avino, M. Souto, M. Masino, J. K. Fischer, I. Ratera, X. Fontrodona, G. Giovannetti, M. J. Verstraete, A. Painelli, P. Lunkenheimer, et al., "Conflicting evidence for ferroelectricity", *Nature* **547** (2017).

[7] V. Podzorov, E. Menard, A. Borissov, V. Kiryukhin, J. A. Rogers, and M. E. Gershenson, "Intrinsic Charge Transport on the Surface of Organic Semiconductors", *Physical Review Letters* **93** (2004).

[8] S. Fratini, D. Mayou, and S. Ciuchi, "The transient localization scenario for charge transport in crystalline organic materials", *Advanced Functional Materials* **26** (2016).

[9] S. Fratini, S. Ciuchi, D. Mayou, G. T. De Laissardière, and A. Troisi, "A map of high-mobility molecular semiconductors", *Nature materials* **16** (2017).

[10] I. Salzmann and G. Heimel, "Toward a comprehensive understanding of molecular doping organic semiconductors (review)", *Journal of Electron Spectroscopy and Related Phenomena* **204** (2015).

[11] K. Walzer, B. Maennig, M. Pfeiffer, and K. Leo, "Highly efficient organic devices based on electrically doped transport layers", *Chemical reviews* **107** (2007).

[12] I. Salzmann, G. Heimel, S. Duhm, M. Oehzelt, P. Pingel, B. M. George, A. Schnegg, K. Lips, R.-P. Blum, A. Vollmer, and N. Koch, "Intermolecular Hybridization Governs Molecular Electrical Doping", *Physical Review Letters* **108** (2012).

[13] I. Salzmann, G. Heimel, M. Oehzelt, S. Winkler, and N. Koch, "Molecular electrical doping of organic semiconductors: fundamental mechanisms and emerging dopant design rules", *Accounts of chemical research* **49** (2016).

[14] K. Kang, S. Watanabe, K. Broch, A. Sepe, A. Brown, I. Nasrallah, M. Nikolka, Z. Fei, M. Heeney, and D. Matsumoto, "2D coherent charge transport in highly ordered conducting polymers doped by solid state diffusion", *Nat. mater.* **15** (2016).

[15] R.-Q. Png, M. C. Ang, M.-H. Teo, K.-K. Choo, C. G. Tang, D. Belaineh, L.-L. Chua, and P. K. Ho, "Madelung and Hubbard interactions in polaron band model of doped organic semiconductors", *Nature communications* **7** (2016).

[16] C. Gaul, S. Hutsch, M. Schwarze, K. S. Schellhammer, F. Bussolotti, S. Kera, G. Cuniberti, K. Leo, and F. Ortmann, "Insight into doping efficiency of organic semiconductors from the analysis of the density of states in n-doped C60 and ZnPc", *Nature Materials* **17** (2018).

[17] M. L. Tietze, J. Benduhn, P. Pahner, B. Nell, M. Schwarze, H. Kleemann, M. Krammer, K. Zojer, K. Vandewal, and K. Leo, "Elementary steps in electrical doping of organic semiconductors", *Nature Communications* **9** (2018).

[18] W. Kohn and J. Luttinger, "Theory of donor states in silicon", *Physical Review* **98** (1955).

[19] C. Kittel, *Introduction to solid state physics*, Vol. 8 (Wiley New York, 1976).

[20] A. Troisi and G. Orlandi, "Band structure of the four pentacene polymorphs and effect on the hole mobility at low temperature", *The Journal of Physical Chemistry B* **109** (2005).

[21] K. Doi, K. Yoshida, H. Nakano, A. Tachibana, T. Tanabe, Y. Kojima, and K. Okazaki, "Ab initio calculation of electron effective masses in solid pentacene", *Journal of applied physics* **98** (2005).

[22] R. C. Hatch, D. L. Huber, and H. Höchst, "HOMO band structure and anisotropic effective hole mass in thin crystalline pentacene films", *Physical Review B* **80** (2009).

[23] S. Tavazzi, L. Raimondo, L. Silvestri, P. Spearman, A. Camposeo, M. Polo, and D. Pisignano, "Dielectric tensor of tetracene single crystals: The effect of anisotropy on polarized absorption and emission spectra", *The Journal of chemical physics* **128** (2008).

