

# Excitons in Polymer-Fullerene Morphologies

**Citation for published version (APA):**

Sundaram, V. (in press). *Excitons in Polymer-Fullerene Morphologies: Multiscale Modeling, Implementation and Application*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Applied Physics and Science Education]. Eindhoven University of Technology.

**Document status and date:**

Accepted/In press: 27/02/2024

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# **Excitons in Polymer-Fullerene Morphologies**

*Multiscale Modeling, Implementation &  
Application*

Vivek Sundaram

*Copyright © 2023 by V.Sundaram. All Rights Reserved.*

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Thesis title: Excitons in Polymer-Fullerene Morphologies  
Multiscale Modeling, Implementation & Application / by V.Sundaram.  
Technische Universiteit Eindhoven, 2023. Proefschrift.

Cover design by V. Sundaram

A catalogue record is available from the Eindhoven University of Technology Library

**ISBN: 978-90-386-5964-0**

The work in this thesis has been sponsored by Netherlands Organisation for Scientific Research (NWO) and the Netherlands eScience Center, within the Joint CSER and eScience program for Energy Research (JCER 2017). Also acknowledge support by the Innovational Research Incentives Scheme Vidi of the NWO with project number 723.016.002.

Printed by Leesterheide Grafisch

# Excitons in Polymer-Fullerene Morphologies

*Multiscale Modeling, Implementation &  
Application*

## PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr. S.K. Lenaerts, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op  
dinsdag 27 februari 2024 om 16:00 uur

door

Vivek Sundaram

geboren te Patna, India

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter: prof. dr. C. Storm

promotor: dr. B. Baumeier

co-promotor: dr. A.V. Lyulin

leden: prof. dr. ir. R.A.J. Janssen

prof. dr. L. Visscher (Vrije Universiteit, Amsterdam)

prof. dr. F. Schmid (Johannes Gutenberg Universität Mainz)

dr. S. Tao

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

# List of Publications

1. **Excited-state electronic structure of molecules using many-body Green's functions: Quasiparticles and electron-hole excitations with VOTCA-XTP**  
G. Tirimbó, V. Sundaram, O. Çaylak, W. Scharpach, J. Sijen, C. Junghans, J. Brown, F.Z. Ruiz, N. Renaud, J. Wehner, B. Baumeier  
*J. Chem Phys.* **152**, 114103 (2020)
2. **Development and Testing of an All-Atom Force Field for Diketopyrrolopyrrole Polymers with Conjugated Substituents**  
V. Sundaram, A.V. Lyulin, B. Baumeier  
*J. Phys. Chem. B* **124** 48, 11030-11039 (2020)
3. **Effect of Solvent Removal Rate and Annealing on the Interface Properties in a Blend of a Diketopyrrolopyrrole-Based Polymer with Fullerene**  
V. Sundaram, A.V. Lyulin, B. Baumeier  
*J. Phys. Chem. B* **126**, 38, 7445-7453 (2022)
4. **Time-Dependent Density Functional Theory and Green's Functions Methods with the Bethe–Salpeter Equation in *Excited States and Photodynamics: From Photobiology to Photomaterials.***  
W. Scharpach, Z. Chen, V. Sundaram, B. Baumeier  
Comprehensive Computational Chemistry, Elsevier, (in press)
5. **Quantum-quantum and quantum-classical schemes with projection-based-embedded *GW*-BSE**

V. Sundaram, B. Baumeier  
*in preparation*

**6. Excitonics at a DPP2Py2T-PCBM interface**

V. Sundaram, B. Baumeier  
*in preparation*

# Contents

<b>List of Publications</b>	<b>1</b>
<b>1 Introduction</b>	<b>6</b>
<b>2 Methodologies</b>	<b>13</b>
2.1 Density-Functional Theory . . . . .	14
2.1.1 Hartree–Fock Theory . . . . .	17
2.1.2 Hohenberg–Kohn Theorems . . . . .	20
2.1.3 Kohn–Sham Theory . . . . .	21
2.1.4 Computational aspects . . . . .	24
2.2 Many-Body Green’s Functions Methods for Electronically Excited States . . . . .	27
2.2.1 Charged Excitations and the Single-Particle Green’s Function . . . . .	28
2.2.2 Hedin’s Equations and the <i>GW</i> Approximation . . . . .	31
2.2.3 Two-Particle Excitations and the Bethe–Salpeter Equation . . . . .	41
2.3 Overview of Classical Molecular Dynamics . . . . .	50
2.3.1 Classical Potentials . . . . .	50
2.3.2 MD trajectories . . . . .	55
2.3.3 Simulation ensembles . . . . .	56
2.4 QM/MM embedding schemes for electronic excitations with <i>GW</i> -BSE . . . . .	56
<b>3 All-atom force field development for DPP2PymT</b>	<b>61</b>
3.1 Introduction . . . . .	61
3.2 Methodology . . . . .	63

3.3	$N_p$ -DPP2PymT force field . . . . .	66
3.3.1	Bond stretching . . . . .	66
3.3.2	Angle bending . . . . .	69
3.3.3	Dihedral angles . . . . .	70
3.3.4	Partial charge calculation . . . . .	73
3.4	$N_p$ -DPP2PymT in solution . . . . .	74
3.5	Melt simulations . . . . .	78
3.5.1	Glass transition . . . . .	78
3.5.2	Segmental mobility . . . . .	82
3.6	Summary . . . . .	84
<b>4</b>	<b>Interface Properties of DPP2PymT-PCBM[60] blends</b>	<b>85</b>
4.1	Selection of system size . . . . .	85
4.2	Methodology . . . . .	92
4.3	Results . . . . .	95
4.3.1	Glass Transition . . . . .	95
4.3.2	Segmental mobility . . . . .	96
4.3.3	Fullerene distribution around polymer . . . . .	99
4.3.4	Influence of side-chain alignment . . . . .	105
4.4	Summary . . . . .	107
<b>5</b>	<b>QM/QM and QM/QM/MM with PbE-GW-BSE</b>	<b>109</b>
5.1	Projection-based-embedded (PbE) Density-Functional Theory . . . . .	110
5.2	Truncation of the full atomic orbital basis (trPbE) . . . . .	113
5.3	Projection-based-embedded GW-BSE calculations . . . . .	117
5.3.1	Projection-based-embedded GW-BSE/MM . . . . .	120
5.4	Frontier orbital and near-gap electron-hole excitation energies . . . . .	121
5.4.1	DPP bicyclic ring with branched alkyl side-chains	122
5.4.2	Prodan in water . . . . .	128
5.4.3	Benzene-TCNE dimer in water . . . . .	134
5.5	Discussion . . . . .	138
5.5.1	Embedding cost . . . . .	141
5.6	Summary . . . . .	142

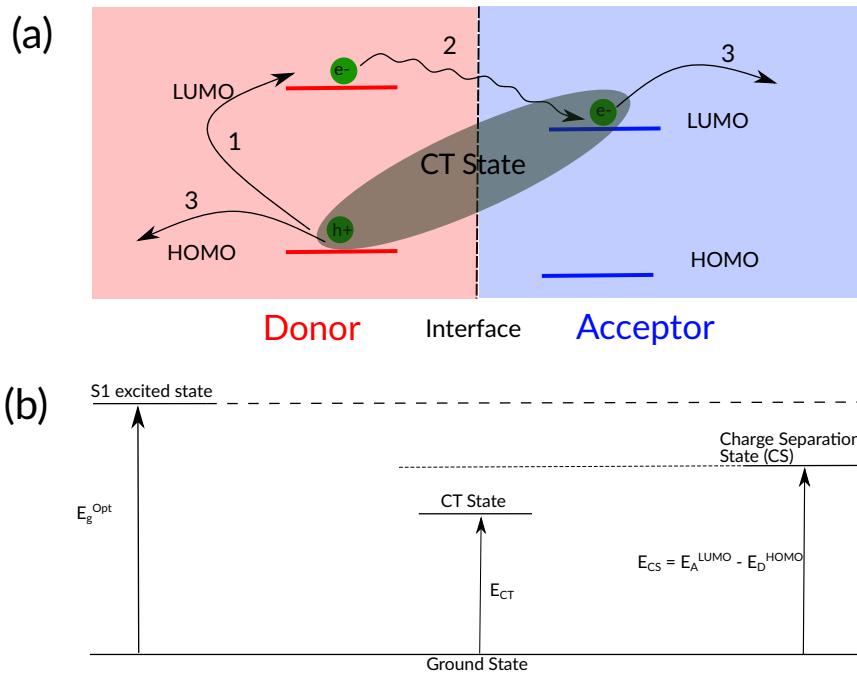
<b>6 Excitonics at a DPP2Py2T-PCBM interface</b>	<b>144</b>
6.1 Introduction . . . . .	144
6.2 Computational details . . . . .	145
6.2.1 Non-adiabatic LE-CT couplings . . . . .	147
6.2.2 Reorganization energy estimates . . . . .	148
6.3 Interface excitonics . . . . .	149
6.4 Summary & Discussion . . . . .	153
<b>7 Summary &amp; Conclusion</b>	<b>156</b>
<b>Acknowledgement</b>	<b>164</b>
<b>About the author</b>	<b>165</b>
<b>References</b>	<b>165</b>
<b>Index</b>	<b>187</b>

# Chapter 1

## Introduction

"We need renewable and green sources of energy." This is a sentence that has kept material scientists and chemists on the hunt for almost five decades now. Among the multiple sources of renewable energy, solar power has emerged as a viable resource. For instance, in 2022, electricity production remained unchanged compared to the previous year. Production from renewable sources rose by 20 percent to 47 billion kWh. The contribution from solar power production increased by 54 percent, while wind power production was up by 17 percent in the Netherlands [1]. Solar photovoltaics is becoming the lowest-cost option for new electricity generation in most of the world, which is expected to propel investment in the coming years.

Silicon is considered the industry standard material for the large scale production of solar cells. The theoretical maximum power conversion efficiency (PCE) of a single junction (electron rich and electron deficient) solar cell is 33.7 % as defined by the Shockley–Queisser limit [2, 3]. The majority of currently available solar cells are based on silicon and deliver a PCE of up to 26 %[4–7] which is about 80 % of the Shockley–Queisser limit. Such high efficiency is because the band-gap of silicon allows most of the light in the visible and UV spectrum to be absorbed efficiently which allows for the generation of an electron-hole pair (or exciton). In silicon, the binding energy for an exciton is 0.01 eV which is low enough for it to be dissociated at room temperature [8, 9]. Doping silicon with electron rich or deficient dopants can allow creation of electron and hole

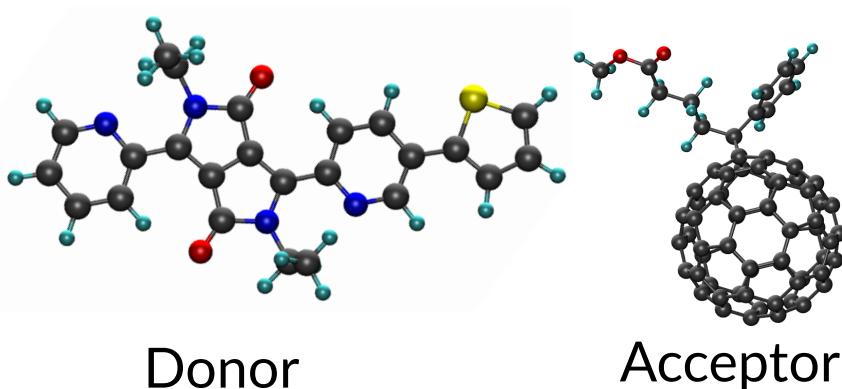


**Figure 1.1:** (a) Diagrammatic representation of the three stages of charge generation and separation in a bulk-heterojunction solar cell.  
 1: Exciton generation in the donor material  
 2: Exciton goes into the CT state.  
 3: Charges separate and move into the CS state beyond which they are transferred to the load.  
 (b) Energy diagram of the different states that play an important role in charge transfer in a bulk-heterojunction solar cell.

carriers respectively in the same material thereby forming a p-n junction. This junction in turn facilitates the transport of charge carriers from the materials to the electrodes. To obtain high conversion efficiencies, 80 % of the theoretical limit, requires, however, monocrystalline wafers whose processing causes high production costs. In contrast, devices made from less expensive polycrystalline or amorphous silicon wafers show a reduction of the PCE to around 10-14 % [10, 11].

Instead of using inorganic materials, the use of organic molecules (either polymers or small molecules) in photovoltaic devices is attractive due to their ease of processing from solution, higher flexibility due to low bulk modulus, and the non-toxicity of the compounds used [12, 13]. However, these materials require significantly higher energy for charge carrier generation. Low dielectric screening and substantial disorder in the materials generally lead to strongly bound, localized electronic excitations, hampering the efficient generation of free charge carriers and their transport processes [14]. Hence, what is sought after, is the use of a different acceptor molecule whose unoccupied levels are lower than that of its donor counterpart. This allows the formation of donor-acceptor heterojunctions. A schematic depiction of generation of local excitons, CT state and the HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital) energy levels, respectively, are shown in Figure 1.1. Light absorption in one of the materials, often in the donor material, generates a localized exciton. This exciton then moves to a state, commonly referred to as the charge transfer state (CT state) in which the excited electron from the LUMO of the donor moves to the LUMO of the acceptor while the hole rests on the HOMO of the donor. This facilitates the excitonic separation and then free charge carriers are generated which diffuse into the respective material to reach the electrodes [15–17]. Under ideal circumstances it is expected that the CT state has a lower energy compared to the localized excitons of both the donor and the acceptor. In short, current generation from a solar cell can be written in three steps broadly as numbered in Figure 1.1.

1. A photon is absorbed, and an electron is promoted from the HOMO to the LUMO of the donor material leaving a hole behind in the HOMO. This is characterized by the excitation energy written as  $E_g^{\text{Opt}}$  which is the difference between the band-gap (HOMO-LUMO gap) of the material and the exciton binding energy in the donor.
2. This step is followed by an exothermic reaction where the exciton moves in the CT state with the energy of  $E_{\text{CT}}$ .



**Figure 1.2:** Structural representation of a DPP-based monomer (donor) and a fullerene based PCBM (acceptor).

- Finally, these charges diffuse into the bulk of the systems and reach a charge-separated state (CS) which is an endothermic reaction to move from the CT state. This charge separated energy is the difference between the HOMO energy of the donor and the LUMO energy of the acceptor material.

It must be understood that the mechanism stated above is a simplification of the actual processes that govern charge separation, other factors such as reorganization energies, non-adiabatic couplings (see Chapter 6) play a vital role as well. However, we stick to the energy picture only for the moment. We can clearly see that it is now an uphill task to find two right materials, one as a donor and one as an acceptor, and ensure that the energy levels of these materials are favorably aligned to improve power conversion efficiency (PCE). To overcome these factors limiting the PCE of organic photovoltaic devices, significant effort is directed towards tuning electronic and structural properties of the materials by altering the architecture of donor and acceptor compounds [12, 18, 19].

Among the intensively studied donor-acceptor materials are oligomers based on diketopyrrolopyrroles (DPP) for donors and fullerene (C<sub>60</sub>) based compounds (PCBM) for acceptors [5, 20–22], see the unit in

Figure 1.2. They can be easily modified by adding various aromatic and  $\pi$ -conjugated substituents to the backbone and different side chains to DPP, affecting both the solubility and crystallinity of the polymer as well as its electronic properties. With suitably chosen substituents, a low band-gap internal donor-acceptor architecture leads to efficient harvesting of the solar energy spectrum and a PCE of around 8% in blends with fullerene acceptors [23–26].

In order to improve this PCE the initial step is to understand how the charge transfer actually happens and what factors play an important role in determining the extent of charge generation and transfer. It is known that charge separation may occur after excitation of either the donor or the acceptor. The efficiency of charge generation depends on the electron affinity and ionization potential of the donor and acceptor and the excited state energy. Energy barriers can be present that regulate the kinetics of the charge transfer process. However, what was observed in these systems is that excitation of the donor and acceptor to the first charge transfer state (CT1) happens with equal rate but for the second charge transfer state (CT2) excitations from the acceptor are orders of magnitude slower [27]. This difference is sometimes directly observable in the external quantum efficiency (EQE) spectra of polymer-fullerene solar cells. The main cause of this difference remained unaccounted for as the charge generation in these large systems are not vividly understood as the size of these systems do not permit a full quantum treatment for simulations within a reasonable time and using limited computational resources. In this thesis we intend to understand the mechanisms driving charge transfer using a multiscale modeling approach ranging from classical simulations using Molecular Dynamics to quantum simulation involving post Density Functional theory methods.

At this point we would like to leave the reader with some questions to ponder upon which will be the focal points of this thesis and shall be revisited at the end.

1. Can we model the different excited states, in particular the charge transfer states, in bulk heterojunction solar cells?

2. How do processing techniques have an influence on the local morphology of the blend?
3. Are current methods to model electronically excited states sufficiently scalable to systems of size which can accurately represent realistic materials?

From the above, more specific questions arise, e.g., what processing techniques are used experimentally, what are appropriate molecular simulation models for materials processing and what can be learned from them, what are the current methods for studying electronic excitations in complex molecular materials, and how are morphology and charge transfer linked [28]. We will address each of these questions in steps in the subsequent chapters.

## Thesis structure

This thesis is divided into three major parts:

- Classical simulation of the polymer-PCBM system to perform structural analysis.
- Understanding a hybrid embedding approach where more precise calculations are embedded in the classical model.
- Combining the knowledge of the polymer-PCBM morphology and embedding to perform charge transfer analysis on the actual polymer model.

Chapter 2 '*Methodologies*' summarizes the theoretical framework that is needed for a complete understanding of the techniques developed and employed in the remainder of the thesis. Topics covered in this part include some essential theory of quantum mechanics for many electron systems, specifics of density-functional theory (DFT), many-body Green's functions based approaches for electronic excitations (*GW*-BSE), classical molecular mechanics models for structure simulations, and combined quantum-classical methods.

In the next step we intend to proceed with the first challenge: classical simulation of the polymer-PCBM system. Chapter 3 '*All-atom Force-field generation using a Classical and Quantum approach*' discusses the primary steps needed to be done before starting a molecular dynamics simulation: Force-Field generation and its validation. Having generated a reliable force-field, Chapter 4 titled '*Structural analysis for DPP2Py2T-PCBM interface using Molecular Dynamics*' discusses how the force-field has been used to study the DPP-PCBM interface and the effects of the processing techniques on the morphology of the blend. This concludes the classical study of the DPP-PCBM interface.

With some representative atomistic morphologies at hand, we now have a building ground to study electronic excitations in them using the *GW*-BSE method. The size of the respective system make such simulations with traditional *GW*-BSE implementations for the whole system computationally infeasible. Hence, we developed a hybrid embedding scheme which embeds a fraction of the system that has been treated with *GW*-BSE within a DFT environment further encapsulating the embedded system into a classical MM environment. Chapter 5 '*Development and testing of a hybrid Quantum-Classical embedding scheme*' discusses how such a model has been developed and tested on three sample systems each depicting a different use case. Once the model has been tested it is finally used in Chapter 6 to study the '*Excitonics at a DPP2Py2T-PCBM interface*' including not only energy level calculations but also estimates for the rates of temperature-activated conversion processes between localized and charge-transfer excitations.

The thesis is concluded with a brief summary in Chapter 7.