[24] M. Dressel, B. Gompf, D. Faltermeier, A. Tripathi, J. Pflaum, and M. Schubert, "Kramers-Kronig-consistent optical functions of anisotropic crystals: generalized spectroscopic ellipsometry on pentacene", *Optics express* **16** (2008).

[25] D. T. Duong, C. Wang, E. Antono, M. F. Toney, and A. Salleo, "The chemical and structural origin of efficient p-type doping in P3HT", *Organic Electronics* **14** (2013).

[26] W. Gao and A. Kahn, "Electrical doping: the impact on interfaces of π-conjugated molecular films", *Journal of Physics: Condensed Matter* **15** (2003).

[27] S. Reineke, F. Lindner, G. Schwartz, N. Seidler, K. Walzer, B. Lüssem, and K. Leo, "White organic light-emitting diodes with fluorescent tube efficiency", *Nature* **459** (2009).

[28] K. Harada, M. Sumino, C. Adachi, S. Tanaka, and K. Miyazaki, "Improved thermoelectric performance of organic thin-film elements utilizing a bilayer structure of pentacene and 2, 3, 5, 6-tetrafluoro-7, 7, 8, 8-tetracyanoquinodimethane (F 4-TCNQ)", *Applied Physics Letters* **96** (2010).

[29] H. Kleemann, C. Schuenemann, A. A. Zakhidov, M. Riede, B. Lüssem, and K. Leo, "Structural phase transition in pentacene caused by molecular doping and its effect on charge carrier mobility", *Organic Electronics* **13** (2012).

[30] I. E. Jacobs and A. J. Moulé, "Controlling molecular doping in organic semiconductors", *Advanced Materials* **29** (2017).

[31] I. E. Jacobs, C. Cendra, T. F. Harrelson, Z. I. Bedolla Valdez, R. Faller, A. Salleo, and A. J. Moulé, "Polymorphism controls the degree of charge transfer in a molecularly doped semiconducting polymer", *Mater. Horiz.* **5** (2018).

[32] P. Pingel and D. Neher, "Comprehensive Picture of p-Type Doping of P3HT With the Molecular Acceptor F_4TCNQ", *Physical Review B* **87** (2013).

[33] B. Maennig, M. Pfeiffer, A. Nollau, X. Zhou, K. Leo, and P. Simon, "Controlled p-type doping of polycrystalline and amorphous organic layers: Self-consistent description of conductivity and field-effect mobility by a microscopic percolation model", *Physical Review B* **64** (2001).

[34] Y. Shen, K. Diest, M. H. Wong, B. R. Hsieh, D. H. Dunlap, and G. G. Malliaras, "Charge transport in doped organic semiconductors", *Physical Review B* **68** (2003).

[35] S. Olthof, S. Mehraeen, S. K. Mohapatra, S. Barlow, V. Coropceanu, J.-L. Brédas, S. R. Marder, and A. Kahn, "Ultralow Doping in Organic Semiconductors: Evidence of Trap Filling", *Physical Review Letters* **109** (2012).

[36] M. Schwarze, C. Gaul, R. Scholz, F. Bussolotti, A. Hofacker, K. S. Schellhammer, B. Nell, B. D. Naab, Z. Bao, D. Spoltore, K. Vandewal, J. Widmer, S. Kera, N. Ueno, F. Ortmann, and K. Leo, "Molecular parameters responsible for thermally activated transport in doped organic semiconductors", *Nature Materials* **18** (2019).

[37] J. Li, G. D'Avino, A. Pershin, D. Jacquemin, I. Duchemin, D. Beljonne, and X. Blase, "Correlated electron-hole mechanism for molecular doping in organic semiconductors", *Physical Review Materials* **1** (2017).

[38] J. Li, I. Duchemin, O. Roscioni, P. Friederich, M. Anderson, E. Da Como, G. Kociok-Köhn, W. Wenzel, C. Zannoni, D. Beljonne, X. Blase, and G. D'Avino, "Host dependence of the electron affinity of molecular dopants", *Materials Horizons* **6** (2019).

[39] A. Privitera, G. Londi, M. Riede, G. D'Avino, and D. Beljonne, "Molecular Quadrupole Moments Promote Ground-State Charge Generation in Doped Organic Semiconductors", *Adv. Funct. Mater.* **30** (2020).

[40] F. Zhang and A. Kahn, "Investigation of the High Electron Affinity Molecular Dopant F6-TCNNQ for Hole-Transport Materials", *Advanced Functional Materials* **28** (2018).

[41] I. E. Jacobs, G. D'Avino, Y. Lin, V. Lemaur, Y. Huang, X. Ren, D. Simatos, W. Wood, C. Chen, T. Harrelson, T. Mustafa, C. A. O'Keefe, L. Spalek, D. Tjhe, M. Statz, L. Lai, P. A. Finn, W. G. Neal, J. Strzalka, C. B. Nielsen, J.-K. Lee, S. Barlow, S. R. Marder, I. McCulloch, S. Fratini, D. Beljonne, and H. Sirringhaus, "Ion-exchange doped polymers at the degenerate limit: what limits conductivity at 100% doping efficiency?", *arXiv* (2021).

[42] M. L. Tietze, P. Pahner, K. Schmidt, K. Leo, and B. Lüssem, "Doped Organic Semiconductors: Trap-Filling, Impurity Saturation, and Reserve Regimes", *Adv. Funct. Mater.* **25** (2015).

[43] S. D. Ha and A. Kahn, "Isolated molecular dopants in pentacene observed by scanning tunneling microscopy", *Physical Review B* **80** (2009).

[44] K. F. Herzfeld, "On Atomic Properties which make an Element a Metal", *Phys. Rev.* **29** (1927).

[45] N. Mott and M. Littleton, "Conduction in polar crystals. I. Electrolytic conduction in solid salts", *Transactions of the Faraday Society* **34** (1938).

[46] T. F. Rosenbaum, R. F. Milligan, M. A. Paalanen, G. A. Thomas, R. N. Bhatt, and W. Lin, "Metal-insulator transition in a doped semiconductor", *Physical Review B* **27** (1983).

[47] M. Comin and L. J. Lewis, "Deep-learning approach to the structure of amorphous silicon", *Physical Review B* **100** (2019).

[48] F. Reinert and S. Hüfner, "Photoemission spectroscopy—from early days to recent applications", *New Journal of Physics* **7** (2005).

[49] S. Hüfner, *Photoelectron spectroscopy: principles and applications* (Springer Science & Business Media, 2013).

[50] H. Fujiwara, *Spectroscopic Ellipsometry: Principles and Applications* (John Wiley & Sons, 2007).

[51] R. D. Mattuck, *A guide to Feynman diagrams in the many-body problem* (Courier Corporation, 1992).

[52] M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln", *Annalen der physik* **389** (1927).

[53] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J.-L. Brédas, "Charge Transport in Organic Semiconductors", *Chemical Reviews* **107** (2007).

[54] F. Neese, "The ORCA program system", *WIREs Comput. Mol. Sci.* **2** (2012).

[55] F. Bechstedt, *Many-body approach to electronic excitations* (Springer, 2016).

[56] R. M. Martin, *Electronic structure: basic theory and practical methods* (Cambridge university press, 2020).

[57] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas", *Physical Review* **136** (1964).

[58] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects", *Physical Review* **140** (1965).

[59] L. H. Thomas, "The calculation of atomic fields", *Mathematical proceedings of the Cambridge philosophical society* **23** (1927).

[60] E. Fermi, "Un metodo statistico per la determinazione di alcune priorieta dell'atome", *Rend. Accad. Naz. Lincei* **6** (1927).

[61] P. A. Dirac, "Note on exchange phenomena in the Thomas atom", *Mathematical proceedings of the Cambridge philosophical society* **26** (1930).

[62] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method", *Physical Review Letters* **45** (1980).

[63] A. Zunger, J. Perdew, and G. Oliver, "A self-interaction corrected approach to many-electron systems: Beyond the local spin density approximation", *Solid State Communications* **34** (1980).

[64] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis", *Canadian Journal of physics* **58** (1980).

[65] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", *Physical Review B* **45** (1992).

[66] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán, "Assessing the performance of recent density functionals for bulk solids", *Physical Review B* **79** (2009).

[67] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, "Ab initio calculation of phonon dispersions in semiconductors", *Physical Review B* **43** (1991).

[68] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", *Physical Review Letters* **77** (1996).