# Chapter 2

## Methodologies

Based in parts on the paper: Gianluca Tirimbo, Vivek Sundaram, Onur Caylak, Wouter Scharpach, Javier Sijen, Christoph Junghans, Joshua Brown, Felipe Zapata Ruiz, Nicolas Renaud, Jens Wehner, Björn Baumeier, *J. Chem. Phys.* **152**, 114103 (2020) and the book chapter: Wouter Scharpach, Zhongquan Chen, Vivek Sundaram, Björn Baumeier, *Time-Dependent Density Functional Theory and Green's Functions Methods with the Bethe-Salpeter Equation* in "Excited States and Photodynamics: from Photobiology to Photomaterials", Elsevier (in print, 2023).

In this chapter we summarize the essentials of the various theoretical frameworks and derived computational methodologies used in this thesis. Beginning with density-functional theory for electronic ground states, we cover the theory underlying the many-body Green's function methods and Bethe–Salpeter Equations to model electron-hole excitations. We then give a brief overview of classical simulations using Molecular Dynamics in different ensembles and finally conclude the chapter with a method describing how a QM system can be coupled with a classical system. Classical simulation is needed for the generation of the bulk morphology (Chapter 3 and Chapter 4). Studies in Chapter 5 and Chapter 6 rely on quantum-quantum and quantum-classical embedding.

## 2.1 Density-Functional Theory

In this section, we briefly recapitulate the basics of density-functional theory as a method to obtain an effective single-particle description of the electronic ground state of an atomic, molecular or solid system made out of  $M$  atoms and  $N$  electrons. For the sake of a simpler presentation, we consider a spin-singlet, closed-shell system: most of the statements can easily be generalized to take the spin variables into account explicitly. Throughout this chapter we will use Hartree atomic units, i.e., with a reduced Planck constant  $\hbar = 1$ , all charges in units of the elementary charge, positions in units of the Bohr radius, and masses in units of the electron mass. The coordinates  $\mathbf{R}_\alpha$  of the individual nuclei with charges  $Z_\alpha$  and  $\mathbf{r}_i$  of the individual electrons are combined into the variables  $\overline{\mathbf{R}} = (\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M)$  and  $\overline{\mathbf{r}} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ , respectively. With this, the (non-relativistic) many-body Hamiltonian reads

$$\hat{H} = \underbrace{-\frac{1}{2} \sum_{\alpha=1}^M M_\alpha^{-1} \Delta_{\mathbf{R}_\alpha}}_{\hat{T}_{\text{nuc}}} + \underbrace{\frac{1}{2} \sum_{\substack{\alpha, \beta=1, \\ \alpha \neq \beta}}^M \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}}_{\hat{V}_{\text{nuc-nuc}}} - \frac{1}{2} \sum_{i=1}^N \Delta_{\mathbf{r}_i} + \underbrace{\frac{1}{2} \sum_{\substack{i, j=1, \\ i \neq j}}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\hat{V}_{\text{el-el}}} - \underbrace{\sum_{\alpha=1}^M \sum_{i=1}^N \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|}}_{\hat{V}_{\text{nuc-el}}}, \quad (2.1)$$

where  $\hat{T}$  and  $\hat{V}$  are the respective operators for the kinetic and potential energies involving the nuclear (nuc) and electronic (el) subsystems. The time evolution of the many-body wave function  $\Psi(\overline{\mathbf{r}}, \overline{\mathbf{R}}, t)$  is obtained by solving the time-dependent Schrödinger equation [29]

$$\hat{H}\Psi(\overline{\mathbf{r}}, \overline{\mathbf{R}}, t) = i \frac{\partial}{\partial t} \Psi(\overline{\mathbf{r}}, \overline{\mathbf{R}}, t) \quad (2.2)$$

and can be used to calculate the expectation value of an observable described by an operator  $\hat{A}$  as

$$A(t) = \langle \hat{A} \rangle_{\Phi} = \int \Psi^*(\bar{\mathbf{r}}, \bar{\mathbf{R}}, t) \hat{A} \Psi(\bar{\mathbf{r}}, \bar{\mathbf{R}}, t) d^{3N} \bar{\mathbf{r}} d^{3M} \bar{\mathbf{R}}. \quad (2.3)$$

This allows – formally – to extract information about structural, electronic and optical properties of the system, as well as the dynamic response to an external perturbation (described by an additional potential term in (2.1)) as measured, e.g., by spectroscopic techniques.

In practice however, (2.2) is exactly solvable only for  $M = N = 1$ , which is a hydrogen atom, so we will need to explore several approximations to make the problem tractable [30]. The standard method of solving a partial differential equation such as (2.2) is the method of separation of variables in which one makes a product function *ansatz*, i.e.,  $\Psi(\bar{\mathbf{r}}, \bar{\mathbf{R}}, t) = \Phi(\bar{\mathbf{r}}, \bar{\mathbf{R}})U(t)$ . If the Hamiltonian in (2.1) is not explicitly time-dependent, its expectation value, the total molecular energy  $E_{\text{mol}}$ , is constant according to (2.3), and the time evolution of the wave function is given by  $U(t) = C \exp(-iE_{\text{mol}}t)$ . The spatial component  $\Phi(\bar{\mathbf{r}}, \bar{\mathbf{R}})$  of the wave function and the total energy are obtained as solutions of the stationary Schrödinger equation

$$\hat{H}_{\text{mol}}\Phi(\bar{\mathbf{r}}, \bar{\mathbf{R}}) = E_{\text{mol}}\Phi(\bar{\mathbf{r}}, \bar{\mathbf{R}}). \quad (2.4)$$

Note that in (2.4) both  $\bar{\mathbf{r}}$  and  $\bar{\mathbf{R}}$  are explicit variables of this eigenvalue problem. Since the nuclei are much heavier than the electrons, one can further assume that the electrons adjust instantaneously to the nuclear motion, i.e., the electrons move adiabatically. To express this situation in formal terms, we consider a *fixed arrangement* of nuclei  $\bar{\mathbf{R}}$ . The Hamiltonian representing the electronic system that interacts with the fixed nuclear configuration reads

$$\hat{H}_{\text{el}} = \hat{H}_{\text{el}}(\bar{\mathbf{R}}) = \underbrace{\hat{T}_{\text{el}} + \hat{V}_{\text{nuc-el}}(\bar{\mathbf{R}})}_{\text{1-electron operator}} + \underbrace{\hat{V}_{\text{el-el}}}_{\text{2-electron operator}}. \quad (2.5)$$

In this situation,  $\bar{\mathbf{R}}$  is no longer a variable of the electronic system, but a fixed parameter for the electronic degrees of freedom. The corresponding stationary electronic Schrödinger equation is given by

$$\hat{H}_{\text{el}}(\bar{\mathbf{R}})\Phi_{\nu}(\bar{\mathbf{r}}; \bar{\mathbf{R}}) = E_{\nu}(\bar{\mathbf{R}})\Phi_{\nu}(\bar{\mathbf{r}}; \bar{\mathbf{R}}), \quad (2.6)$$

where  $\{\Phi_{\nu}(\bar{\mathbf{r}}; \bar{\mathbf{R}})\}$  is a set of adiabatic electronic wave functions. Those can be used as a basis to expand the molecular wave function  $\Phi(\bar{\mathbf{r}}, \bar{\mathbf{R}})$  according to

$$\Phi(\bar{\mathbf{r}}, \bar{\mathbf{R}}) = \sum_{\nu} \Xi_{\nu}(\bar{\mathbf{R}})\Phi_{\nu}(\bar{\mathbf{r}}; \bar{\mathbf{R}}). \quad (2.7)$$

Entering this *Born–Oppenheimer separated* [31] wave function into (2.4) yields – after some steps [32] – a coupled set of equations for the coefficients  $\{\Xi_{\nu}(\bar{\mathbf{R}})\}$ :

$$E_{\text{mol}}\Xi_{\nu}(\bar{\mathbf{R}}) = \left( E_{\nu}(\bar{\mathbf{R}}) + \hat{T}_{\text{nuc}} + \hat{V}_{\text{nuc-nuc}} \right) \Xi_{\nu}(\bar{\mathbf{R}}) + \sum_{\mu} A_{\nu\mu} \Xi_{\mu}(\bar{\mathbf{R}}), \quad (2.8)$$

where the integrals  $A_{\nu\mu}$  are matrix elements of the transition between electronic states  $\nu$  and  $\mu$  induced by the dynamics of the nuclei. In the adiabatic approximation it is assumed that  $A_{\mu\nu} = 0$ , i.e., there are no transitions between different electronic states, and the nuclear motion for each electronic state  $\nu$  is determined by

$$E_{\text{mol}}\Xi_{\nu}(\bar{\mathbf{R}}) = \left( E_{\nu}(\bar{\mathbf{R}}) + \hat{T}_{\text{nuc}} + \hat{V}_{\text{nuc-nuc}} \right) \Xi_{\nu}(\bar{\mathbf{R}}). \quad (2.9)$$

This describes the motion of the nuclei in an effective potential

$$U_{\nu}(\bar{\mathbf{R}}) = E_{\nu}(\bar{\mathbf{R}}) + \hat{V}_{\text{nuc-nuc}}(\bar{\mathbf{R}}). \quad (2.10)$$

It is to be noted by the reader that further electronic theory will not be treating the nuclear dynamics. The nuclear motion will be picked up again in Section 2.3 when we discuss the classical molecular mechanics

where we map the nuclear potential energy surfaces to the combination of electronic and nuclear potential energy surfaces.

### 2.1.1 Hartree–Fock Theory

The electronic Schrödinger equation (2.6) is in practice still not solvable for many-body systems, due to the presence of the electron-electron interaction  $\hat{V}_{\text{el-el}}$ . Without it, the electronic Hamiltonian is simply the sum of non-interacting single-particle Hamiltonians, i.e.,  $\hat{H}_{\text{el}} = \sum_{i=1}^N \hat{h}_i(\mathbf{r}_i)$ .

As  $[\hat{h}_i(\mathbf{r}_i), \hat{h}_j(\mathbf{r}_j)] = 0$ , the corresponding  $N$ -electron wave function  $\Phi_\nu^0$  is simply a product of single-particle functions

$$\Phi_\nu^0(\bar{\mathbf{r}}) = \prod_{i=1}^N \phi_{\nu_i}^0(\mathbf{r}_i) \quad (2.11)$$

and the total energy is given by

$$E_\nu^0 = \sum_{i=1}^N \varepsilon_{\nu_i}^0. \quad (2.12)$$

However, according to the Pauli principle, the electronic wave functions must be antisymmetric with respect to particle exchange, and therefore must change sign whenever the coordinates of two electrons are interchanged. To meet this requirement, the electronic wave function is constructed from single-particle functions  $\phi_i$  as a so-called Slater determinant [33]:

$$\Phi(\bar{\mathbf{r}}; \bar{\mathbf{R}}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \cdots & \phi_1(\mathbf{r}_N) \\ \phi_2(\mathbf{r}_1) & \cdots & \phi_2(\mathbf{r}_N) \\ \vdots & \ddots & \vdots \\ \phi_N(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix} \quad (2.13)$$

The idea of the Hartree–Fock theory is that instead of starting from predetermined single-particle functions and enforcing antisymmetry, we can start from the requirement of antisymmetry and use the variational principle[34] to derive a set of equations that determine suitable effective single particles for the interacting case. Let us take a closer look at the expression for the total electronic energy (suppressing the parametric dependence on  $\bar{\mathbf{R}}$ ) for any given wave function  $\Phi$ :

$$E[\Phi] = \int \Phi^*(\bar{\mathbf{r}}) \hat{H}_{\text{el}} \Phi(\bar{\mathbf{r}}) d^{3N} \bar{\mathbf{r}}. \quad (2.14)$$

The variational theorem states that this energy functional is minimal for the true ground-state wave function, i.e.,  $E[\Phi] \geq E_0$ , where  $E_0$  is the ground-state energy. Denote now  $\Phi^{\text{HF}}$  as the many-body wave function in Hartree–Fock theory as a Slater-determinant ansatz. By a variational principle, the energy as a functional of the determinant approximates the true ground-state energy  $E_0$ :

$$E[\Phi^{\text{HF}}] = \frac{\langle \Phi^{\text{HF}} | \hat{H}_{\text{el}} | \Phi^{\text{HF}} \rangle}{\langle \Phi^{\text{HF}} | \Phi^{\text{HF}} \rangle} \geq E_0. \quad (2.15)$$

Minimizing the above energy functional via the effective single-particle functions  $\phi_j$  under the constraint that they are normalized is achieved by

$$\frac{\delta}{\delta \phi_j^*(\mathbf{r})} \left\{ E[\Phi^{\text{HF}}] - \sum_{i=1}^N \varepsilon_i^{\text{HF}} \left( \int \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) d\mathbf{r} - 1 \right) \right\} = 0 \quad (2.16)$$

and yields a set of equations that allow to determine the functions  $\phi_j$ :

$$\begin{aligned} & \left\{ -\frac{\Delta_{\mathbf{r}}}{2} + \hat{V}_{\text{ext}}(\mathbf{r}) + \int n(\mathbf{r}') v_C(\mathbf{r}, \mathbf{r}') d\mathbf{r}' \right\} \phi_j^{\text{HF}}(\mathbf{r}) \\ & - \int n(\mathbf{r}, \mathbf{r}') v_C(\mathbf{r}, \mathbf{r}') \phi_j^{\text{HF}}(\mathbf{r}') d\mathbf{r}' = \varepsilon_j^{\text{HF}} \phi_j^{\text{HF}}(\mathbf{r}), \end{aligned} \quad (2.17)$$

with  $v_C = |\mathbf{r} - \mathbf{r}'|^{-1}$ . Here, we have also introduced the electronic densities

$$n(\mathbf{r}) = \sum_{i=1}^N \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}) \quad n(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^N \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}'). \quad (2.18)$$

The first integral in (2.17) corresponds to the classical Hartree integral [35] of the Coulomb interaction  $V_H(\mathbf{r}) = \int n(\mathbf{r}')v_C(\mathbf{r}, \mathbf{r}')d\mathbf{r}'$ , and the second integral defines the exchange potential operator  $\hat{V}_x(\mathbf{r}) = \int n(\mathbf{r}, \mathbf{r}')v_C(\mathbf{r}, \mathbf{r}') \cdots d\mathbf{r}'$ . The  $N$ -electron problem has thus been mapped on a set of effective single-particle problems with the Hartree–Fock potential

$$\hat{V}_{HF}(\mathbf{r}) = \hat{V}_{ext}(\mathbf{r}) + \hat{V}_H(\mathbf{r}) + \hat{V}_x(\mathbf{r}). \quad (2.19)$$

Considering the double counting of  $i$  and  $j$  interactions  $\phi_i^{HF}$  and  $\phi_j^{HF}$ , the total energy of the ground state is

$$E_0^{HF} = \sum_{i=1}^N \varepsilon_i^{HF} - \frac{1}{2}(E_H + E_x) \quad (2.20)$$

where

$$\begin{aligned} E_H &= \int n(\mathbf{r})v_C(\mathbf{r}, \mathbf{r}')n(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \\ E_x &= - \int n(\mathbf{r}, \mathbf{r}')v_C(\mathbf{r}, \mathbf{r}')n(\mathbf{r}', \mathbf{r})d\mathbf{r}d\mathbf{r}'. \end{aligned} \quad (2.21)$$

In summary, the Hartree–Fock theory assumes that the many-electron wave function takes the form of a Slater determinant. Since the exact wave functions cannot be expressed as single determinants, the problem with this assumption is that Hartree–Fock methods cannot fully represent the solution of the exact many-electron Schrödinger equation (2.6) and the corresponding total energy differs from the true ground-state energy. This difference is often referred to as *correlation* energy.

### 2.1.2 Hohenberg–Kohn Theorems

With a solution to the  $N$ -electron problem in (2.6), potentially from the Hartree–Fock approximation or some other theory, one can now determine observables from the wave function of the electronic ground state alone. However, the electronic Hamiltonian in (2.5) comprises only one- and two-electron operators, that is, operators that act on either a single electronic coordinate or two. This raises whether it is necessary to look for a solution in terms of an  $N$ -electron wave function after all.

Consider the second-order density matrix, defined as

$$P_2(\mathbf{r}'_1, \mathbf{r}'_2; \mathbf{r}_1, \mathbf{r}_2) = \frac{N(N-1)}{2} \times \int \Phi^*(\mathbf{r}'_1, \mathbf{r}'_2, \dots, \mathbf{r}_N) \Phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3\mathbf{r}_3 \dots d^3\mathbf{r}_N. \quad (2.22)$$

Its diagonal elements  $\bar{P}_2(\mathbf{r}_1, \mathbf{r}_2) = P_2(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2)$  form the *two-particle density matrix*. A first-order density matrix can be written in terms of  $P_2$  elements as

$$P_1(\mathbf{r}'_1; \mathbf{r}_1) = \frac{2}{N-1} \int P_2(\mathbf{r}'_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2) d^3\mathbf{r}_2, \quad (2.23)$$

whose diagonal element  $\bar{P}_1(\mathbf{r}_1) = P_1(\mathbf{r}_1; \mathbf{r}_1)$  is the *charge density*. Instead of using (2.14) to determine the total energy from the full  $3N$ -dimensional wave function  $\psi$ , one can obtain the same via

$$E = \int (\hat{T}_{\text{el}} + \hat{V}_{\text{nuc-el}}) P_1(\mathbf{r}_1) d^3\mathbf{r}_1 + \int \hat{V}_{\text{el-el}} \bar{P}_2(\mathbf{r}_1, \mathbf{r}_2) d^3\mathbf{r}_1 d^3\mathbf{r}_2, \quad (2.24)$$

which requires only information about a six-dimensional object. It seems attractive to minimize (2.24) directly by finding optimal density matrices  $P_1$  and  $P_2$ , under the constraint that they are constructible from a proper  $\psi$  that is antisymmetric with respect to exchange of electrons. However, this has in practice not been achieved reliably.

Hohenberg and Kohn realized that one does not even need  $\bar{P}_2$  to find the ground-state energy, and that it is instead *completely* determined by

the charge density  $n(\mathbf{r}) = \bar{P}_1(\mathbf{r})$  alone. Two theorems relate the ground state to the electron density:

1. The density  $n_0$ , which minimizes the ground-state energy, uniquely determines the external potential  $\hat{V}_{\text{ext}}(\mathbf{r})$  acting on the electronic system. The ground state  $\Phi_0$  is a one-to-one functional of the particle density  $n(\mathbf{r})$ .

(Note that in this theorem, external potentials are considered equivalent if they only differ an additional constant which is independent of space.)

2. The energy functional

$$E[n] = \int v_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \langle \Phi | \hat{T} + \hat{V}_{\text{el-el}} | \Phi \rangle \quad (2.25)$$

obeys a variational principle with respect to the particle density  $n(\mathbf{r})$  and is minimal for the ground-state density  $n_0(\mathbf{r})$ :

$$E_0 = E[n_0] \leq E[n]. \quad (2.26)$$

For proofs of these two theorems, the interested reader is referred to Ref. [36]. Above theorems restrict density-functional theory to studies of the ground state.