[69] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Physical Review B* **37** (1988).

[70] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids", *Physical Review Letters* **91** (2003).

[71] Y. Zhao and D. G. Truhlar, "A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions", *The Journal of chemical physics* **125** (2006).

[72] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential", *The Journal of chemical physics* **118** (2003).

[73] S. Grimme, "Density functional theory with London dispersion corrections", *WIREs Computational Molecular Science* **1** (2011).

[74] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko, "Density functional theory is straying from the path toward the exact functional", *Science* **355** (2017).

[75] J. F. Janak, "Proof that $\frac{\partial E}{\partial n_i} = \epsilon$ in density-functional theory", *Physical Review B* **18** (1978).

[76] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, "Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy", *Physical Review Letters* **49** (1982).

[77] J. P. Perdew and M. Levy, "Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities", *Physical Review Letters* **51** (1983).

[78] C. Li and W. Yang, "On the piecewise convex or concave nature of ground state energy as a function of fractional number of electrons for approximate density functionals", *Journal of Chemical Physics* **146** (2017).

[79] I. Dabo, A. Ferretti, N. Poilvert, Y. Li, N. Marzari, and M. Cococcioni, "Koopmans' condition for density-functional theory", *Physical Review B* **82** (2010).

[80] L. Kronik, T. Stein, S. Refaely-Abramson, and R. Baer, "Excitation Gaps of Finite-Sized Systems from Optimally Tuned Range-Separated Hybrid Functionals", *Journal of Chemical Theory and Computation* **8** (2012).

[81] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory* (Courier Corporation, 2012).

[82] M. J. Frisch, J. A. Pople, and J. S. Binkley, "Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets", *The Journal of chemical physics* **80** (1984).

[83] T. H. Dunning Jr, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen", *The Journal of chemical physics* **90** (1989).

[84] X. Blase, C. Attaccalite, and V. Olevano, "First-principles *GW* calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications", *Physical Review B* **83** (2011).

[85] X. Blase and C. Attaccalite, "Charge-transfer excitations in molecular donor-acceptor complexes within the many-body Bethe-Salpeter approach", *Applied Physics Letters* **99** (2011).

[86] V. M. Galitskii and A. B. Migdal, "Application of quantum field theory methods to the many body problem", *Sov. Phys. JETP* **7** (1958).

[87] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: density-functional versus many-body Green's-function approaches", *Rev. Mod. Phys.* **74** (2002).

[88] P. C. Martin and J. Schwinger, "Theory of Many-Particle Systems. I", *Phys. Rev.* **115** (1959).

[89] L. Hedin, "New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem", *Phys. Rev.* **139** (1965).

[90] M. van Schilfgaarde, T. Kotani, and S. Faleev, "Quasiparticle Self-Consistent GW Theory", *Physical Review Letters* **96** (2006).

[91] F. Kaplan, M. E. Harding, C. Seiler, F. Weigend, F. Evers, and M. J. van Setten, "Quasi-Particle Self-Consistent GW for Molecules", *Journal of Chemical Theory and Computation* **12** (2016).

[92] D. Golze, M. Dvorak, and P. Rinke, "The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy", *Frontiers in Chemistry* **7** (2019).

[93] R. W. Godby, M. Schlüter, and L. J. Sham, "Self-energy operators and exchange-correlation potentials in semiconductors", *Physical Review B* **37** (1988).

[94] N. H. March, *Electron Correlations in the Solid State* (World Scientific Publishing Company, 1999).

[95] M. J. van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie, C. Yang, F. Weigend, J. B. Neaton, F. Evers, and P. Rinke, "GW100: Benchmarking G0W0 for Molecular Systems", *Journal of Chemical Theory and Computation* **11** (2015).

[96] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, "Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions", *New Journal of Physics* **14** (2012).

[97] I. Duchemin, J. Li, and X. Blase, "Hybrid and Constrained Resolution-of-Identity Techniques for Coulomb Integrals", *Journal of Chemical Theory and Computation* **13** (2017).

[98] J. W. R. Macdonald, G. Piana, M. Comin, E. v. Hauff, G. Kociok-Köhn, C. Bowen, P. Lagoudakis, G. D'Avino, and E. D. Como, "Charge transfer excitons in a donor–acceptor amphidynamic crystal: the role of dipole orientational order", *Materials Horizons* (2020).

[99] L. Hedin and S. Lundqvist, "Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids", in *Frederick Seitz and David Turnbull and Henry Ehrenreich*, Vol. 23, Solid State Physics (Academic Press, 1970).

[100] G. Bussi, "Effects of the Electron-Hole Interaction on the Optical Properties of Materials: the Bethe-Salpeter Equation", *Physica Scripta* **T109** (2004).

[101] X. Blase, I. Duchemin, and D. Jacquemin, "The Bethe–Salpeter equation in chemistry: relations with TD-DFT, applications and challenges", *Chem. Soc. Rev.* **47** (2018).

[102] I. Tamm, "Relativistic interaction of elementary particles", in *Selected Papers* (Springer, 1991).

[103] S. Dancoff, "Non-adiabatic meson theory of nuclear forces", *Physical Review* **78** (1950).

[104] A. Dreuw and M. Head-Gordon, "Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene Complexes", *Journal of the American Chemical Society* **126**, PMID: 15038755 (2004).

[105] P.-F. Loos, M. Comin, X. Blase, and D. Jacquemin, "Reference Energies for Intramolecular Charge-Transfer Excitations", *Journal of Chemical Theory and Computation* **17** (2021).

[106] E. V. Tsiper and Z. G. Soos, "Charge redistribution and polarization energy of organic molecular crystals", *Physical Review B* **64** (2001).

[107] G. D'Avino, L. Muccioli, C. Zannoni, D. Beljonne, and Z. Soos, "Electronic Polarization in Organic Crystals: A Comparative Study of Induced Dipoles and Intramolecular Charge Redistribution Schemes", *Journal of Chemical Theory and Computation* **10** (2014).

[108] G. D'Avino, L. Muccioli, F. Castet, C. Poelking, D. Andrienko, Z. G. Soos, J. Cornil, and D. Beljonne, "Electrostatic phenomena in organic semiconductors: fundamentals and implications for photovoltaics", *Journal of Physics: Condensed Matter* **28** (2016).

[109] F. Castet, P. Aurel, A. Fritsch, L. Ducasse, D. Liotard, M. Linares, J. Cornil, and D. Beljonne, "Electronic polarization effects on charge carriers in anthracene: A valence bond study", *Physical Review B* **77** (2008).

[110] P.-O. Löwdin, "On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals", *The Journal of Chemical Physics* **18** (1950).

[111] R. Mulliken, "Electronic population analysis on LCAO–MO molecular wave functions. II. Overlap populations, bond orders, and covalent bond energies", *The Journal of Chemical Physics* **23** (1955).

[112] C. M. Breneman and K. B. Wiberg, "Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis", *Journal of Computational Chemistry* **11** (1990).

[113] B. H. Besler, K. M. Merz Jr., and P. A. Kollman, "Atomic charges derived from semiempirical methods", *Journal of Computational Chemistry* **11** (1990).

[114] B. T. Thole, "Molecular polarizabilities calculated with a modified dipole interaction", *Chemical Physics* **59** (1981).

[115] P. Ewald, "Evaluation of optical and electrostatic lattice potentials", *Ann. Phys.* **64** (1921).

[116] Z. Soos, E. Tsiper, and R. Pascal Jr, "Charge redistribution and electronic polarization in organic molecular crystals", *Chemical physics letters* **342** (2001).

[117] C. Poelking and D. Andrienko, "Long-range embedding of molecular ions and excitations in a polarizable molecular environment", *Journal of chemical theory and computation* **12** (2016).

[118] F. Lipparini, L. Lagardère, B. Stamm, E. Cancès, M. Schnieders, P. Ren, Y. Maday, and J.-P. Piquemal, "Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: I. Toward Massively Parallel Direct Space Computations", *Journal of Chemical Theory and Computation* **10** (2014).

[119] M. Comin, "Tequila: A Micro-Electrostatic Software for the Spectral Analysis of Polarization Modes", (2022).

[120] S. Duhm, G. Heimel, I. Salzmann, H. Glowatzki, R. L. Johnson, A. Vollmer, J. P. Rabe, and N. Koch, "Orientation-dependent ionization energies and interface dipoles in ordered molecular assemblies", *Nature materials* **7** (2008).