### 2.1.3 Kohn–Sham Theory

From the Hohenberg–Kohn theorems, it follows that the exact ground-state energy and density can be found by minimizing a universal energy functional  $E[n]$  under the constraint that  $\int n(\mathbf{r})d^3\mathbf{r} = N$ , so

$$\delta \left[ E[n] - \mu \left( \int n(\mathbf{r})d^3\mathbf{r} - N \right) \right] = 0 \quad (2.27)$$

for  $n(\mathbf{r}) = n_0(\mathbf{r})$ . The Lagrange multiplier  $\mu$  can physically be interpreted as a chemical potential. However, the expression of the functional in (2.25) is unsuitable for this purpose, as terms for the kinetic energy

and the electron-electron interaction are not expressed as functionals of the density. To work around this problem, Kohn and Sham proposed two approximations. First, they introduced a fictitious system of  $N$  non-interacting electrons reproducing the exact ground-state density, described by single-electron wave functions  $\phi_i(\mathbf{r})$ . They are used to construct an antisymmetric (with respect to electron exchange)  $N$ -electron wave function as a Slater determinant like in (2.13). For such a system, the kinetic energy is

$$T_s[n] = -\frac{1}{2} \sum_{i=1}^N \int \phi_i^*(\mathbf{r}) \Delta \phi_i(\mathbf{r}) d^3\mathbf{r} \quad (2.28)$$

and the density simply as in (2.18). For the electron-electron interaction, one splits off the classical Coulomb interaction, called the Hartree energy

$$E_H[n] = \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}', \quad (2.29)$$

so that the Kohn–Sham (KS) energy functional reads

$$E_{\text{KS}}[n] = T_s[n] + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + E_H[n] + E_{\text{xc}}[n]. \quad (2.30)$$

Here, the *exchange-correlation functional*  $E_{\text{xc}}[n]$  combines the differences of the true electron-electron interaction energy with all quantum effects and the classical Hartree energy as well as errors made in the kinetic energy expression:

$$E_{\text{xc}}[n] = (T[n] - T_s[n]) + (E_{\text{el-el}}[n] - E_H[n]). \quad (2.31)$$

Instead of finding the ground-state energy via variation with respect to the density (2.27), one can now perform a variation of (2.30) to find a set of equations to determine the orbitals  $\phi_i$  such that the density of the form  $n(\mathbf{r})$  as in (2.18) minimizes  $E_{\text{KS}}[n]$ . This yields a set of effective single-particle equations known as the Kohn–Sham equations [37, 38]

$$\left\{ -\frac{1}{2} \Delta + v_{\text{ext}}(\mathbf{r}) + v_H[n](\mathbf{r}) + \frac{\delta E_{\text{xc}}[n]}{\delta n}(\mathbf{r}) \right\} \phi_i^{\text{KS}}(\mathbf{r}) = \varepsilon_i^{\text{KS}} \phi_i^{\text{KS}}(\mathbf{r}), \quad (2.32)$$

with the effective Kohn–Sham Hamiltonian

$$\hat{H}_{\text{KS}} = -\frac{1}{2}\Delta + v_{\text{KS}}[n](\mathbf{r}) = -\frac{1}{2}\Delta + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}[n](\mathbf{r}) + \frac{\delta E_{\text{xc}}[n]}{\delta n}(\mathbf{r}). \quad (2.33)$$

Note that the  $\varepsilon_i^{\text{KS}}$  are formally only Lagrangian multipliers used to introduce the constraints that  $\int \phi_i^*(\mathbf{r})\phi_i(\mathbf{r})d^3\mathbf{r} = 1$  in the minimization, but are often interpreted as effective single-particle energies. There is little formal justification to equate them to actual excitations. Two points are noteworthy about (2.32).

First, the variation of the exchange-correlation energy functional defines a multiplicative exchange-correlation potential

$$v_{\text{xc}}[n](\mathbf{r}) = \frac{\delta E_{\text{xc}}[n]}{\delta n}(\mathbf{r}). \quad (2.34)$$

As the exact form of  $E_{\text{xc}}[n]$  is unknown, one has to resort to physically motivated approximations with varying accuracy. A thorough discussion of specific functional choices is beyond the scope of this work, and we therefore refer the reader to Refs. [39, 40]. Instead, we briefly mention three main types of approximations commonly used among the DFT community. The simplest of these approximations is the *local density approximation* (LDA). The assumption behind this approximation is that the charge density of the system, which is not homogeneous overall, is locally similar to the one of the homogeneous electron gas, whose exchange-correlation energy is known [37]. An improvement upon the LDA can be obtained by semi-local Generalized Gradient Approximation (GGA) functionals [41, 42]. These depend not just on the value of the density at a point (as in the LDA case), but also on its gradient. The last popular type of approximation is called hybrid functionals [43, 44]. Hybrid functionals are based on the ansatz that the exact exchange energy is situated between the GGA exchange energy functional and the Hartree–Fock exchange integral. In these, the Hartree–Fock exchange integral is mixed with GGA exchange functionals at a constant ratio.

Second, both  $v_{\text{H}}[n](\mathbf{r})$  and  $v_{\text{xc}}[n](\mathbf{r})$ , needed to determine the  $\phi_i$ , depend on the density, which is in turn determined from the  $\phi_i$ , requiring a self-consistent procedure to find the solutions.

### 2.1.4 Computational aspects

Solving the Kohn–Sham equations (2.32) to find the ground-state energy as the minimum of (2.30) in practice for any molecular system of interest needs to be done numerically. This requires in turn the use of several computational techniques, whose specific choices can affect the overall accuracy of the obtained result. This section cannot be exhaustive in this context, but we want to mention the real-space representation of the Kohn–Sham orbitals in more detail, as it is also relevant for the later discussions in the chapter.

Numerical solutions to the Kohn–Sham equations are typically obtained by expanding the effective single-particle wave functions  $\phi_i(\mathbf{r})$  in some finite basis with  $N_b$  basis functions  $\zeta_j(\mathbf{r})$  and expansion coefficients  $c_{ij}$  according to

$$\phi_i(\mathbf{r}) = \sum_{j=1}^{N_b} c_{ji} \zeta_j(\mathbf{r}). \quad (2.35)$$

Inserting this into (2.32), multiplying from the left with  $\zeta_k(\mathbf{r})$ , and subsequent integration yields

$$\sum_{j=1}^{N_b} \underbrace{\int \zeta_k(\mathbf{r}) \hat{H}_{\text{KS}} \zeta_j(\mathbf{r}) d^3 \mathbf{r}}_{F_{kj}} c_{ji} = \sum_{j=1}^{N_b} \underbrace{\int \zeta_k(\mathbf{r}) \zeta_j(\mathbf{r}) d^3 \mathbf{r}}_{S_{kj}} c_{ji} \varepsilon_i^{\text{KS}}, \quad (2.36)$$

with the integrals  $F_{kj}$  and  $S_{kj}$  defining the elements of the so called Fock matrix  $\underline{\mathbf{F}}$  and overlap matrix  $\underline{\mathbf{S}}$ , respectively. The Fock matrix comprises terms from the one-electron contributions to  $\hat{H}_{\text{KS}}$  (kinetic energy and external potential) also referred to as core Hamiltonian  $\underline{\mathbf{h}}_{\text{core}}$ , from the classical Hartree contribution  $\underline{\mathbf{J}}$ , explicit exchange  $\underline{\mathbf{K}}$  in hybrid DFT,

and from the exchange-correlation potential  $\mathbf{V}_{\text{xc}}$ . Note that in general, the basis functions are normalized but not orthogonal to each other, and  $\underline{\mathbf{S}} \neq \underline{\mathbf{1}}$ . With this matrix notation, (2.36) can be written as

$$\underline{\mathbf{F}}\mathbf{c}_i = \varepsilon_i^{\text{KS}} \underline{\mathbf{S}}\mathbf{c}_i \quad (2.37)$$

and the electron density can then be rewritten as

$$n(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 = \sum_{i=1}^N \sum_{j=1}^{N_b} c_{ji} \zeta_j(\mathbf{r}) \sum_{k=1}^{N_b} c_{ki} \zeta_k(\mathbf{r}) = \sum_{j,k=1}^{N_b} D_{jk} \zeta_j(\mathbf{r}) \zeta_k(\mathbf{r}). \quad (2.38)$$

Here, we introduce the density matrix  $\underline{\mathbf{D}}$  as

$$D_{jk} = \sum_{i=1}^N c_{ji} c_{ki}. \quad (2.39)$$

The basis set expansion turns the Kohn–Sham equations to a generalized eigenvalue problem as in (2.37), which can be solved with standard numerical methods. Several aspects need to be considered for the choice of the actual functions  $\zeta(\mathbf{r})$ . Obviously, the closer they are to the eventual  $\phi(\mathbf{r})$ , the fewer functions are needed and  $N_b$  – and with that the dimension of the eigenvalue problem – remains small. At the same time, the expressions for  $\zeta(\mathbf{r})$  should facilitate easy computations of the elements of the Fock and overlap matrices. A third consideration is how to approach the limit of an exact basis. Preferably increasing the number of basis functions  $N_b$  should converge results towards the limit of a complete basis. This is not necessarily true for non-orthogonal basis sets, where  $\underline{\mathbf{S}} \neq \underline{\mathbf{1}}$ . Adding more functions to a non-orthogonal basis can lead to linear dependencies and finally numerical problems.

Most commonly used implementations of DFT (and related methods) for finite systems make use of atom-centered Gaussian Type Orbitals (GTOs). They have the functional form  $\zeta_{lm\gamma}(\mathbf{r}) = N_{l,\gamma} Y_{lm}(\bar{\theta}, \bar{\varphi}) |\mathbf{r} - \mathbf{R}_\alpha|^l \exp(-\gamma |\mathbf{r} - \mathbf{R}_\alpha|^2)$ , where  $\mathbf{R}_\alpha$  is the position of the particular atom,  $Y_{lm}(\bar{\theta}, \bar{\varphi})$  is a spherical harmonic, and  $N_{l,\gamma}$  ensures normalization.

tion. These functions are named as  $s, p, d, f, \dots$  depending on the value of  $l$  ( $0, 1, 2, 3, \dots$ ). All functions with the same  $l$  and  $\gamma$  are collectively referred to as a shell. GTOs are attractive because the elements of the Fock and overlap matrices can be calculated analytically, see e.g. [45], except for the contribution from the exchange-correlation potential, for which a numerical integration procedure is required. However, compared to the real solution of atomic wave functions (or Slater orbitals), whose radial ( $r = |\mathbf{r} - \mathbf{R}_\alpha|$ ) dependence is  $\exp(-\gamma r)$ , GTOs decay faster and cannot reproduce the cusp at  $r = 0$ . Basis functions with the correct radial dependence are also known as Slater-type Orbitals (STOs), but they come with the disadvantage of the difficulty to calculate the matrix elements. To approximate the  $r$ -dependence of STOs, one often has use several GTOs with different values of the decay constant  $\gamma$ , where a suitable mix of fast and slow decaying GTOs ensures better (but not exact) reproduction of the cusp-like feature for  $r \rightarrow 0$  and the slow decay for  $r \rightarrow \infty$ . One can bundle a few of these functions in one *contracted* function

$$\zeta_{lm}(\mathbf{r}) = \sum_i d_i \zeta_{lm\gamma_i}(\mathbf{r}) \quad (2.40)$$

with fixed contraction coefficients  $d_i$ . Eventually, many GTO basis sets can be constructed by the choice of the  $\gamma_i$ ,  $d_i$ , and values of  $l$ . In the literature, one will find a plethora of families of such pre-defined sets, with more or less descriptive names. An example for one widely used basis set is the **def2** family [46]. For carbon, the smallest among this family, **def2-svp**, consists of three contracted  $s$ -shells, two contracted  $p$ -shells, and one uncontracted  $d$ -shell, for a total of  $3 \cdot 1 + 2 \cdot 3 + 1 \cdot 5 = 14$  variable coefficients  $c_{ij}$  per molecular orbital  $i$  in (2.35). The functions of the  $d$ -shell are referred to as polarization functions, as they are not needed for the isolated atom, but add more degrees of freedom to allow the electron cloud to be polarized in a molecular environment. It should be stressed that it is the responsibility of any practitioner performing DFT calculations to carefully check the convergence of their results with the choice of the basis set. The same notion applies to other computational

choices, such as the selection of an initial density guess for the start of the self-consistency procedure, the method to ensure convergence to the self-consistent solution, convergence criteria, numerical grids for the  $v_{xc}$  and  $E_{xc}$  integrations, or the use of efficiency-improving auxiliary basis sets for Hartree and exchange integrals. For a full account of those choices, we refer to reader to Ref. [47].

## 2.2 Many-Body Green's Functions Methods for Electronically Excited States

DFT calculations are great for ground state energy calculations. Whenever there is an incidence of electron promotion to an excited state DFT calculations no longer accurately represent the underlying physics [48]. One common method to study excited state phenomena is time-dependent DFT (TD-DFT). However, the real-time TDDFT is too complicated for calculation purposes, and the linear-response theory only allows for minor perturbations in the system. In 1965, Hedin derived a set of equations based on the many-body perturbation theory that enables us to model addition/removal of electrons within a system [49]. He showed the use of a Green's function formalism to the time-dependent Schrödinger equation to calculate an exact description of single-particle excitations and many other observables described by single-particle operators, like the ground-state density and the total energy. Hedin's equations, however, are computationally very challenging. This is where the  $GW$  approximation plays an important role: it enables us obtain a simpler set of equations to solve. In this section, we will discuss the  $GW$  approximation, its significance, physical meaning and practicality.

In Section 2.2.1, we look at the Green's function approach to formulating the equations needed to model charge transfer, which begins with the addition, propagation and finally removal of an extra particle in the system. We arrive at a set of closed form of equations in Section 2.2.2, and look at the approximations needed to solve them. We also

introduce the self-energy term, and briefly discuss the challenges involved in computing it before looking at a fully analytical approach to calculate it. Finally, in Section 2.2.3 we introduce two-particle excitations, that are treated with the Bethe–Salpeter equation (BSE), connecting the independent two-particle Green’s function and the interacting two-particle Green’s function. Towards the end, we also discuss how the BSE can be cast into a matrix form, revealing Casida-like matrices.

### 2.2.1 Charged Excitations and the Single-Particle Green’s Function

Before we recapitulate the major concept behind the  $GW$  approach, let us emphasize that to keep the notation simple we restrict the discussion to the following case. We consider a closed shell system, with a total spin of 0, allowing us to neglect spin, as we did in the previous sections.

Let us assume a particle was added to the  $N$ -particle system at  $(\mathbf{r}_2, t_2) \equiv (2)$  and allowed to propagate through the system until it was removed at  $(\mathbf{r}_1, t_1) \equiv (1)$ . The Green’s function  $G(12)$  is defined as the quantum-mechanical probability amplitude for this process [50, 51]. Note that we will use shorthand notations  $(\mathbf{r}_1, t_1) \equiv 1$  etc. for visual presentation.

To describe this phenomena for electrons, we need to define annihilation ( $\hat{\psi}$ ) and creation ( $\hat{\psi}^\dagger$ ) operators, which help to remove or create an electron in the system respectively. The final state of the system represented in the many-body wave function  $|\Psi^N(\mathbf{r}_1, t_1)\rangle$  from the initial state  $|\Psi^N(\mathbf{r}_2, t_2)\rangle$ , can be written as

$$|\Psi^N(\mathbf{r}_1, t_1)\rangle = \hat{\psi}(\mathbf{r}_1) \hat{U}(t_1, t_2) \hat{\psi}^\dagger(\mathbf{r}_2) |\Psi^N(\mathbf{r}_2, t_2)\rangle, \quad (2.41)$$

where  $\hat{U}(t_1, t_2)$  is the evolution operator that propagates the  $N + 1$ -particle state in time and is defined as

$$\hat{U}(t_1, t_2) = \exp(-i\hat{H}(t_1 - t_2)). \quad (2.42)$$

The Green's function for an electron to be found at (2) if added at (1) is given by

$$G^e(1, 2) = -i \langle \Psi^N(t_1) | \hat{\psi}(\mathbf{r}_1) \hat{U}(t_1, t_2) \hat{\psi}^\dagger(\mathbf{r}_2) | \Psi^N(t_2) \rangle \Theta(t_1 - t_2), \quad (2.43)$$

where  $\Theta(t_1 - t_2)$  is the Heaviside function, which is 1 (0) for  $t_1 > t_2$  ( $t_1 < t_2$ ).

Let us briefly discuss the Schrödinger and the Heisenberg pictures, in general terms, before we begin to use them for further simplification. The state vector  $|\psi(t)\rangle$  which is a solution to the Schrödinger equation is described by action of the time-evolution operator  $\hat{U}(t, t_0)$  on the initial-state vector  $|\psi(t_0)\rangle$ . The state vector at time  $t$  can be written as

$$|\psi(t)\rangle = \exp(-i\hat{H}(t - t_0)) |\psi(t_0)\rangle. \quad (2.44)$$

The term  $\langle \psi(t)|\psi(t)\rangle$  can be written as

$$\langle \psi(t)|\psi(t)\rangle = \langle \psi(t_0)| \exp(i\hat{H}(t - t_0)) \exp(-i\hat{H}(t - t_0)) |\psi(t_0)\rangle \quad (2.45)$$

because  $\hat{H}$  is Hermitian. Since the Hamiltonian commutes with itself ( $[\hat{H}, \hat{H}] = 0$ ), we can write  $\exp(i\hat{H}(t - t_0))\exp(-i\hat{H}(t - t_0)) = \hat{1}$ , and we have  $\langle \psi(t)|\psi(t)\rangle = \langle \psi(t_0)|\psi(t_0)\rangle$ . Any observable in the Schrödinger picture should be the same in the Heisenberg picture as well. Hence, we can write

$$\langle A \rangle(t) = \langle \psi(t) | \hat{A}_S | \psi(t) \rangle = \langle \psi_H | \hat{A}_H(t) | \psi_H \rangle, \quad (2.46)$$

where  $\hat{A}_S$  and  $\hat{A}_H(t)$  are operators in the Schrödinger picture and Heisenberg picture, respectively. Using (2.44) in (2.46), we see that  $\psi_H = \psi(t_0)$  and  $\hat{A}_H(t) = \exp(i\hat{H}(t - t_0))\hat{A}_S\exp(-i\hat{H}(t - t_0))$ . What we essentially see here, is that the time dependence of the state vectors in the Schrödinger picture can be transferred to the operator in the Heisenberg picture keeping the state vectors independent of time.

Using the knowledge of the Schrödinger and the Heisenberg picture from above, we change the representation to the Heisenberg picture via

$$\begin{aligned} |\Psi_H^N\rangle &= \hat{U}(0, t) |\Psi_S^N(t)\rangle \\ \hat{A}_H(t) &= \hat{U}(0, t) \hat{A}_S \hat{U}(t, 0), \end{aligned} \quad (2.47)$$

which gives us

$$G^e(1, 2) = -i \langle \Psi_H^N | \hat{\psi}(\mathbf{r}_1, t_1) \hat{\psi}^\dagger(\mathbf{r}_2, t_2) | \Psi_H^N \rangle \Theta(t_1 - t_2). \quad (2.48)$$

Similarly, the definition of the Green's function for a hole to propagate from (2) to (1) follows from (2.48):

$$G^h(1, 2) = -i \langle \Psi_H^N | \hat{\psi}^\dagger(\mathbf{r}_1, t_1) \hat{\psi}(\mathbf{r}_2, t_2) | \Psi_H^N \rangle \Theta(t_2 - t_1). \quad (2.49)$$

From the definition of the state in the Heisenberg picture, we know that the Heisenberg state can be written as the initial state of the Schrödinger picture  $\Psi_H = \Psi_0$ . The combination of the electron and hole Green's function gives the one-particle Green's function as  $G(1, 2) = G^e(1, 2) - G^h(2, 1)$ : so,

$$G_1(1, 2) = -i \langle \Psi_0^N | \hat{\mathcal{T}}[\hat{\psi}(1) \hat{\psi}^\dagger(2)] | \Psi_0^N \rangle, \quad (2.50)$$

where  $\hat{\mathcal{T}}$  is the fermionic time-ordering operator, which, with the help of the Heaviside function  $\Theta(t - t')$ , can be written as

$$\hat{\mathcal{T}} [\hat{\psi}(1) \hat{\psi}^\dagger(2)] = \Theta(t_1 - t_2) \hat{\psi}(1) \hat{\psi}^\dagger(2) - \Theta(t_2 - t_1) \hat{\psi}(2) \hat{\psi}^\dagger(1). \quad (2.51)$$

Depending on the time-ordering operator, the Green's function describes the propagation of an electron or a hole. Let us convert  $t_1 - t_2 = \tau$  for simplicity. The state vectors for  $N \pm 1$  particle systems are defined as

$$\begin{aligned} g_i(\mathbf{r}) &= \langle \Psi_i^{N-1} | \hat{\psi}(\mathbf{r}) | \Psi_0^N \rangle, \\ f_i^*(\mathbf{r}) &= \langle \Psi_0^N | \hat{\psi}(\mathbf{r}) | \Psi_i^{N+1} \rangle, \end{aligned} \quad (2.52)$$

and the single particle excitation energies as

$$\begin{aligned}\varepsilon_i^{N-1} &= E_0^N - E_i^{N-1}, \\ \varepsilon_i^{N+1} &= E_i^{N+1} - E_0^N.\end{aligned}\quad (2.53)$$

Upon rewriting the Green's function in the Schrödinger picture with these definitions, and supplementing it with the completeness relation  $\sum_i |\Psi_i^{N\pm 1}\rangle \langle \Psi_i^{N\pm 1}| = \hat{1}$ , we get the Lehmann representation

$$\begin{aligned}G_1(\mathbf{r}_1, \mathbf{r}_2; \tau) &= -i \sum_i^{\text{unocc}} f_i(\mathbf{r}_1) f_i^*(\mathbf{r}_2) \exp[-i\varepsilon_i^{N+1}\tau] \Theta(\tau) \\ &\quad + i \sum_i^{\text{unocc}} g_i(\mathbf{r}_1) g_i^*(\mathbf{r}_2) \exp[-i\varepsilon_i^{N-1}\tau] \Theta(-\tau).\end{aligned}\quad (2.54)$$

Let us assume  $t_1 > t_2$  (so  $\Theta(-\tau) = 0$ ) and Fourier-transform (2.54) to the frequency domain, to obtain

$$G_1(\mathbf{r}_1, \mathbf{r}_2; \omega) = \sum_i^{\text{unocc}} \frac{f_i(\mathbf{r}_1) f_i^*(\mathbf{r}_2)}{\omega - \varepsilon_i^{N+1} - i\eta}, \quad (2.55)$$

where we introduced a small imaginary part  $\eta$  to ensure convergence of the Fourier transform. This representation of the Green's function has its poles at  $\varepsilon_i + i\gamma_i$ , where the real part gives the single-particle excitation energies, and the imaginary part yields the particle lifetimes  $\gamma_i = 2\pi/\tau^{N+1}$ . In case of a closed-shell non-interacting systems, with no restrictions on  $t_1$  and  $t_2$ ,  $g_i(\mathbf{r})$  and  $f_i(\mathbf{r})$  are simply the occupied and unoccupied single-particle wave functions. We shall use this later in the chapter, when we discuss the implementation of the Green's function from a computational perspective.

## 2.2.2 Hedin's Equations and the $GW$ Approximation

In this section, we look to sketch out a closed set of equations which define the many-body Green's function for an interacting set of electrons. We begin with the many-body Hamiltonian as described in (2.5). In

terms of field operators, the Hamiltonian can be written as

$$\begin{aligned}\hat{H}_{\text{el}} = & \int \hat{\psi}^\dagger(\mathbf{r}) \hat{h}(\mathbf{r}) \hat{\psi}(\mathbf{r}) d^3\mathbf{r} \\ & + \frac{1}{2} \iint \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}^\dagger(\mathbf{r}') v_C(\mathbf{r}, \mathbf{r}') \hat{\psi}(\mathbf{r}) \hat{\psi}(\mathbf{r}') d^3\mathbf{r} d^3\mathbf{r}',\end{aligned}\quad (2.56)$$

with  $\hat{h}(\mathbf{r})$  the single-particle Hamiltonian operator

$$\hat{h}(\mathbf{r}) = -\frac{\Delta}{2} + \hat{V}_{\text{ext}}(\mathbf{r}).\quad (2.57)$$

From the equation of motion of the field operators in the Heisenberg picture

$$i \frac{\partial}{\partial t} \hat{\psi}(\mathbf{r}, t) = [\hat{\psi}(\mathbf{r}, t), \hat{H}],\quad (2.58)$$

we can derive an equation of motion for the Green's function  $G_1$ :

$$\begin{aligned}\left[ i \frac{\partial}{\partial t} - \hat{h} \right] G_1(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) + i \int d^3\mathbf{r}_3 v_C(\mathbf{r}_1, \mathbf{r}_3) G_2(\mathbf{r}_3 t_1, \mathbf{r}_1 t_1, \mathbf{r}_2 t_2) \\ = \delta(\mathbf{r}_1 - \mathbf{r}_2) \delta(t_1 - t_2),\end{aligned}\quad (2.59)$$

where  $G_2(\mathbf{r}_3 t_1, \mathbf{r}_2 t_2, \mathbf{r}_2 t_2, \mathbf{r}_1 t_1)$  is the two-particle Green's function. It can be seen that the equation of motion for the 1-particle Green's function contains a 2-particle Green's function and then if we write the equation of motion for 2-particle Green's function we will see that it contains 3-particle Green's function and so on infinitely. To overcome this problem we define the self-energy  $\Sigma(\mathbf{r}_1 t_1, \mathbf{r}_3 t_3)$  which is an effective non-local, non-Hermitian potential. It accounts for all the many-body exchange and correlation terms that are beyond the scope of Hartree contributions. It is given by the Green's function via

$$\begin{aligned}i \int v_{\text{el-el}}(\mathbf{r}_2 - \mathbf{r}_3) G_2(\mathbf{r}_3 t_1, \mathbf{r}_1 t_1, \mathbf{r}_2 t_2) d^3\mathbf{r}_3 \\ = - \int \Sigma(\mathbf{r}_1 t_1, \mathbf{r}_3 t_3) G_1(\mathbf{r}_3 t_3, \mathbf{r}_1 t_1) d^3\mathbf{r}_3 dt_3.\end{aligned}\quad (2.60)$$

This simplifies (2.59) to:

$$\begin{aligned} \left[ i \frac{\partial}{\partial t} - \hat{h} \right] G_1(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2) - \int d^3 \mathbf{r}_3 dt_3 \Sigma(\mathbf{r}_1 t_1, \mathbf{r}_3 t_3) G_1(\mathbf{r}_3 t_3, \mathbf{r}_1 t_1) \\ = \delta(\mathbf{r}_1 - \mathbf{r}_2) \delta(t_1 - t_2). \end{aligned} \quad (2.61)$$

The independent particle Green's function  $G_0$ , derived from KS calculations can be written as shown in 2.2.1 as

$$G_0(\mathbf{r}_1, \mathbf{r}_2, \omega) = \sum_i \frac{\phi_i^{\text{KS}}(\mathbf{r}_1) \phi_i^{\text{KS}}(\mathbf{r}_2)}{\omega - \varepsilon_i^{\text{KS}} \pm i\eta}. \quad (2.62)$$

We will skip further derivation for the self-energy and write down the closed set of coupled equations, known as the Hedin's equations to obtain  $\Sigma$  [52].

$$G_1(1, 2) = G_0(1, 2) + \int G_0(1, 3) \Sigma(3, 4) G_1(4, 2) d(34) \quad (2.63)$$

$$\Sigma(1, 2) = i \int G_1(1, 3) W(1, 4) \Gamma(4, 2, 3) d(34) \quad (2.64)$$

$$\Gamma(1, 2, 3) = \delta(1, 2) \delta(1, 3) + \int \frac{\delta \Sigma(1, 2)}{\delta G_1(4, 5)} G_1(4, 6) G_1(5, 7) \Gamma(6, 7, 3) d(4567) \quad (2.65)$$

$$P(1, 2) = -i \int G_1(1, 3) G_1(4, 1) \Gamma(3, 4, 2) d(34) \quad (2.66)$$

$$W(1, 2) = v_C(1, 2) + \int v_C(1, 3) P(3, 4) W(4, 2) d(34). \quad (2.67)$$

Here  $G_0$  is the non-interacting Green's function,  $\Gamma$  the vertex function,  $P$  the polarizability,  $W$  the screened Coulomb interaction and  $v_C$  the Coulomb interaction. (2.63) is referred to as the Dyson's equation for the Green's function which relates the Green's function  $G_0$  to the interacting particle Green's function  $G_1$ . In matrix notation we can write the Dyson equation as  $G_1 = G_0 + G_0 \Sigma G_1$  which upon rearranging gives the interacting particle Green's function as

$$G_1 = (G_0^{-1} - \Sigma)^{-1}. \quad (2.68)$$

The screened interaction  $W(12)$  describes the effective potential of an interacting electron at (1) at (2) including all polarization effects. This essentially means that when an electron moves in an interacting system it carries a ring of positive charge around it due to repulsion from other electrons. This positive charge screens the bare Coulomb interaction on the electron at (2). Such screened electrons are called quasi-particles (QP). Now we invert (2.67) to find the relation of the screened coulomb interaction to the microscopic polarizability. With the dielectric function

$$\bar{\epsilon}(1, 2) = \delta(1, 2) - \int v_C(1, 3)P(3, 2)d(3) \quad (2.69)$$

the (2.67) can be cast into a convenient form

$$W(1, 2) = \int \bar{\epsilon}^{-1}(1, 3)v_C(3, 2)d(3). \quad (2.70)$$

The density response  $\chi$  is calculated by taking the partial derivative of the perturbation w.r.t. the external potential in the ground state as the dielectric function is a function of the ground state only. Further details can be found in the book chapter this chapter has been derived from. It is related to the inverse of the dielectric constant by

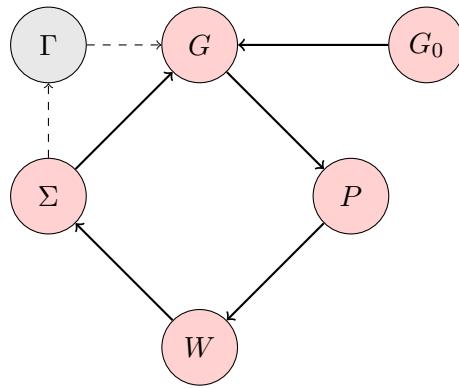
$$\bar{\epsilon}^{-1}(1, 2) = \delta(1, 2) + \int v_C(1, 3)\chi(3, 2)d(3). \quad (2.71)$$

Hence, the screened coulomb interaction  $W$  can be calculated from the dielectric response  $\chi$  via

$$W(1, 2) = v_C(1, 2) + \int v_C(1, 3)\chi(3, 4)v_C(4, 2)d(34). \quad (2.72)$$

Evaluating this system of coupled equations poses severe computational problems, especially the calculation of the derivative with respect to  $G$  in  $\Gamma$  as shown in (2.64). This is when we introduce the  $GW$  formalism which neglects all higher vertex corrections which reduces the vertex correction to

$$\Gamma(1, 2, 3) = \delta(1, 2)\delta(1, 3). \quad (2.73)$$



**Figure 2.1:** Schematic representation of the self-consistent  $GW$  cycle. The vertex correction  $\Gamma$  is approximated to  $\delta(1-2)\delta(2-3)$  and hence it is put in gray. The arrows denote the flow of calculations and  $G_0$  is typically obtained from DFT or HF calculations.

This is the RPA approximation and the self energy is merely a product of the Green's function and the screened Coulomb interaction:  $GW$ . This simplifies the Hedin's equations to a self-consistent set of equations

$$\Sigma(1, 2) = iG_1(1, 2)W(1, 2) \quad (2.74)$$

$$P(1, 2) = -iG_1(1, 2)G_1(2, 1) \quad (2.75)$$

$$\epsilon(1, 2) = \delta(1, 2) - \int v_C(1, 3)P(3, 2)d(3) \quad (2.76)$$

$$W(1, 2) = \int \epsilon^{-1}(1, 3)v_C(3, 2)d(3). \quad (2.77)$$

Having seen the expressions for self-energy using the Green's function let us see how it is used in practical calculations. We know that to calculate Green's function we have to extend frequency ( $\omega$ ) from real to a complex term  $z = \omega + i\gamma$ . The Green's function can be written in the Lehmann representation as

$$G_1(\mathbf{r}, \mathbf{r}', z) = \sum_i \frac{\varphi_i(\mathbf{r}, z)\bar{\varphi}_i(\mathbf{r}', z)}{z - E_i(z)}, \quad (2.78)$$

where  $\varphi_i(\mathbf{r}, z)$  ( $\bar{\varphi}_i(\mathbf{r}', z)$ ) are the right (left) eigenvectors of the non-Hermitian operator  $\hat{H} = \hat{h}_0 + \Sigma(\mathbf{r}, \mathbf{r}', z)$  and  $E_i(z)$  the corresponding eigenvalues,

$$\hat{h}_0\varphi_i(\mathbf{r}, z) + \int \Sigma(\mathbf{r}, \mathbf{r}', z)\varphi_i(\mathbf{r}', z)d^3\mathbf{r}' = E_i(z)\varphi_i(\mathbf{r}, z). \quad (2.79)$$

We have seen the concept of quasiparticles which are electrons shielded by a cloud of positive charge producing an effective screened Coulomb potential  $W$ . These quasiparticles describe low-lying excitations which have energies corresponding to the poles of  $G_1$  in (2.78). We can write down the quasiparticle energy expression from its Hamiltonian in the bra-ket notation as

$$[\hat{h}_0 + \Sigma(\varepsilon_i^{QP})] |\phi_i^{QP}\rangle = \varepsilon_i^{QP} |\phi_i^{QP}\rangle, \quad (2.80)$$

where  $\phi_i^{QP} = \varphi_i(\mathbf{r}, z = \varepsilon_i^{QP})$ . Now that we have all the theoretical framework to proceed with the  $GW$  approximation let us see how we use this set of equations for practical calculations. We start with KS orbitals and the initial non-interacting Green's function  $G_0$  is obtained from (2.62). Then we calculate the self-energy  $\Sigma$  using (2.74)-(2.77). The Hamiltonian  $\hat{h}_0$  can be written as  $\hat{H}_{KS} - \hat{V}_{xc}$  which makes the quasiparticle Hamiltonian as  $\hat{H}_{KS} - \hat{V}_{xc} + \Sigma(\varepsilon_i)$ . The QP eigenvalues and eigenvectors are calculated from

$$[\hat{H}_{KS} + \Sigma(\varepsilon_i^{QP}) - \hat{V}_{xc}] |\phi_i^{QP}\rangle = \varepsilon_i^{QP} |\phi_i^{QP}\rangle. \quad (2.81)$$

One can expand the  $\phi_i^{QP}(\mathbf{r})$  in the basis of the Kohn–Sham orbitals. i.e., using

$$\phi_i^{QP}(\mathbf{r}) = \sum_j c_{ji}^{QP} \phi_j^{KS}(\mathbf{r}). \quad (2.82)$$

With the expansion, the quasiparticle equations (2.81) turn into an eigenvalue problem in matrix form:

$$\underline{\mathbf{H}}^{\text{QP}} \mathbf{c}_i^{\text{QP}} = \varepsilon_i^{\text{QP}} \mathbf{c}_i^{\text{QP}}. \quad (2.83)$$

Here, the elements of the matrix  $\underline{\mathbf{H}}^{\text{QP}}$  are given as

$$\begin{aligned} H_{kj}^{\text{QP}} &= \langle \phi_k^{\text{KS}} | \left[ \hat{H}_{\text{KS}} + \hat{\Sigma}(\varepsilon_i^{\text{QP}}) - \hat{V}_{\text{xc}} \right] | \phi_j^{\text{KS}} \rangle \\ &= \varepsilon_j^{\text{KS}} \delta_{kj} + \langle \phi_k^{\text{KS}} | \hat{\Sigma}(\varepsilon_i^{\text{QP}}) - \hat{V}_{\text{xc}} | \phi_j^{\text{KS}} \rangle \end{aligned} \quad (2.84)$$

and depend via  $\hat{\Sigma}(\varepsilon_i^{\text{QP}})$  on the state of interest, which makes finding the solutions of (2.83) cumbersome.

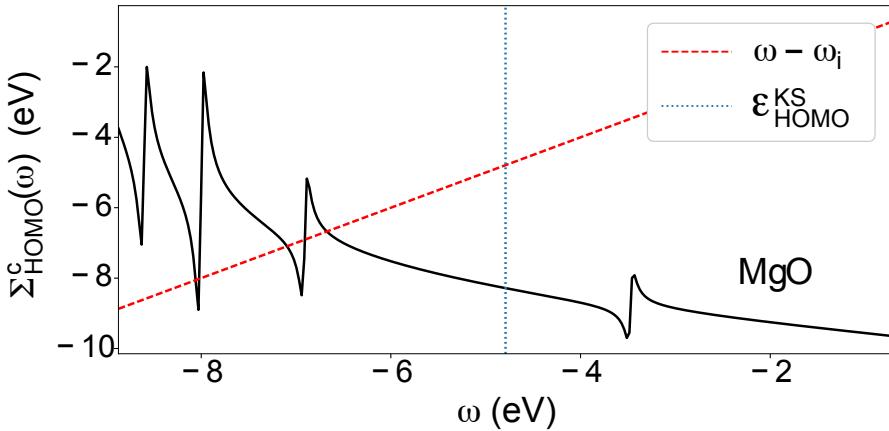
Instead, often a perturbative approach is adopted, in which it is assumed that the off-diagonal elements of (2.84) are negligibly small and quasiparticle wave functions are identical to the Kohn–Sham wave functions,  $\phi^{\text{KS}}(\mathbf{r}) = \phi^{\text{QP}}(\mathbf{r})$ . Then, only the quasi-particle corrections to the  $\varepsilon_i^{\text{KS}}$  are calculated according to

$$\varepsilon_i^{\text{QP}} = \varepsilon_i^{\text{KS}} + \langle \phi_i^{\text{KS}} | \hat{\Sigma}(\varepsilon_i^{\text{QP}}) - \hat{V}_{\text{xc}} | \phi_i^{\text{KS}} \rangle. \quad (2.85)$$

As the self-energy is energy-dependent, and thus depends on  $\varepsilon_i^{\text{QP}}$ , the solution of (2.85) must be found self-consistently. From (2.76) and (2.77) in the  $GW$  approach, it is possible to split the self-energy  $\Sigma = iG\mathcal{W}$  into its bare exchange part  $\Sigma^x = iGv_C$  and its correlation part  $\Sigma^c = iG\widetilde{W}$ , where  $\widetilde{W} = W - v_C$ , which explicitly depends on the screened electron-electron interaction. With  $\omega_i = \varepsilon_i^{\text{KS}} + \langle \phi_i^{\text{KS}} | \Sigma^x - V_{\text{xc}} | \phi_i^{\text{KS}} \rangle$  and  $\langle \phi_i^{\text{KS}} | \Sigma^c(\omega) | \phi_i^{\text{KS}} \rangle = \Sigma_i^c(\omega)$ , we can rewrite (2.85) into the following fixed-point problem

$$\omega - \omega_i = \Sigma_i^c(\omega). \quad (2.86)$$

Figure 2.2 illustrates this fixed-point problem for the highest-occupied molecular orbital (HOMO) of MgO. In the initial step,  $\Sigma^c(\omega)$  is evaluated at  $\omega = \varepsilon_{\text{HOMO}}^{\text{KS}}$ , from which a first  $\varepsilon_{\text{HOMO}}^{\text{QP}}$  is determined, its value used in



**Figure 2.2:** Frequency dependence of the correlation part of the self-energy of the HOMO,  $\Sigma^c_{\text{HOMO}}(\omega) = \langle \phi_{\text{HOMO}}^{\text{KS}} | \hat{\Sigma}^c(\omega) | \phi_{\text{HOMO}}^{\text{KS}} \rangle$ , of the MgO molecule.

the next evaluation of  $\Sigma^c(\omega)$  until finally convergence is reached. Due to the pole structure of the self-energy there are in general several solutions to Eq. (2.86). In this situation, the spectral weight, defined as

$$Z(\omega) = \left( 1 - \frac{d\Sigma^c(\omega)}{d\omega} \right)^{-1}, \quad (2.87)$$

is used to identify the "true" QP energy by  $Z(\omega) \approx 1$ , or  $|d\Sigma^c(\omega)/d\omega| \approx 0$ .