[121] I. Salzmann, S. Duhm, G. Heimel, M. Oehzelt, R. Kniprath, R. L. Johnson, J. P. Rabe, and N. Koch, "Tuning the ionization energy of organic semiconductor films: the role of intramolecular polar bonds", *Journal of the American Chemical Society* **130** (2008).

[122] G. D'Avino, S. Duhm, R. G. Della Valle, G. Heimel, M. Oehzelt, S. Kera, N. Ueno, D. Beljonne, and I. Salzmann, "Electrostatic Interactions Shape Molecular Organization and Electronic Structure of Organic Semiconductor Blends", *Chemistry of Materials* **32** (2020).

[123] M. L. Tiago, J. E. Northrup, and S. G. Louie, "Ab initio calculation of the electronic and optical properties of solid pentacene", *Physical Review B* **67** (2003).

[124] K. Hummer, P. Puschnig, and C. Ambrosch-Draxl, "Lowest Optical Excitations in Molecular Crystals: Bound Excitons versus Free Electron-Hole Pairs in Anthracene", *Physical Review Letters* **92** (2004).

[125] S. Sharifzadeh, A. Biller, L. Kronik, and J. B. Neaton, "Quasiparticle and optical spectroscopy of the organic semiconductors pentacene and PTCDA from first principles", *Physical Review B* **85** (2012).

[126] P. Cudazzo, M. Gatti, and A. Rubio, "Excitons in molecular crystals from first-principles many-body perturbation theory: Picene versus pentacene", *Physical Review B* **86** (2012).

[127] M. Karplus, M. Levitt, and A. Warshel, "The nobel prize in chemistry 2013", *Nobel Media* (2013).

[128] A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme", *Journal of molecular biology* **103** (1976).

[129] I. Duchemin, D. Jacquemin, and X. Blase, "Combining the GW formalism with the polarizable continuum model: A state-specific non-equilibrium approach", *The Journal of chemical physics* **144** (2016).

[130] J. Li, G. D'Avino, I. Duchemin, D. Beljonne, and X. Blase, "Combining the Many-Body *GW* Formalism with Classical Polarizable Models: Insights on the Electronic Structure of Molecular Solids", *The Journal of Physical Chemistry Letters* **7** (2016).

[131] J. Li, G. D'Avino, I. Duchemin, D. Beljonne, and X. Blase, "Accurate description of charged excitations in molecular solids from embedded many-body perturbation theory", *Physical Review B* **97** (2018).

[132] I. Duchemin, C. A. Guido, D. Jacquemin, and X. Blase, "The Bethe–Salpeter formalism with polarisable continuum embedding: reconciling linear-response and state-specific features", *Chemical Science* **9** (2018).

[133] T. Siegrist, C. Kloc, J. H. Schön, B. Batlogg, R. C. Haddon, S. Berg, and G. A. Thomas, "Enhanced Physical Properties in a Pentacene Polymorph", *Angewandte Chemie International Edition* **40** (2001).

[134] T. Sakanoue and H. Sirringhaus, "Band-like temperature dependence of mobility in a solution-processed organic semiconductor", *Nature Materials* **9** (2010).

[135] H. Yoshida, K. Yamada, J. Tsutsumi, and N. Sato, "Complete description of ionization energy and electron affinity in organic solids: Determining contributions from electronic polarization, energy band dispersion, and molecular orientation", *Physical Review B* **92** (2015).

[136] M. Comin, V. Lemaur, A. GIUNCHI, D. Beljonne, X. Blase, and G. D'Avino, "Doping of semicrystalline conjugated polymers: dopants within alkyl chains do it better", *Journal of Materials Chemistry C* (2022).

[137] H. Mendez, G. Heimel, S. Winkler, J. Frisch, A. Opitz, K. Sauer, B. Wegner, M. Oehzelt, C. Rothel, S. Duhm, D. Tobbens, N. Koch, and I. Salzmann, "Charge-transfer crystallites as molecular electrical dopants", *Nature Communications* **6** (2015).

[138] J. E. Cochran, M. J. N. Junk, A. M. Glaudell, P. L. Miller, J. S. Cowart, M. F. Toney, C. J. Hawker, B. F. Chmelka, and M. L. Chabinyc, "Molecular Interactions and Ordering in Electrically Doped Polymers: Blends of PBTTT and F $_4$ TCNQ", *Macromolecules* **47** (2019).