In the above, when evaluating the self-energy in the *GW* approximation, the self-energy takes the form

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int G_1(\mathbf{r}, \mathbf{r}', \omega + \omega') W(\mathbf{r}, \mathbf{r}', \omega) d\omega', \quad (2.88)$$

i.e., it is a convolution of  $G_1$  with the screened Coulomb interaction  $W = \bar{\epsilon}^{-1} v_C$ , where  $v_C(\mathbf{r}, \mathbf{r}') = |\mathbf{r} - \mathbf{r}'|^{-1}$  is the bare Coulomb interaction and  $\bar{\epsilon}^{-1}(\mathbf{r}, \mathbf{r}', \omega)$  is the inverse dielectric function calculated in the Random-Phase Approximation (RPA) [53, 54]. Single-particle wave functions and energies are needed to determine both  $G$  and  $W$ . Constructing both based on the KS eigenvalues and eigenfunctions is known as a

”one-shot”  $G_0W_0$  calculation. Alternatively, it is possible to use updated QP energies in  $G_1$  and  $W$  until eigenvalue self-consistency is reached (ev $GW$ ) [52, 55, 56].

Performing the frequency integration in (2.88) for  $GW$  calculations is numerically challenging. Both the Green’s function and the screened Coulomb interaction inside the integrand have poles infinitesimally above and below the real frequency axis. This makes a standard numerical integration infeasible as the integrand is ill-behaved close to the pole regions, and alternative methods are needed.

In the contour deformation (CD) approach [57, 58], the real-frequency integration is carried out by using the contour integral extending the integrand to the complex plane. In this way, the numerically unstable integration along the real-frequency axis, where the poles of  $G_1$  and  $W$  are located, is avoided. Integration over the full-frequency range can also be performed using Analytic Continuation (AC). This technique makes use of the smoothness of the integral of the self-energy along the imaginary frequency axis. Along this axis, the integral is thus easier to evaluate than along the real-frequency axis. In practice, first the self-energy is determined for a set of imaginary frequencies  $i\omega$  and then continued to the real-frequency axis by fitting the matrix elements  $\Sigma(i\omega)$  to a multipole model (either the so-called 2-pole-model [59, 60] or the Padé approximant [61, 62]).

In what is also known as a fully-analytical approach (FAA) [63, 64] can be used to perform the integration in (2.88). In fact the screened Coulomb interaction can be written as a Dyson-like equation  $W(\omega) = v_C + v_C \cdot \chi(\omega) \cdot v_C$  with  $\chi(\omega)$  the frequency-dependent susceptibility. The poles of the latter can be found solving a Casida-like eigenvalue problem

$$\begin{pmatrix} \underline{\mathbf{H}}^{\text{res}} & \underline{\mathbf{K}} \\ -\underline{\mathbf{K}} & -\underline{\mathbf{H}}^{\text{res}} \end{pmatrix} \begin{pmatrix} \mathbf{A}_n \\ \mathbf{B}_n \end{pmatrix} = \Omega_n \begin{pmatrix} \mathbf{A}_n \\ \mathbf{B}_n \end{pmatrix}. \quad (2.89)$$

with the elements of the exchange-correlation kernel  $K_{\text{xc}}$  omitted. The self-energy integral can then be solved analytically and a closed expression

**Table 2.1:** Comparison of calculated HOMO energies (in eV) based on Kohn–Sham DFT (KS),  $G_0W_0$  and ev $GW$  with the PPM and FAA methods, respectively, based on PBE and PBE0 ground state calculations, and the def2-tzvp basis set. Experimental reference value is -9.2 eV [66].

	PPM			FAA	
	KS	$G_0W_0$	ev $GW$	$G_0W_0$	ev $GW$
PBE	-6.31	-8.74	-9.10	-8.85	-9.24
PBE0	-7.30	-8.92	-9.10	-9.06	-9.24

for  $\Sigma(\omega)$  is obtained. This method is computationally expensive because solving the eigenvalue problem to obtain the poles of  $\chi$  scales as  $\mathcal{O}(N^6)$ .

Instead of computationally demanding full-frequency approaches like the above, the frequency dependence of the dielectric function  $\bar{\epsilon}^{-1}$  and thus the screened Coulomb interaction can be approximated by a plasmon pole model (PPM) [53, 65]. This approximation is motivated by the observation that  $\bar{\epsilon}^{-1}$  is usually dominated by a pole at the plasma frequency  $\omega_p$ , which corresponds to a collective charge-neutral excitation (a plasmon) in the material. The PPM is computationally the most efficient method, because the dielectric function  $\epsilon$  used to compute  $W$  has to be calculated only at a few frequency points to determine the parameters of the PPM.

We close this section on single-particle excitations by illustrating the different discussed methods for the self-energy evaluation in  $GW$  using the example of the HOMO energy of benzene. In Table 2.1, we list the results of plain Kohn–Sham DFT calculations, and combinations of  $G_0W_0$  and ev $GW$  with PPM and FAA frequency integration for the self-energy. We also compare the use of GGA (PBE) and hybrid (PBE0) exchange-correlation functionals in the DFT step. In experiment, the HOMO energy has been measured as -9.2 eV. On KS level, one can observe the typical overestimation of this value, which is larger for the PBE than for the PBE0 functional. When considering the  $G_0W_0$  results,

it is apparent that the HOMO energy is lowered significantly by around 2.5 eV starting from PBE and by around 1.8 eV starting from the PBE0 functional, respectively. One can also clearly see that the energies based on PBE0 are roughly 0.2 eV lower than the ones based on PBE. In other words, the  $G_0W_0$  results show a slight dependence on the underlying Kohn–Sham calculation for the ground state. When the quasiparticle energies are calculated self-consistently in the ev $GW$  variant, the HOMO energies result about 0.4 eV lower than the  $G_0W_0$  ones and are with a value of -9.24 eV very close to the experimental reference. It is particularly noteworthy that the values are identical for PBE and PBE0 ground state calculations. However, these observations should not be generalized to all cases.

### 2.2.3 Two-Particle Excitations and the Bethe–Salpeter Equation

The single particle Green’s function, as discussed in Section 2.2, proves excellent for modeling single particle excitations and hence we can accurately obtain electron addition or removal energies. However, this approach fails for charge-neutral excitations, two-particle scattering or for any property that involves excitonic effects (electron-hole pair interaction). These excitonic effects are essentially encapsulated in the dielectric function  $\bar{\epsilon}(\omega)$ , which describes the response of the system to small perturbations. However, we have computed  $\bar{\epsilon}$  in the  $GW$  approximation which neglects the vertex function  $\Gamma$ , see (2.74) - (2.77). Reintroducing the vertex function would require the whole set of Hedin’s equation (2.63) - (2.67) be solved. Instead, we can write a two-particle Green’s function and solve for it. This two-particle Green’s function depends on four points in space and time as it propagates one particle from  $\mathbf{r}_3, t_3$  to  $\mathbf{r}_1, t_1$  and another particle from  $\mathbf{r}_4, t_4$  to  $\mathbf{r}_2, t_2$ . For simplicity, we will compact notation once again where  $G(\mathbf{r}_1, t_1) = G(1)$ .

Let us begin by assuming that the two particles whose propagation need to be calculated have no correlation. Hence for independent particles

we can write the two-body Green's function as

$$L_0(12, 34) = G(1, 3)G(2, 4). \quad (2.90)$$

Next step is to move to the interacting particle picture where we calculate the two-particle Green's function  $L(12, 34)$  from the existing knowledge of the one-particle Green's function. Let the system be perturbed by a small external potential  $v_{\text{ext}}$  which is non-local in space and time. The two-particle Green's function  $L(12, 34)$  can be written as the change in the Green's function  $G(1, 3)$  w.r.t the external potential

$$L(12, 34) = \frac{\delta G(1, 3)}{\delta v_{\text{ext}}(2, 4)}. \quad (2.91)$$

We rewrite (2.91) such that instead of having the functional derivative of  $G(1, 3)$  we will have the functional derivative of  $G^{-1}$

$$L(12, 34) = G(1, 5) \frac{\delta G^{-1}(5, 6)}{\delta v_{\text{ext}}(2, 4)} G(6, 3). \quad (2.92)$$

Then we insert the Dyson equation for Green's function (2.68) in the presence of an external perturbing potential as

$$L(12, 34) = G(1, 5) \frac{\delta [G_{\text{non-inter}}^{-1}(5, 6) - v_H(5, 6) - v_{\text{ext}}(5, 6) - \Sigma(5, 6)]}{\delta v_{\text{ext}}(2, 4)} G(6, 3) \quad (2.93)$$

where  $G_{\text{non-inter}}$  is the Green's function for non-interacting electrons. The derivative of  $v_{\text{ext}}(5, 6)$  w.r.t  $v_{\text{ext}}(34)$  only gives  $\delta(2, 5)\delta(4, 6)$ . Hence the equation can be written as

$$L(12, 34) = L_0(12, 34) + L_0(16, 53) \left[ \frac{\delta v_H(5)\delta(5, 6)}{\delta v_{\text{ext}}(2, 4)} - \frac{\delta \Sigma(5, 6)}{\delta v_{\text{ext}}(2, 4)} \right]. \quad (2.94)$$

We now use a chain rule and convert the functional derivatives w.r.t  $v_{\text{ext}}$  to functional derivatives w.r.t  $G$ . The functional derivative of the

Hartree potential w.r.t the Green's function is  $-iv_C$ . Thus we can write

$$L(12,34) = L_0(12,34) + L_0(16,53) \left[ v(5,7)\delta(5,6)\delta(7,8) - \frac{\delta\Sigma(5,6)}{\delta G(7,8)} \right] L(72,84). \quad (2.95)$$

Until now, we have made no approximation and this, in principle, is an exact way to calculate the two-particle Green's function for correlated particles. (2.95) is the Dyson equation for the Bethe–Salpeter equation [67, 68]. At this point we have to calculate the functional derivative of the self-energy  $\Sigma$  w.r.t the Green's function  $G$ . This is not an easy task, and we will make our first approximation here, the  $GW$  approximation, and write the self-energy as a product of the Green's function and the screened coulomb interaction, see (2.74), from which it follows that

$$\frac{\delta(iGW)}{\delta G} = iW + iG \frac{\delta W}{\delta G}. \quad (2.96)$$

Here, it is often assumed that  $\delta W/\delta G = 0$  as the change of the screening due to the excitation is small. With that it follows that

$$L(12,34) = L_0(12,34) + L_0(12,56) \left[ \underbrace{-iv_C(5,7)\delta(5,6)\delta(7,8)}_{K^x(56,78)} + \underbrace{i\delta(5,8)W(5,6)\delta(6,7)}_{K^d(56,78)} \right] L(78,34). \quad (2.97)$$

$K^d$  is normally called the direct interaction and originates from the screened interaction  $W$  between electron and hole and is responsible for the binding in the electron-hole pair.  $K^x$  on the other hand originates from the unscreened interaction  $v_C$  and is responsible for the singlet-triplet splitting. It is normally denoted as the exchange interaction.  $K^x$  and  $K^d$  are collectively called the BSE kernel. In the matrix notation this can be written as:

$$\underline{L}(\omega) = \underline{L}_0(\omega) + \underline{L}_0(\omega) \underline{\mathbf{K}}(\omega) \underline{L}(\omega). \quad (2.98)$$

In the Lehmann representation  $L_0$  can be written as a combination of independent quasi-particle excitations, which in the position and frequency space yields:

$$-iL_0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \omega) = \sum_{v,c} \left[ \frac{\phi_c(\mathbf{r}_1)\phi_v(\mathbf{r}_2)\phi_v^*(\mathbf{r}_3)\phi_c^*(\mathbf{r}_4)}{\omega - (\varepsilon_c - \varepsilon_v) + i\eta} - \frac{\phi_v(\mathbf{r}_1)\phi_c(\mathbf{r}_2)\phi_c^*(\mathbf{r}_3)\phi_v^*(\mathbf{r}_4)}{\omega + (\varepsilon_c - \varepsilon_v) - i\eta} \right] \quad (2.99)$$

where  $c$  runs over all unoccupied electronic states and  $v$  over all occupied electron states. It is very challenging to invert a four-point function for each frequency, hence we will move to the orbital basis. It is noteworthy that  $\varepsilon_c$  and  $\varepsilon_v$  here are the quasiparticle corrected energies instead of the Kohn-Sham energies. Transforming this  $L_0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \omega)$  in transition space can be written as

$$L_0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \omega) = \sum_{\substack{(n_1, n_2) \\ (n_3, n_4)}} \phi_{n_1}(\mathbf{r}_1)\phi_{n_2}^*(\mathbf{r}_2)L_0^{n_1, n_2, n_3, n_4}(\omega)\phi_{n_3}(\mathbf{r}_3)\phi_{n_4}^*(\mathbf{r}_4) \quad (2.100)$$

where

$$-iL_0^{n_1, n_2, n_3, n_4}(\omega) = \sum_{v,c} \frac{\delta_{n_1c}\delta_{n_2v}\delta_{n_3c}\delta_{n_4v}}{\omega - (\varepsilon_{n_2} - \varepsilon_{n_1}) + i\eta} - \frac{\delta_{n_1v}\delta_{n_2c}\delta_{n_3v}\delta_{n_4c}}{\omega + (\varepsilon_{n_1} - \varepsilon_{n_2}) - i\eta} \quad (2.101)$$

So  $L_0$  is non-zero only for the diagonal elements when  $n_1 = n_3$  and  $n_2 = n_4$  else the RHS of (2.100) would vanish. Now we assume only transitions from occupied  $\rightarrow$  unoccupied and unoccupied  $\rightarrow$  occupied carry a weight. This implies that only those diagonal elements in the matrix are non-zero which have a transition from  $v$  to  $c$  or vice-versa. In the transition basis, assuming that the transition happens from  $n_1/n_3 \rightarrow$

$n_2/n_4$ ,  $L_0$  is a matrix written as

$$\underline{\mathbf{L}}_0(\omega) = \begin{array}{l} (n_1 n_2) | (n_3 n_4) \rightarrow \\ \downarrow \\ vv \\ cc \\ vc \\ cv \end{array} \begin{pmatrix} v'v' & c'c' & v'c' & c'v' \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta\varepsilon_{n_2 n_1} - \omega} & 0 \\ 0 & 0 & 0 & \frac{-1}{\Delta\varepsilon_{n_2 n_1} - \omega} \end{pmatrix}, \quad (2.102)$$

where  $\Delta\varepsilon_{n_2 n_1} = \varepsilon_{n_2} - \varepsilon_{n_1}$ . Now we can rewrite  $L_0^{n_1, n_2, n_3, n_4}(\omega)$  using occupation factors  $f_i$  ( $f_v = 1$ ,  $f_c = 0$ ):

$$iL_0^{n_1, n_2, n_3, n_4}(\omega) = \frac{(f_{n_1} - f_{n_2})\delta(n_1, n_3)\delta(n_2, n_4)}{(\varepsilon_{n_2} - \varepsilon_{n_1}) - \omega}. \quad (2.103)$$

This allows us to introduce an occupation matrix  $\underline{\mathbf{F}}$ , with components:

$$F^{n_1, n_2, n_3, n_4} = (f_{n_2} - f_{n_1})\delta(n_1, n_3)\delta(n_2, n_4). \quad (2.104)$$

We now restrict the further derivation of  $L_0$  to the physically meaningful subspace  $(vc)(cv)$  belonging to occupied→unoccupied and unoccupied→occupied transitions so that we look only at the bottom right of the matrix in (2.102). We will redefine  $\underline{\mathbf{L}}_0$  including the occupation numbers such that

$$-i\underline{\mathbf{L}}_0 = -i\tilde{\underline{\mathbf{L}}}_0 \underline{\mathbf{F}} \text{ with } \underline{\mathbf{F}} = \begin{pmatrix} \underbrace{f_2 - f_1}_{=-1} & 0 \\ 0 & \underbrace{f_2 - f_1}_{=1} \end{pmatrix}. \quad (2.105)$$

Then we can solve (2.98) for  $\underline{\mathbf{L}}$ :

$$-i\tilde{\underline{\mathbf{L}}}(\omega) = \left[ -i\underline{\mathbf{L}}_0^{-1}(\omega) - i\underline{\mathbf{F}} \underline{\mathbf{K}}(\omega) \right]^{-1}. \quad (2.106)$$

Note that  $\underline{\mathbf{L}}(\omega)$  here is actually defined as  $\underline{\mathbf{L}} = \tilde{\mathbf{L}}\mathbf{F}$ . As  $\tilde{\mathbf{L}}_0(\omega)$  is diagonal the inverse is readily calculated:

$$-\mathrm{i}\tilde{\mathbf{L}}_0^{-1}(\omega) = \begin{pmatrix} \Delta\varepsilon_{n_2n_1} - \omega & 0 \\ 0 & \Delta\varepsilon_{n_2n_1} - \omega \end{pmatrix} \quad (2.107)$$

Therefore, we can single out the frequency dependence:

$$-\mathrm{i}\tilde{\mathbf{L}}_0^{-1}(\omega) = \begin{pmatrix} \Delta\varepsilon_{n_2n_1} & 0 \\ 0 & \Delta\varepsilon_{n_2n_1} \end{pmatrix} - \omega \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.108)$$

which enables us to rewrite the two particle Green's function  $\underline{\mathbf{L}}$  introducing an effective two-particle Hamiltonian  $H^{\text{BSE}}$ :

$$\begin{aligned} \underline{\mathbf{H}}^{\text{BSE}} &= \begin{pmatrix} \Delta\varepsilon_{n_2n_1} & 0 \\ 0 & \Delta\varepsilon_{n_2n_1} \end{pmatrix} - \mathrm{i}\mathbf{F}\underline{\mathbf{K}}(\omega) \\ &= \begin{pmatrix} \Delta\varepsilon_{n_2n_1} & 0 \\ 0 & \Delta\varepsilon_{n_2n_1} \end{pmatrix} - \mathrm{i} \begin{pmatrix} K_{vc,v'c'} & K_{vc,c'v'} \\ -K_{cv,v'c'} & -K_{cv,c'v'} \end{pmatrix} \end{aligned} \quad (2.109)$$

and  $\underline{\mathbf{L}}(\omega) = (\underline{\mathbf{H}}^{\text{BSE}} - \omega)^{-1}$ .