[139] I. E. Jacobs, C. Cendra, T. F. Harrelson, Z. I. Bedolla Valdez, R. Faller, A. Salleo, and A. J. Moulé, "Polymorphism controls the degree of charge transfer in a molecularly doped semiconducting polymer", *Material Horizons* (2018).

[140] A. Hamidi-Sakr, L. Biniek, J.-L. Bantignies, D. Maurin, L. Herrmann, N. Leclerc, P. Lévêque, V. Vijayakumar, N. Zimmermann, and M. Brinkmann, "A Versatile Method to Fabricate Highly In-Plane Aligned Conducting Polymer Films with Anisotropic Charge Transport and Thermoelectric Properties: The Key Role of Alkyl Side Chain Layers on the Doping Mechanism", *Advanced Functional Materials* **27** (2017).

[141] D. T. Scholes, P. Y. Yee, J. R. Lindemuth, H. Kang, J. Onorato, R. Ghosh, C. K. Luscombe, F. C. Spano, S. H. Tolbert, and B. J. Schwartz, "The Effects of Crystallinity on Charge Transport and the Structure of Sequentially Processed F4TCNQ-Doped Conjugated Polymer Films", *Advanced Functional Materials* **27** (2017).

[142] H. Hase, K. O'Neill, J. Frisch, A. Opitz, N. Koch, and I. Salzmann, "Unraveling the Microstructure of Molecularly Doped Poly(3-hexylthiophene) by Thermally Induced Dedoping", *The Journal of Physical Chemistry C* **122** (2018).

[143] V. Vijayakumar, P. Durand, H. Zeng, V. Untilova, L. Herrmann, P. Algayer, N. Leclerc, and M. Brinkmann, "Influence of dopant size and doping method on the structure and thermoelectric properties of PBTTT films doped with F6TCNNQ and F4TCNQ", *J. Mater. Chem. C* **8** (2020).

[144] V. Untilova, H. Zeng, P. Durand, L. Herrmann, N. Leclerc, and M. Brinkmann, "Intercalation and Ordering of F6TCNNQ and F4TCNQ Dopants in Regioregular Poly(3-hexylthiophene) Crystals: Impact on Anisotropic Thermoelectric Properties of Oriented Thin Films", *Macromolecules* **54** (2021).

[145] B. J. Topham and Z. G. Soos, "Ionization in organic thin films: Electrostatic potential, electronic polarization, and dopants in pentacene films", *Physical Review B* **84** (2011).

[146] C. P. Theurer, M. Richter, D. Rana, G. Duva, D. Lepple, A. Hinderhofer, F. Schreiber, P. Tegeder, and K. Broch, "Coexistence of Ion Pairs and Charge-Transfer Complexes and Their Impact on Pentacene Singlet Fission", *The Journal of Physical Chemistry C* **125** (2021).

[147] M. Brinkmann, V. S. Videva, A. Bieber, J. J. André, P. Turek, L. Zuppiroli, P. Bugnon, M. Schaer, F. Nuesch, and R. Humphry-Baker, "Electronic and Structural Evidences for Charge Transfer and Localization in Iodine-Doped Pentacene", *The Journal of Physical Chemistry A* **108** (2004).

[148] R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro, and B. Kirtman, "Quantum-mechanical condensed matter simulations with CRYSTAL", *WIREs Computational Molecular Science* **8** (2018).

[149] D. Niedzialek, I. Duchemin, T. B. de Queiroz, S. Osella, A. Rao, R. Friend, X. Blase, S. Kümmel, and D. Beljonne, "First Principles Calculations of Charge Transfer Excitations in Polymer–Fullerene Complexes: Influence of Excess Energy", *Advanced Functional Materials* **25** (2015).

[150] H. Tanaka, A. Wakamatsu, M. Kondo, S. Kawamura, S.-i. Kuroda, Y. Shimoi, W.-T. Park, Y.-Y. Noh, and T. Takenobu, "Microscopic observation of efficient charge transport processes across domain boundaries in donor-acceptor-type conjugated polymers", *Communications Physics* **2** (2019).