This Hamiltonian is in general non-Hermitian and is typically labeled:

$$\underline{\mathbf{H}}^{\text{BSE}} = \begin{pmatrix} H^{\text{res}} & K \\ -K & -H^{\text{res}} \end{pmatrix} \quad (2.110)$$

where  $H^{\text{res}}$  is the resonant part of the BSE Hamiltonian that treats transitions between occupied and unoccupied states, whereas  $-H^{\text{res}}$  is the anti-resonant part that treats unoccupied to occupied transitions with negative frequencies:

$$H_{vc,v'c'}^{\text{res}} = \Delta\varepsilon_{vc,v'c'} + K_{vc,v'c'}. \quad (2.111)$$

$\underline{\mathbf{K}}$  is a two-point function which the difference between the coulomb repulsion and the screened Coulomb interaction.

Any Hermitian matrix  $\underline{\mathbf{M}}$  can be written in terms of their eigenvalues  $\varepsilon_i$  and their eigenvectors  $\lambda_i$  as

$$\underline{\mathbf{M}} = \sum_i \varepsilon_i |\lambda_i\rangle \langle \lambda_i| \quad (2.112)$$

$$\underline{\mathbf{M}}^{-1} = \sum_i \frac{1}{\varepsilon_i} |\lambda_i\rangle \langle \lambda_i| \quad (2.113)$$

Hence, effectively  $\underline{\mathbf{L}}$  can be written as  $(\underline{\mathbf{H}}^{\text{BSE}} - \omega)^{-1}$ . We would rather look to solve this instead of trying to solve the Dyson equation (2.106). In terms of the spectral representation for  $\underline{\mathbf{L}}(\omega)$

$$-i\underline{\tilde{\mathbf{L}}}(\omega) = (\underline{\mathbf{H}}^{\text{BSE}} - \omega)^{-1} = \sum_S \frac{|\bar{\zeta}_S\rangle S_{\zeta}^{\bar{\zeta}} \langle \zeta_S|}{\Omega_S - \omega} \quad (2.114)$$

where  $S_{\zeta}^{\bar{\zeta}}$  is the overlap between the left and the right eigenvectors and  $\Omega_S$  and  $\zeta_S$  are the eigenvalues and the eigenvectors to the eigenvalue problem

$$\underline{\mathbf{H}}^{\text{BSE}} |\zeta_S\rangle = \Omega_S |\zeta_S\rangle, \quad (2.115)$$

which, using (2.110), is typically written as

$$\begin{pmatrix} \underline{\mathbf{H}}^{\text{res}} & \underline{\mathbf{K}} \\ -\underline{\mathbf{K}} & -\underline{\mathbf{H}}^{\text{res}} \end{pmatrix} \begin{pmatrix} \mathbf{A}^S \\ \mathbf{B}^S \end{pmatrix} = \Omega_S \begin{pmatrix} \mathbf{A}^S \\ \mathbf{B}^S \end{pmatrix}, \quad (2.116)$$

with:

$$\zeta_S(\mathbf{r}_1, \mathbf{r}_2) = A_{vc}^S \phi_c(\mathbf{r}_1) \phi_v^*(\mathbf{r}_2) + B_{vc}^S \phi_v(\mathbf{r}_1) \phi_c^*(\mathbf{r}_2). \quad (2.117)$$

For many systems the off-diagonal blocks  $K$  in (2.116) are small and can be neglected known as the Tamm–Danoff approximation (TDA). This makes the BSE matrix a Hermitian matrix. Furthermore, we can also ignore backward transitions which leads to:

$$\underline{\mathbf{H}}^{\text{res}} \mathbf{A}^{S,\text{TDA}} = \Omega_S^{\text{TDA}} \mathbf{A}^{S,\text{TDA}} \quad (2.118)$$

and the resulting electron-hole amplitude:

$$\zeta_S^{\text{TDA}}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{vc} A_{vc}^{S,\text{TDA}} \phi_c(\mathbf{r}_1) \phi_v^*(\mathbf{r}_2). \quad (2.119)$$

$H^{\text{res}}$  and  $K$  are given by:

$$H_{vc,v'c'}^{\text{res}}(\omega) = D_{vc,v'c'} + K_{vc,v'c'}^x + K_{vc,v'c'}^d \quad (2.120)$$

$$K_{cv,v'c'}(\omega) = K_{cv,v'c'}^x + K_{cv,v'c'}^d. \quad (2.121)$$

Transforming again from the transition basis to the non-interacting quasi-particle basis the respective matrix elements are:

$$D_{vc,v'c'} = (\varepsilon_c - \varepsilon_v) \delta_{vv'} \delta_{cc'} \quad (2.122)$$

$$K_{vc,v'c'}^x = \int \phi_c^*(\mathbf{r}) \phi_v(\mathbf{r}) v_C(\mathbf{r}, \mathbf{r}') \phi_{c'}(\mathbf{r}') \phi_{v'}^*(\mathbf{r}') d^3\mathbf{r} d^3\mathbf{r}' \quad (2.123)$$

$$K_{vc,v'c'}^d = \int \phi_c^*(\mathbf{r}) \phi_{c'}(\mathbf{r}) \phi_v(\mathbf{r}') \phi_{v'}^*(\mathbf{r}') W(\mathbf{r}, \mathbf{r}', \omega = 0) d^3\mathbf{r} d^3\mathbf{r}'. \quad (2.124)$$

Here, as described earlier, we assume that the dynamic properties of  $W(\omega)$  are negligible and use the static approximation, which reduces the computational cost significantly by not demanding self-consistency.

The spin structure of the BSE solutions depends on the spin-orbit coupling. If the ground state is a spin singlet state and spin-orbit coupling is small compared to the electron-hole coupling, the single particles can be classified as spin-up state or spin-down state depending on how the spins of the electron and hole are oriented. Spin-up state indicates a spin-up for both electrons and holes  $|\uparrow\uparrow\rangle$  and vice-versa. Note that a spin of  $\downarrow$  for a hole corresponds to  $\uparrow$  of an electron, that could be there. This yields  $\langle \uparrow_e \rangle \downarrow_h = 1$  and  $\langle \uparrow_e \rangle \uparrow_h = 0$ . The Hilbert space of the electron-hole pairs has four distinct subspaces:

$$|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle,$$

where the first arrow indicates the spin of the electron and the second arrow the spin of the hole. Most of matrix elements defined by (2.122)-

(2.124) between these subspaces are zero. The BSE Hamiltonian then obtains the form

$$\underline{\mathbf{H}}^{\text{BSE}} = \begin{pmatrix} D + K^x + K^d & 0 & 0 & K^x \\ 0 & D + K^d & 0 & 0 \\ 0 & 0 & D + K^d & 0 \\ K^x & 0 & 0 & D + K^x + K^d \end{pmatrix} \begin{array}{l} |\uparrow\downarrow\rangle \\ |\uparrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \end{array} \quad (2.125)$$

This Hamiltonian can be diagonalized, yielding two distinct solutions corresponding to the spin-triplet class and the spin-singlet class. If we look to solve this Hamiltonian we get three eigenvalues that are equal,  $D + K^d$  and one eigenvalue that is  $D + K^d + 2K^x$  which correspond to the adiabatic triplet states and the singlet state.

$$\underbrace{|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)}_{S=1 \text{ triplets}}, \quad \underbrace{\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)}_{S=0 \text{ singlet}}.$$

What we notice here is that if we have an electron that is spin-up and a hole that is spin-down (effectively two electrons in the spin-up state), we get a contribution from their energy difference, hartree-exchange and also their effective screening. When the electron and hole both are spin-up (one spin-up, one spin-down electron) we do not get an exchange term as stated in the Hartree-Fock theory. This allows to solve the BSE equation separately for both spin types.  $K^x$  vanishes for spin-triplet solutions, whereas it is non-zero and typically repulsive for singlets. The exchange interaction is thus responsible for the singlet-triplet energy splitting which yields that the spin-singlet excitation energies are usually higher than the spin-triplet ones. We therefore get

$$H_{\text{singlet}}^{\text{BSE}} = D + K^d + 2K^x \quad (2.126)$$

$$H_{\text{triplet}}^{\text{BSE}} = D + K^d \quad (2.127)$$

If spin-orbit coupling is large, the matrices no longer decouple and instead the BSE Hamiltonian must be evaluated using the full spin structure, which increases the size of the Hamiltonian by a factor of four.

It is to note that traditionally the *GW*-BSE method has been used within the solid-state community [69], however in the last decade it has gained attention for the treatment of electronically excited states of molecular systems [70–73].

## 2.3 Overview of Classical Molecular Dynamics

As we saw in Section 2.1, the nuclear dynamics was ignored once the Born-Oppenheimer approximation [31, 74] was introduced. In this section we will examine the classical treatment of the nuclei to characterize its motion primarily using Newton’s Law of motion. Hence, we see how Equation 2.10 is treated in the classical regime.

### 2.3.1 Classical Potentials

The basic principles for molecular dynamics (MD) arise from the definition of a conservative force acting on a particle  $i$  given by

$$\mathbf{F}_i = -\nabla_{\bar{\mathbf{R}}_i} U(\bar{\mathbf{R}}) \quad (2.128)$$

where  $U(\bar{\mathbf{R}})$  is a potential that defines the interaction between the particle  $i$  and all other particles in the system and  $\bar{\mathbf{R}} \in \mathbb{R}^{3N}$  is a vector describing the configurations of all  $N$  particles in a 3-dimensional space.  $U(\bar{\mathbf{R}})$  thus describes the potential energy landscape in which the atom  $i$  is located. The gradient is calculated with regard to the position of atom  $i$ . The second order ordinary differential equation (ODE)

$$m_i \ddot{\mathbf{R}}_i = -\nabla_{\bar{\mathbf{R}}_i} U(\bar{\mathbf{R}}) \quad (2.129)$$

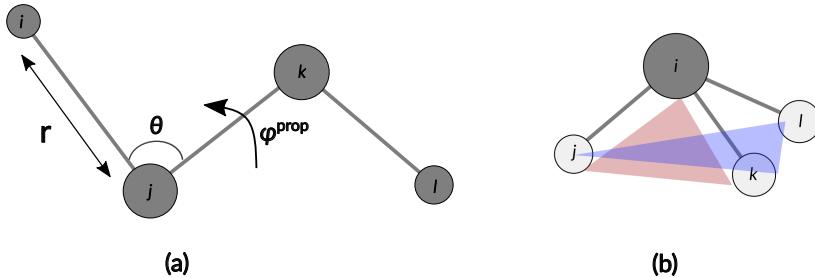
is created by combining the definition of conservative forces with Newton's second law where  $m_i$  is the mass of atom  $i$ . The interactions between the atoms are classically described in the potential energy landscape  $U(\mathbf{R})$ . These result from the Born-Oppenheimer approximation, in which we assume that the degrees of freedom for the electronic and nuclear systems are distinct. Consequently, continuous potentials that simply depend on the distance between atoms can be used to describe interactions between atoms. The potential is also assumed to be completely additive, which means that the total potential energy surface can be effectively described as a sum of individual potential energy terms [75]. Owing to this nature we split the potential into two categories: bonded potentials and non-bonded potentials,

$$U(\overline{\mathbf{R}}) = U^B(\overline{\mathbf{R}}) + U^{NB}(\overline{\mathbf{R}}). \quad (2.130)$$

Covalent bonds are used in the first case to describe the bonded interactions between atoms of the same molecule, while dispersion, repulsion, and/or electrostatic interactions are used in the second case to describe non-bonded interactions between all atoms. *Force Fields* describe the precise definition of the potentials, and their parameters for various atoms and molecules. There are multiple published force fields such as CHARMM [76, 77] AMBER [78, 79], GROMOS [80], and OPLS [81, 82].

Firstly, for the bonded interactions, they are decomposed into three further subcategories arising from (a) the distance between the bonded atoms  $U^b(\overline{\mathbf{R}})$ , (b) the angle formed between three consecutively bonded atoms  $U^\theta(\overline{\mathbf{R}})$ , and (c) the torsion formed between four successive bonded atoms  $U^\phi(\overline{\mathbf{R}})$ . Illustration in Figure 2.3 shows the three subcategories. This decomposition can be written in mathematical terms as

$$U^B(\overline{\mathbf{R}}) = \sum_{bonds} U^b(\overline{\mathbf{R}}) + \sum_{angles} U^\theta(\overline{\mathbf{R}}) + \sum_{dihedrals} U^\phi(\overline{\mathbf{R}}). \quad (2.131)$$



**Figure 2.3:** All bonded degrees of freedom are depicted.  $\mathbf{r}$  is the distance between atoms  $i$  and  $j$  whereas  $\theta$  is the angle formed between atoms  $i, j$  and  $k$ . The proper dihedral  $\phi^{\text{prop}}$  is shown in (a) as the angle between the planes defined by  $i, j, k$  and  $j, k, l$ . The improper dihedral is shown in (b) which is used to keep atoms in a plane especially useful in case of planar rings. It is defined as the angle formed between the planes of atoms  $i, j, k$  and  $j, k, l$ .

Each of these terms can individually be approximated as harmonic potentials and hence we can write their expressions as:

$$\begin{aligned}
 U^b(\bar{\mathbf{R}}) &= \frac{1}{2}k_{ij}^b(r_{ij} - r_{ij}^0)^2 \\
 U^\theta(\bar{\mathbf{R}}) &= \frac{1}{2}k_{ijk}^\theta(\theta_{ijk} - \theta_{ijk}^0)^2 \\
 U^{\phi,\text{imp}}(\bar{\mathbf{R}}) &= \frac{1}{2}k_{ijkl}^{\phi,\text{imp}}(\phi_{ijkl} - \phi_{ijkl}^0)^2 \\
 U^{\phi,\text{prop}}(\bar{\mathbf{R}}) &= k_{ijkl}^{\phi,\text{prop}}[1 + \cos(n\phi_{ijkl} - \delta_{ijkl})]
 \end{aligned} \tag{2.132}$$

We can also see that dihedrals are divided into improper ( $\phi, \text{imp}$ ) and proper dihedrals ( $\phi, \text{prop}$ ) which denote out of plane angles and torsional angles respectively. All these expressions will be revisited in Chapter 3 when we discuss force-field development for the system in use.

Next we consider the non-bonded terms which contain dispersion interactions arising from the transient dipolar interactions between the atoms. The repulsive term arises from repulsive interactions between electrons when they overlap. Atoms can also interact electrostatically to each other when charged. Repulsion and Van der Waals interactions

are combined into a single approximation given by the Lennard–Jones potential.

$$U_{\text{LJ}} = \sum_{i,j>1} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.133)$$

where  $\epsilon_{ij}$  is the depth of the potential well for the two atoms i and j while  $\sigma_{ij}$  is the zero of the LJ potential for i and j and  $r_{ij}$  is the distance between the two atoms. The two constants  $\sigma_{ij}$  and  $\epsilon_{ij}$  are fed in as parameters to the simulation. They are normally fitted to the interatomic interaction of atoms of the same element  $\epsilon_{ii}$  and  $\sigma_{jj}$ . Mixing rules exist [83] that approximate the Lennard-Jones constants for different elements such as the Lorentz-Berthelot [75] rule where  $\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj})$  and  $\epsilon_{ij} = (\epsilon_{ii}\epsilon_{jj})^{\frac{1}{2}}$ .

It is to be understood that Molecular Dynamics is a tool to simulate and understand an experimental system that usually consist of millions to billions of atoms. Given the computational limitations available at disposal simulation of the whole bulk system is impossible. Hence, we intend to be looking at a smaller comparable system, but it would immediately lead to what we know as finite-size effects [84–87]. This poses a challenge to imitate the bulk system using a smaller system but also countering the finite-size effects. The obvious solution is to make periodic images of the smaller finite system to emulate the larger bulk system. Let us begin by putting our system of interest into a finite square box of side  $L$ . Now each particle interacts with other particles within the box and also periodic images of other particles outside the box that are projected in the box. It is to be known that the bonded potentials are not affected but the non-bonded ones are. Given that the MD potentials are additive we can write the total potential of the system as [84]

$$U_{\text{tot}} = \frac{1}{2} \sum_{i,j,\mathbf{n}} U(\mathbf{R} + \mathbf{n}L) \quad (2.134)$$

where  $\mathbf{n} \in \mathbb{N}^3$  is a vector describing the coordinate of each box in units of  $L$ . The value of  $\mathbf{n}$  can range upto  $\infty$ . This leads to infinite size problems in the system most of which can be insignificant. Hence, we resolve this problem in two ways:

- Simple truncation and shift of the Lennard-Jones interaction. The shifted and truncated Lennard-Jones potential acting on particle  $i$  due to particle  $j$  is defined as

$$U_{\text{LJ}}(r_{ij}) = \begin{cases} U_{\text{LJ}}(r_{ij}) - U_{\text{LJ}}(r_{ij}^c), & \text{for } r_{ij} \leq r_{ij}^c \\ 0 & \text{for } r_{ij} > r_{ij}^c, \end{cases} \quad (2.135)$$

where  $r_{ij}^c$  is the threshold truncation range. The truncation range is often set to  $r_{ij}^c = 2.5\sigma_{ij}$  as the Lennard-Jones potential at  $r_{ij} = 2.5\sigma_{ij}$  would have decayed to approximately 1/60th of the well depth  $\epsilon_{ij}$ . Another case is that if the box size is not too large then the threshold is set to half the box size so that periodic images are not counted twice.

- Minimum image convention. Atoms at the edge of the simulation boxes would have to interact with virtual particles due to the PBCs, even if the distance between them is greater than  $r_{ij}^c$ . This is done by correcting the distance between the two particles  $i$  and  $j$  by

$$r_{ij}^x = r_{ij}^x - L \left\lfloor \left( \frac{r_{ij}^x}{L} + 0.5 \right) \right\rfloor \quad (2.136)$$

where  $\lfloor a \rfloor$  is the greatest integer smaller than  $a$ . Here,  $r_{ij}^x$  is the  $x$ -compnent of the vector between the two particles  $\mathbf{r}_{ij}$ .

The other term in the non-bonded potential is the electrostatic interaction between the atomic charges. Similar problem of having infinite boxes arises here as well. However, this time we cannot simply assume a cut-off distance like in case of Lennard-Jones because electrostatics are long-range interactions and does not decay quickly enough to be

truncated. This problem is tackled by using Ewald summation which will not be discussed in detail in this thesis. References for the same can be found in Refs. [88–93].

### 2.3.2 MD trajectories

A collection of potentials that describe the interaction between atoms are approximated using the methods above, and the ODEs described in Eq. (2.128) and Eq. (2.129) are solved as a series of initial value problems. Initial conditions are given by momentum  $\mathbf{p}_i = m_i \frac{dx_i}{dt}$  position  $x_i$  of the atom  $i$ . Solving the ODE yields a topological space (6N-dimensional space) which describes the development of a system with the above possibilities. The complexity of interatomic potentials makes the solution of ODEs analytically impossible. However, numerical methods can be used to approximate the solution. The MD formulation above arose from the conservative definition of force, where the system in question is isolated from the environment and free from external influences. This includes the total energy in the system is conserved. Newtonian equations of motion in this case are time-reversible and requires the use of the time-reversible integrator. This means that we need to sample the space by going over the trajectory in phase space by solving the differential equation. The volume generated for the system should be incompressible thereby making it symplectic.

Details of the theory of symplectic integrators and their relationships to MD are not the subject of this work. More details on these integrators can be found in [94–96] A symplectic integrator which we summarize here is the Verlet integrator defined as:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\mathbf{F}_i(t)}{m_i} \Delta t^2 + \mathcal{O}(\Delta t^4) \quad (2.137)$$

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2) \quad (2.138)$$

where  $\mathbf{r}_i$  is the position,  $\mathbf{v}_i$  is the velocity,  $\Delta t$  is the time step.

### 2.3.3 Simulation ensembles

In the above formulation, the system under consideration is completely isolated from its environment, contains a certain number of particles ( $N$ ), the volume of the box ( $V$ ) is constant and the total energy ( $E$ ) is conserved. Simulation is started using initial positions of atoms which can be randomly oriented within a simulation box or on a regular grid like a crystal. A force-field suitable for the underlying system is chosen and the system is propagated forward in time using a symplectic integrator to generate a trajectory.

We formulate MD as a system guided only by conservative forces thereby conserving energy which leads to the simulation of the microcanonical ensemble [84, 97], also known as NVE simulation. The underlying system is assumed to be an ergodic, therefore time averaging of all measurements is similar to ensemble averaging. It is very difficult to do physical experiments in complete isolation from the area and control/measure the energy of the system. Therefore, we need to develop methods where we can exercise more control for instance temperature and pressure of the system. Hence, the above formulation of the dynamic equations of motion is extensible to create constant volume and temperature (NVT) or constant pressure and temperature (NPT) simulations [84]. Temperature and pressure coupling algorithms are used to create such simulations. Discussion of how they are done, implemented and the derivation of their equations of motion is described in [98–102].

All MD simulation procedures in the thesis, unless stated otherwise, contain energy minimization, NVT and NPT in chronological order.

## 2.4 QM/MM embedding schemes for electronic excitations with *GW*-BSE

Excited states in complex molecular systems are in general multiscale, in the sense that intrinsic quantum-mechanical properties of the basic

units, i.e. isolated molecules, and the local and global morphology of the large-scale molecular system are intertwined. For typical sizes of a system like the polymer-fullerene mixture at the center of this thesis, a full description of the problem by a quantum-mechanical method of the complexity like *GW*-BSE is infeasible and approximations need to be employed instead.

In embedding approaches, the full system is partitioned into two (or more) subsystems, where the one of particular interest is treated at a higher level of theory than the other(s). Examples for such approaches in which the environment of an electronically active region is replaced by a polarizable continuum [103–107], or when it is represented by a classical, molecular mechanics parametrization (QM/MM) [108, 109]. Partitioning in subsystems can also be performed on the level of DFT itself [103, 110–114], and we will focus on aspects of these approaches in more detail in Chapter 5.

In this section, we outline the steps involved in evaluating the effects of the environment on electronic excitation energies in a QM/MM approach combining the *GW*-BSE framework and a polarizable atomistic environment, as implemented in the VOTCA-XTP package.

Assuming a weak coupling between the QM and the MM regions, they affect each other via electrostatic interactions only: the MM region is typically represented by a set of static atomic multipoles, and they form a local field to which the electronic density of the QM region can react, it gets polarized. If one additionally allows for inducible atomic multipole moments in the MM region, this region will also react to the changed density in the QM region. The use of a polarizable distributed atomic multipole representation for molecules in the MM region is an extension of the static charge representation used for MD as introduced in Section 2.3.1. It allows treatment of both the effects of static electric fields and the polarization response as a self-consistent reaction field.

We indicate static atomic multipole moments [115] as  $Q_t^m$ , where  $t$  is the multipole rank and  $m$  the associated atom in the molecule  $M$ . The

tensor  $T_{tu}^{mm'}$  describes the interactions between the multipoles moments  $Q_t^m$  and  $Q_u^{m'}$ . In the polarizable representation, each atom is additionally assigned a polarizability  $\alpha_{tu}^{mm'}$  with which induced moments  $\Delta Q_t^m$  due to the field generated by moments  $u$  on a different atom  $m'$  can be created. Let us consider now a purely classical (MM) system  $\mathcal{S}$  in state  $s$ , which has been split into regions  $\mathcal{R}$  and  $\mathcal{R}'$  with  $\mathcal{S} = \mathcal{R} \cup \mathcal{R}'$ . Molecules in region  $\mathcal{R}$  ( $\mathcal{R}'$ ) are indicated by  $M$  ( $M'$ ), and atoms in molecule  $M$  ( $M'$ ) by  $m$  ( $m'$ ). The total classical electrostatic energy of the system is then given by

$$E_{\text{class}}^{(s)}(\mathcal{S}) = E^{(s)}(\mathcal{R}) + E^{(s)}(\mathcal{R}') + E^{(s)}(\mathcal{R}, \mathcal{R}'), \quad (2.139)$$

where

$$E^{(s)}(\mathcal{R}) = \frac{1}{2} \sum_{M \in \mathcal{R}} \sum_{\substack{M' \in \mathcal{R} \\ M' \neq M}} E_{MM'}^{(s)} + \frac{1}{2} \sum_{M \in \mathcal{R}} E_M^{(s)} \quad (2.140)$$

$$E^{(s)}(\mathcal{R}, \mathcal{R}') = \sum_{M \in \mathcal{R}} \sum_{M' \in \mathcal{R}'} E_{MM'}^{(s)} \quad (2.141)$$

with

$$\begin{aligned} E_{MM'}^{(s)} &= \sum_{m \in M} \sum_{m' \in M'} \sum_{tu} (Q_t^{m(s)} + \Delta Q_t^{m(s)}) \\ &\quad \times T_{tu}^{mm'} (Q_u^{m'(s)} + \Delta Q_u^{m'(s)}) \end{aligned} \quad (2.142)$$

and

$$E_M^{(s)} = \sum_{m \in M} \sum_{\substack{m' \in M \\ m' \neq m}} \sum_{tu} \Delta Q_t^{m(s)} (\alpha_{tu}^{-1})_{(s)}^{mm'} \Delta Q_u^{m'(s)}. \quad (2.143)$$

Eq. (2.139) follows a variational principle with respect to the induced moments, and a preconditioned conjugate gradient method is used to find the  $\Delta Q_t^m$ , which give the minimum energy. Induced interactions are modified using Thole's damping functions [116, 117] to avoid overpolarization.

Now turning towards the QM/MM setup, one can write its total energy as the sum of QM, MM, and QM/MM coupling terms:

$$E_{\text{QM/MM}}^{\text{add}} = E_1^{\text{QM}} + E_2^{\text{MM}} + V_{12}^{\text{QM/MM}}. \quad (2.144)$$

The interaction between the two regions is explicitly evaluated via the term  $V_{12}^{\text{QM/MM}}$ , implying that electrostatic interactions between the two subsystems are handled during the computation of the electronic wave function by adding the external potential of the multipoles in the MM region to the QM Hamiltonian as one-electron operators. In turn, the explicit electrostatic field from the QM density is used to polarize the MM region. For the *GW*-BSE approach several such QM densities can be employed, depending on the state of interest ( $s$ ). If  $s$  is a quasiparticle excitation, we define

$$n_{\text{QP}}^{(s)}(\mathbf{r}) = n_{\text{DFT}}(\mathbf{r}) + f_s |\phi_s^{\text{QP}}(\mathbf{r})|^2, \quad (2.145)$$

with  $f_s = -1$  for occupied and  $f_s = +1$  for unoccupied QPs. If  $s$  is an electron-hole excitation, its total density is evaluated as

$$n^{(s)}(\mathbf{r}) = n_{\text{DFT}}(\mathbf{r}) + n_e^{(s)}(\mathbf{r}) - n_h^{(s)}(\mathbf{r}). \quad (2.146)$$

Here the electron (hole) contribution of the exciton to the density is computed by integrating the squared excited-state wavefunction  $\zeta_S$  with respect to the hole (electron) coordinates, i.e.,

$$\begin{aligned} \rho_e^{(s)}(\mathbf{r}) &= \rho_e^{(s)}(\mathbf{r}_e) = \int d\mathbf{r}_h |\zeta_S(\mathbf{r}_e, \mathbf{r}_h)|^2 \\ \rho_h^{(s)}(\mathbf{r}) &= \rho_h^{(s)}(\mathbf{r}_h) = \int d\mathbf{r}_e |\zeta_S(\mathbf{r}_e, \mathbf{r}_h)|^2. \end{aligned} \quad (2.147)$$

Inclusion of a polarizable MM region requires a self-consistent procedure to evaluate the total QM/MM energy of a system in a given state  $s$ . At a single step  $p$  within this self-consistent procedure, first a QM level calculation (DFT for the ground state  $s = g$ , DFT+*GW*-BSE for electron-hole excited  $s = x$  states) is performed in the electric field

generated by the total moments in the MM region. The resulting QM energy then reads

$$E_{\text{QM}}^{(s),p} = E_{\text{DFT}}^{(s),p} + \delta_{sx} \Omega_s^p \quad (2.148)$$

and the associated total electron density and its electric field at the positions of the atoms in the MM region is evaluated. New atomic dipole moments are induced in the MM region minimizing the classical energy  $E_{\text{class}}^{(s),p}$ . It is used to determine the total energy of the coupled QM/MM system as

$$E_{\text{QM/MM}}^{(s),p} = E_{\text{QM}}^{(s),p} + E_{\text{class}}^{(s),p}. \quad (2.149)$$

The whole procedure is repeated until the change of total energy is less than a preselected accuracy, typically  $10^{-5}$  Ha. From the state-dependence of the total QM/MM energy in Eq. (2.149), it also follows that the excitation energy  $\Omega^{(s)}$  of a complex in the polarizable environment needs to be obtained from a "Δ-QM/MM-SCF" formulation as the difference

$$\Omega^{(s)} = E_{\text{QM/MM}}^{(s)} - E_{\text{QM/MM}}^{(n)}. \quad (2.150)$$

An equivalent formulation for quasiparticle excitation energies uses the respective  $\varepsilon^{\text{QP}}$  instead of  $\Omega_S$  in Eq. (2.148).

# Chapter 3

## All-Atom Force Field for Diketopyrrolopyrrole Polymers with Conjugated Substituents

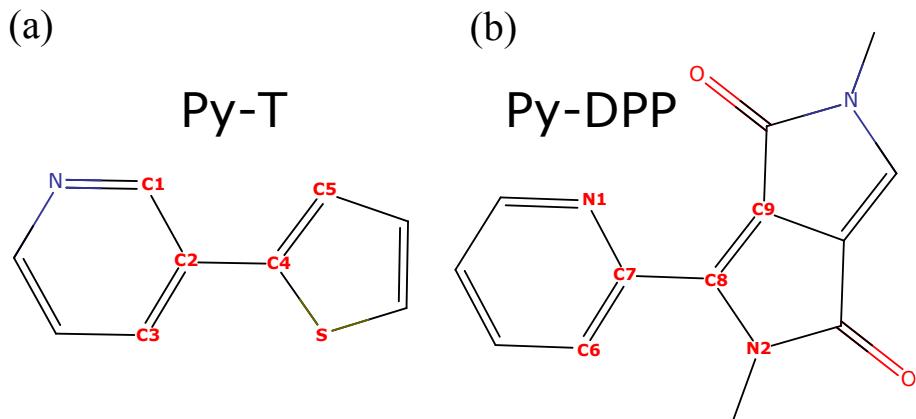
Based on the paper: Vivek Sundaram, Alexey V. Lyulin, Bjoern Baumeier, J. Phys. Chem. B **124**, 11030 (2020).

### 3.1 Introduction

Now that we have described the system undertaken for study and the procedure of the study is intended to advance, we begin with the first step. One must realize that a key element of such approaches is the simulation of representative structures with atomistic detail, which on the needed scale can only be obtained with classical Molecular Dynamics (MD). The resulting morphologies must not only reliably reflect the overall conformational details of the polymer chains in the bulk and at the interface to the acceptor, but also accurately capture the dynamics of torsional degrees of freedom between the segments in the individual chains, as those can massively influence the localization characteristics of electronic states and, hence, their dynamics [118–120]. This, in turn, requires the availability of precise classical force fields which are

parameterized for the specific chemical composition of the materials at hand. In this chapter we discuss how we develop an all-atom force field for a series of diketopyrrolopyrrole polymers with two aromatic pyridine substituents and variable number of  $\pi$ -conjugated thiophene units in the backbone (DPP2PymT) [27]. With the end goal to study this dichotomous role of donor or acceptor excitation, we begin performing classical simulations to study structural and segmental properties of the polymer itself. We develop an all-atom force field for this series by starting out from available *intra*-fragment parameterizations of the individual fragment building blocks. The missing bond, angle, and dihedral bonded parameters for the *inter*-fragment degrees of freedom are derived in an iterative procedure from density-functional theory (DFT) calculations using standard protocols [121–125]. DFT is also used to refine atomic partial charges for the fully assembled DPP2PymT monomers. Guaranteeing *a priori* the transferability of gas-phase optimized force fields to a wide range of conditions is in general a challenging task. The bonded potentials, representing the intra-polymer interactions and – unless the condensed phase dramatically changes these in the polymer by strongly affecting its electronic structure – we expect them to transfer to the condensed phase. Certain torsional motion may be constrained by inter-polymer interactions, which are typically non-bonded interactions, or excluded volume effects. To scrutinize *a posteriori* the applicability of our force field in different conditions, we perform classical MD simulations of a single  $N_p$ -DPP2PymT chain with  $N_p = 50$ ,  $m = 1, 2, 3$ , and branched 2-hexyldecyl side chains both in solutions with chloroform and water, and a few of such chains as melts. We investigate in particular the swelling behavior depending on solvent quality, the glass-transition temperature, and the mobility of the conjugated segments. Available intra-fragment parameterizations of the individual fragment building blocks are combined with inter-fragment bonded and non-bonded parameters explicitly derived from density-functional theory calculations. To validate the force field we perform classical molecular dynamics simulations of single polymer chains with  $m = 1, 2, 3$  in good and bad solvents, and of melts.

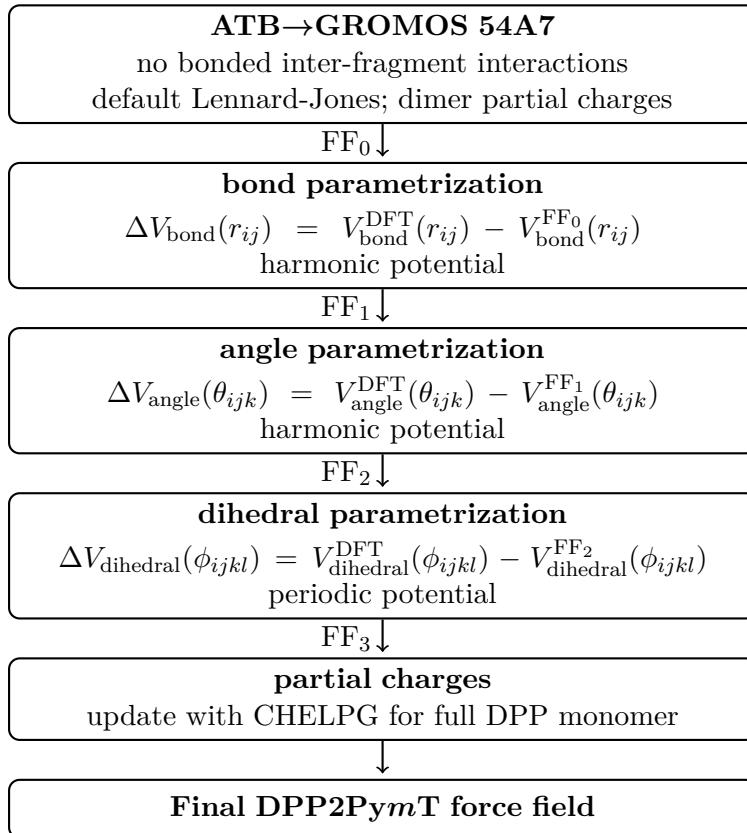
### 3.2 Methodology



**Figure 3.1:** Structure of the (a) Py-T and (b) Py-DPP dimers, respectively, with explicit labels for atoms involved in the inter-fragment degrees-of-freedom for which force field parameters need to be determined.

A force field represents the quantum-mechanical many-body energy of a molecular conformation, defined by the set of atomic coordinates  $\{\mathbf{R}\}$ , as a sum of classical interactions. The potential energy surface (PES) is then written in terms of bonded and non-bonded interactions  $V(\{\mathbf{R}\}) = V_{\text{bonded}}\{\mathbf{R}\} + V_{\text{non-bonded}}\{\mathbf{R}\}$ . The non-bonded ones (Lennard-Jones and electrostatic) are pairwise additive, i.e.,  $V_{\text{non-bonded}}\{\mathbf{R}\} = \sum_{i < j} V_{ij}(\mathbf{R}_{ij})$ , while the bonded ones comprise 2- (bonds  $r_{ij}$ ), 3- (angles  $\theta_{ijk}$ ), and 4-body (dihedral angles  $\phi_{ijkl}$ ) interactions.