[151] M. Comin, S. Fratini, X. Blase, and G. D'Avino, "Doping-Induced Dielectric Catastrophe Prompts Free-Carrier Release in Organic Semiconductors", *Advanced Materials* (2021).

[152] J. Gerratt and I. M. Mills, "Force constants and dipole-moment derivatives of molecules from perturbed Hartree–Fock calculations. I", *The Journal of Chemical Physics* **49** (1968).

[153] R. S. Mulliken, "Molecular compounds and their spectra. II", *Journal of the American Chemical Society* **74** (1952).

[154] R. S. Mulliken and W. B. Person, "Molecular compounds and their spectra. XXI. Some general considerations", *Journal of the American Chemical Society* **91** (1969).

[155] E. F. Valeev, V. Coropceanu, D. A. da Silva Filho, S. Salman, and J.-L. Brédas, "Effect of Electronic Polarization on Charge-Transport Parameters in Molecular Organic Semiconductors", *Journal of the American Chemical Society* **128** (2006).

[156] G. M. Castellan and F. Seitz, *Semiconducting Materials* (1951).

[157] S. Dhar and A. H. Marshak, "Static dielectric constant of heavily doped semiconductors", *Solid-state electronics* **28** (1985).

[158] G. D'Avino, D. Vanzo, and Z. G. Soos, "Dielectric properties of crystalline organic molecular films in the limit of zero overlap", *The Journal of chemical physics* **144** (2016).

[159] H. Mueller, "Theory of the Photoelastic Effect of Cubic Crystals", *Physical Review* **47** (1935).

[160] C. B. Clark, "Coulomb Interactions in the Uniform-Background Lattice Model", *Phys. Rev.* **109** (1958).

[161] R. A. Coldwell-Horsfall and A. A. Maradudin, "Zero-Point Energy of an Electron Lattice", *Journal of Mathematical Physics* **1** (1960).

[162] J. Colpa, "Dipole fields and electric-field gradients in their dependence on the macroscopic and microscopic crystal parameters for orthorhombic and hexagonal lattices. I", *Physica* **56** (1971).

[163] C. K. Purvis and P. L. Taylor, "Dipole-field sums and Lorentz factors for orthorhombic lattices, and implications for polarizable molecules", *Physical Review B* **26** (1982).

[164] D. Vanzo, B. J. Topham, and Z. G. Soos, "Dipole-Field Sums, Lorentz Factors, and Dielectric Properties of Organic Molecular Films Modeled as Crystalline Arrays of Polarizable Points", *Adv. Funct. Mater.* **25** (2015).

[165] D. Pines, *Elementary excitations in solids* (CRC Press, 2018).

[166] B. A. Gregg, "Entropy of Charge Separation in Organic Photovoltaic Cells: The Benefit of Higher Dimensionality", *J. Phys. Chem. Lett.* **2** (2011).

[167] N. Mott and M. Kaveh, "Metal-insulator transitions in non-crystalline systems", *Adv. Phys.* **34** (1985).

[168] A. Fediai, F. Symalla, P. Friederich, and W. Wenzel, "Disorder compensation controls doping efficiency in organic semiconductors", *Nature Communications* **10** (2019).

[169] M. Koopmans, M. A. T. Leiviskä, J. Liu, J. Dong, L. Qiu, J. C. Hummelen, G. Portale, M. C. Heiber, and L. J. A. Koster, "Electrical Conductivity of Doped Organic Semiconductors Limited by Carrier–Carrier Interactions", *ACS Applied Materials & Interfaces* **12** (2020).

[170] A. Mityashin, Y. Olivier, T. Van Regemorter, C. Rolin, S. Verlaak, N. G. Martinelli, D. Beljonne, J. Cornil, J. Genoe, and P. Heremans, "Unraveling the Mechanism of Molecular Doping in Organic Semiconductors", *Advanced Materials* **24** (2012).

[171] W. Setyawan and S. Curtarolo, "High-throughput electronic band structure calculations: Challenges and tools", *Computational Materials Science* **49** (2010).

[172] N. W. Ashcroft and N. D. Mermin, *Solid state physics* (Holt, Rinehart and Winston, New York London, 1976).

[173] A. Bagchi, "Dielectric Properties of the Wigner and Related Dipole Lattices", *Phys. Rev.* **178** (1969).