For developing the DPP2PymT force field, we start from available parameters in GROMOS 54A7 form for the fragment building blocks of pyridine (Py), thiophene (T) and diketopyrrolopyrrole (DPP) [126], see Figure 3.1. These were obtained from the The Automated force field Topology Builder [127], which extracts bond and angle parameters from the Hessian of the optimised structure and fits partial charges through the Kollmann-Singh scheme [128]. Lennard-Jones parameters are taken from the GROMOS 96 force-field [129]. These *intra-fragment* parametrizations



**Figure 3.2:** Illustration of the development of a  $N_p$ -DPP2PymT force field based using a step-by-step procedure to add inter-monomer potentials to a baseline GROMOS 54A7 monomer parametrization.

need to be supplemented by *inter-fragment* interactions related to the linking of the units in the polymer, i.e.,  $V_{\text{bonded}} = V_{\text{bonded}}^{\text{intra}} + V_{\text{bonded}}^{\text{inter}}$ . To obtain

$$\begin{aligned}
 V_{\text{bonded}}^{\text{inter}} &= \sum_{ij}^{\text{inter}} V_{ij}^{\text{bond}}(r_{ij}) + \sum_{ijk}^{\text{inter}} V_{ijk}^{\text{angle}}(\theta_{ijk}) \\
 &\quad + \sum_{ijkl}^{\text{inter}} V_{ijkl}^{\text{dihedral}}(\phi_{ijkl})
 \end{aligned} \tag{3.1}$$

we follow a step-by-step refinement procedure based on Py-T and DPP-Py dimer structures (see Figure 3.1 for atom definitions), respectively, as illustrated in Figure 3.2. From the available fragment parameterizations, an initial force field ( $\text{FF}_0$ ) is created in which all inter-monomer bonded potentials are zero, dimer CHELPG partial charges [130] are assigned from a density-functional theory (DFT) calculation, and Lennard-Jones parameters are taken from the existing GROMOS force-field with geometric averaging for atom pairs. Starting from  $\text{FF}_0$ , first harmonic bond potentials are added ( $\text{FF}_1$ ), then harmonic angle potentials ( $\text{FF}_2$ ), and finally periodic dihedral potentials ( $\text{FF}_3$ ). To determine these additions, the relevant part of the PES must be known for each individual coordinate at each step of the process. A series of calculations must be performed in which one of the structural parameters ( $x = \{r_{ij}; \theta_{ijk}; \phi_{ijkl}\}$ ) is constrained to a certain value, while all other degrees-of-freedom are optimized. Two of such relaxed PES scans are required: one with the current force field ( $\text{FF}_n$ ), yielding  $V^{\text{FF}_n}(x)$ , and one based on DFT calculations, yielding a reference  $V^{\text{DFT}}(x)$ . Then, the difference  $\Delta V(x) = V^{\text{DFT}}(x) - V^{\text{FF}_n}(x)$  is calculated and used to fit the functional form of the respective potential term. With its parameters, e.g., the equilibrium value and force constant for a bond, determined, the terms are added to the force field. Details will be discussed below in the specific sections.

All DFT calculations used in this work have been performed with the ORCA software [131], employing the B3LYP hybrid functional [132] and the def2-TZVP basis set [133]. The VOTCA-XTP package [134, 135] has been used for the assignment of atomic partial charges with the CHELPG [130] methods. GROMACS 2018.3 [136] with double precision has been used for all calculations with the classical force field. In the PES scans the energy has been minimized until the force on each is below 0.1 kJ/mol/nm.

### 3.3 $N_p$ -DPP2PymT force field

#### 3.3.1 Bond stretching

As mentioned in Section 3.2, we first consider the two bond stretching potentials between Py-T and Py-DPP, and the results of the respective PES scan on DFT/B3LYP and force field levels are shown in Figure 3.3. As expected, the scan with the initial force field (gray open circles) does not exhibit any minimum in both cases, while its faint linear behavior can be attributed to non-bonded interactions as the distance between the two fragments is varied. In contrast, the DFT scans (red squares) clearly show harmonic behavior with an equilibrium bond length of  $r_{ij}^0 = 1.46 \text{ \AA}$  for the bonds highlighted in the respective insets.

The difference  $\Delta V_{\text{bond}}(r_{ij}) = V_{\text{bond}}^{\text{DFT}}(r_{ij}) - V_{\text{bond}}^{\text{FF}_0}(r_{ij})$  between the energy values obtained by B3LYP and  $\text{FF}_0$  was modelled as a harmonic potential

$$V_{\text{bond}}(r_{ij}) = \frac{1}{2}k_{ij}(r_{ij} - r_{ij}^0)^2 \quad (3.2)$$

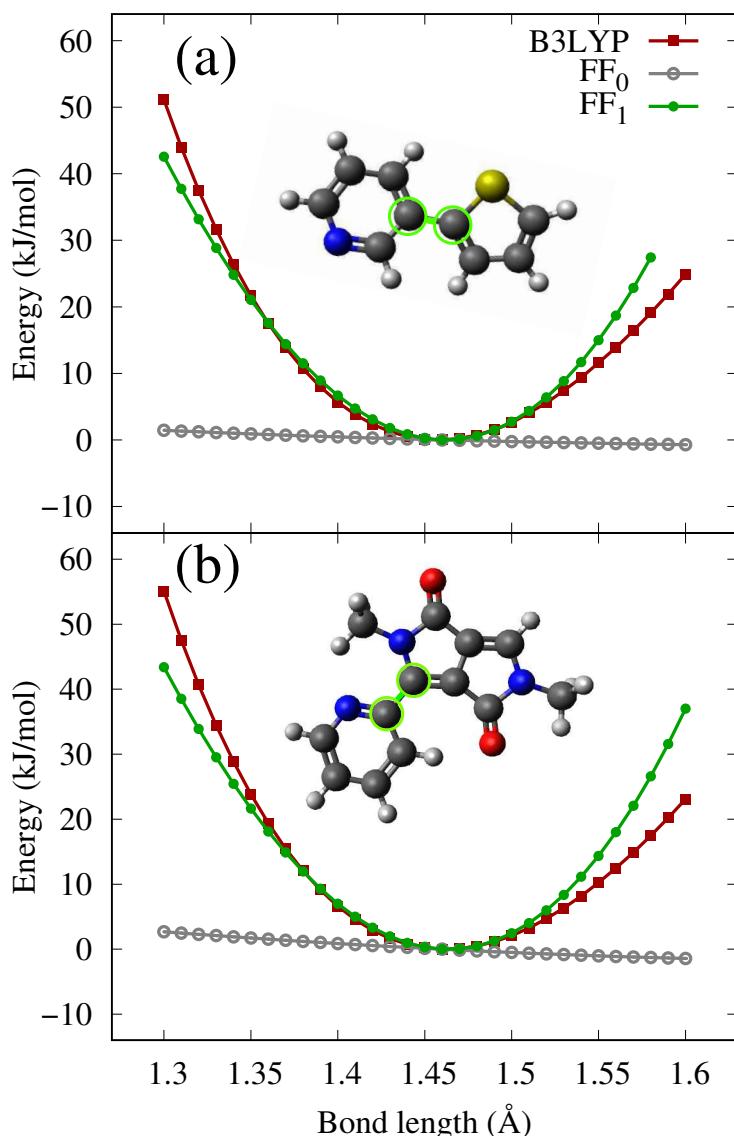
where  $k_{ij}$  the harmonic force constant. However, GROMOS-96 uses a fourth power potential

$$V_{\text{bond}}(r_{ij}) = \frac{1}{4}\tilde{k}_{ij}(r_{ij}^2 - (r_{ij}^0)^2)^2 \quad (3.3)$$

so the determined force constant is converted according to

$$\tilde{k}_{ij} = \frac{k_{ij}}{2(r_{ij}^0)^2}. \quad (3.4)$$

After fitting, the determined parameters as given in Table 3.1 have been added to the force field, defining  $\text{FF}_1$ . To validate this step, the PES scans are repeated with the updated force field yielding results shown as green filled circles in Figure 3.3. For both Py-T and Py-DPP, the agreement with the DFT/B3LYP reference is very good within  $\pm 0.05 \text{ \AA}$  of the equilibrium bond length. The observed deviations for larger



**Figure 3.3:** PES scans (in kJ/mol) for the (a) C2-C4 bond in Py-T and (b) C7-C8 bond in Py-DPP, resulting from DFT/B3LYP (red squares), the initial FF<sub>0</sub> parametrization (grey open circles), and the updated FF<sub>1</sub> (green filled circles) force fields, respectively. The unknown bonds and the atoms involved are highlighted in the insets.

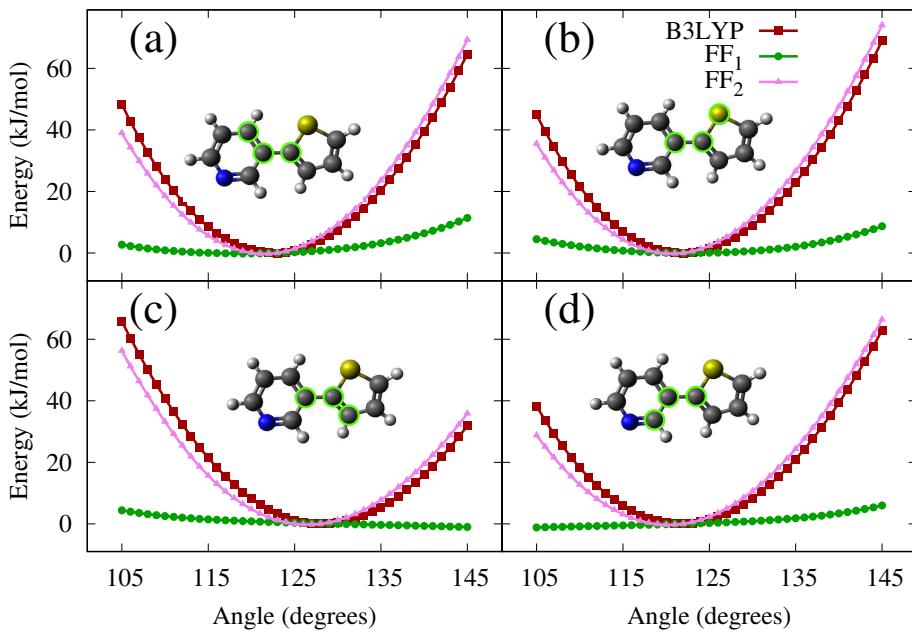
	$i-j$	$\tilde{k}_{ij}$ (kJ/mol/nm <sup>4</sup> )	$b_{ij}$ (nm)
Py-T	C2-C4	$8.44 \cdot 10^6$	0.146
Py-DPP	C7-C8	$3.80 \cdot 10^6$	0.146

**Table 3.1:** Determined harmonic parameters for the inter-fragment bond stretching potentials as used in Eq. (3.3)

variations from the equilibrium can be attributed to anharmonicities that can by construction not be captured by the classical force field. We emphasize that such deviations are associated with energies of about 10 kJ/mol or higher, which corresponds to an equivalent temperature of 1200 K. In practice, MD simulations will be performed at much lower temperatures, and the deviations between FF<sub>1</sub> and the reference can be considered insignificant in this case.

	$i-j-k$	$\tilde{\kappa}_{ijk}$ (kJ/mol)	$\theta_{ijk}^0$ (°)
Py-T	C1-C2-C4	912.28	123
	C2-C4-S	807.07	122
	C2-C4-C5	802.55	128
	C3-C2-C4	896.77	122
Py-DPP	N1-C7-C8	2047.80	124
	C6-C7-C8	1860.80	127
	C7-C8-N2	1658.20	118
	C7-C8-C9	1623.20	120

**Table 3.2:** Determined harmonic parameters for the inter-fragment angle bending potentials.



**Figure 3.4:** PES scans (in kJ/mol) for the angles (a) C1-C2-C4, (b) C2-C4-S, (c) C2-C4-C5, and (d) C3-C2-C4 in the Py-T structure, resulting from DFT/B3LYP (red squares), the FF<sub>1</sub> parametrization (green filled circles), and the updated FF<sub>2</sub> (pink filled triangles) force fields, respectively. The atoms defining the respective angles are also highlighted in the insets.

### 3.3.2 Angle bending

With the potential for the Py-T and Py-DPP bonds included in force field, we now turn to the addition of the angle bending potentials. Figure 3.4 shows the results of PES scans of the four angles in the Py-T structure based on the DFT/B3LYP reference and the FF<sub>1</sub> parametrization. The respective figure for the Py-DPP structure is not shown in this thesis. However, the equilibrium angle and bending potential values are presented in Table 3.2. In contrast to the case of the bond potential, the harmonic force constants for the angle bending cannot be obtained from a simple fit of  $\Delta V_{\text{angle}}(\theta_{ijk}) = V_{\text{angle}}^{\text{DFT}}(\theta_{ijk}) - V_{\text{angle}}^{\text{FF}_1}(\theta_{ijk})$  since the four angles are

not independent of each other. Instead, we determine the eight unknown parameters (equilibrium angles  $\theta_{ijk}^0$  and force constants  $\kappa_{ijk}$  for the four angles) simultaneously by minimizing the Frobenius norm of the matrix

$$F_{ijk,\alpha} = \Delta V_{\text{angle}}(\theta_{ijk,\alpha}) - \frac{1}{2}\kappa_{ijk}(\theta_{ijk,\alpha} - \theta_{ijk}^0)^2, \quad (3.5)$$

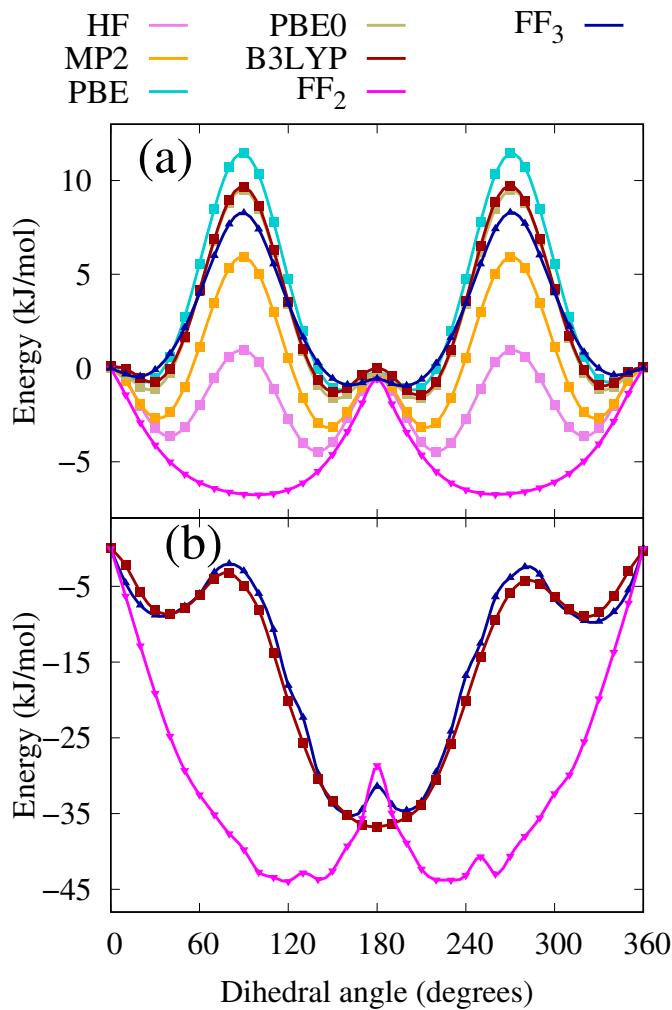
where  $\alpha$  indicates the angle values for which the PES has been scanned. Specifically, we used 41 points in the interval [105°, 145°]. In GROMOS-96, the angle potential is implemented in the form

$$V_{\text{angle}}(\theta_{ijk}) = \frac{1}{2}\tilde{\kappa}_{ijk}(\cos(\theta_{ijk}) - \cos(\theta_{ijk}^0))^2, \quad (3.6)$$

where  $\tilde{\kappa}_{ijk} = \kappa_{ijk}/\sin^2(\theta_{ijk}^0)$ . The obtained parameters in the form are given in Table 3.2. Adding those values and the corresponding equilibrium angles defines FF<sub>2</sub>. The validation scan of the PES shown as pink triangles in Figure 3.4 shows a good agreement with the DFT/B3LYP reference. Similar to the bond potential, significant deviations are only observed that would occur at energies associated to extremely high equivalent temperatures and can therefore be considered insignificant for the intended applications in MD simulations of DPP2PymT.

### 3.3.3 Dihedral angles

The final bonded potentials that remain to be determined are the periodic dihedrals formed by C1-C2-C4-S in Py-T and N1-C7-C8-N2 in Py-DPP. The respective PES scans with DFT/B3LYP (red squares) and FF<sub>2</sub> (pink triangles) for the former are shown in Figure 3.5(a). The *cis* (*trans*) orientation corresponds to 0° (180°), respectively. Without an explicit dihedral potential as in the FF<sub>2</sub> force field, an orthogonal arrangement of the Py and T fragments is preferred. In the DFT reference calculation, in contrast, a nearly planar conformation of Py-T is energetically most favorable, with the *trans* orientation slightly lower in energy by 0.67 kJ/mol. In both cases, there are two out-of-planarity conformations due to steric repulsion of the sulphur atom in the thiophene and hydrogen



**Figure 3.5:** PES scan (in kJ/mol) for the (a) C1-C2-C4-S dihedral formed at the Py-T junction and (b) N1-C7-C8-N2 dihedral formed at the Py-DPP junction, resulting from DFT/B3LYP (red squares), the FF<sub>2</sub> parametrization (pink solid triangles) and FF<sub>3</sub> (blue solid triangles) force fields, respectively. 0° is defined for the *cis* configuration with respect to the nitrogen of Py and sulphur of T while 180° is the *trans* configuration. Likewise, *cis* is defined for the N1 of Py and N2 of DPP being on the same side and *trans* when they are opposite to each other.

atom from the adjacent ring, separated by a barrier of 0.14 kJ/mol [137]. The energy barrier between cis- and trans-conformations, in contrast, is about 10 kJ/mol. For the Py-DPP dihedral, we observe at FF<sub>2</sub> level (pink triangles in Figure 3.5(b)) in non-planar *trans* orientations with minima at 120° and 230°. The energy barrier is around 44 kJ/mol to the planar *cis* and 15 kJ/mol to the planar *trans* configuration, respectively. In the DFT/B3LYP reference, however, there is a clear absolute minimum at 180° configuration, and two local minima at 40° and 320°. For a transition from this twisted *cis* to the planar *trans* arrangement, a barrier of about 5 kJ/mol is found, while it is 32 kJ/mol for the inverse transition.

To obtain the classical potential associated with the dihedral rotation, the difference  $\Delta V_{\text{dihedral}}(\phi_{ijkl}) = V_{\text{dihedral}}^{\text{DFT}}(\phi_{ijkl}) - V_{\text{dihedral}}^{\text{FF}_2}(\phi_{ijkl})$  is represented as

$$V_{\text{dihedral}}(\phi_{ijkl}) = k_{\phi}^{ijkl}[1 + (\cos(n\phi_{ijkl} - \delta_{ijkl}))] \quad (3.7)$$

where  $\phi_{ijkl}$  is the value of the dihedral angle between *ijk* and *klj* planes,  $n$  is the multiplicity which arises due to the fact that multiple points of minima can be obtained with a separation of a small energy barrier between 0° and 360°, and  $\delta_{ijkl}$  is a phase shift. From the shape of both DFT/B3LYP reference curves (Py-T and Py-DPP), we see that they are symmetric about 180° and hence  $n = 2$  and the  $\delta_{ijkl} = 0^\circ$ . Adding the parametrized Eq. (3.7) as in Table 3.3 yields the updated force field FF<sub>3</sub>. A PES scan with this parametrization is shown as blue triangles in Figure 3.5. The overall shapes of the reference PES is reproduced and FF<sub>3</sub> is a massive improvement over FF<sub>2</sub>. Barrier heights, especially the one for the cis-trans transition in Py-T, are slightly underestimated, which is due to the limitations set by the functional form of Eq. (3.7).

	$i-j-k-l$	$k_{\phi}^{ijkl}$ (kJ/mol)	$\delta_{ijkl}$ (°)
Py-T	C1-C2-C4-S	7.5	180
Py-DPP	N1-C7-C8-N2	18.50	180

**Table 3.3:** Determined parameters for the inter-fragment periodic dihedral angle potentials.

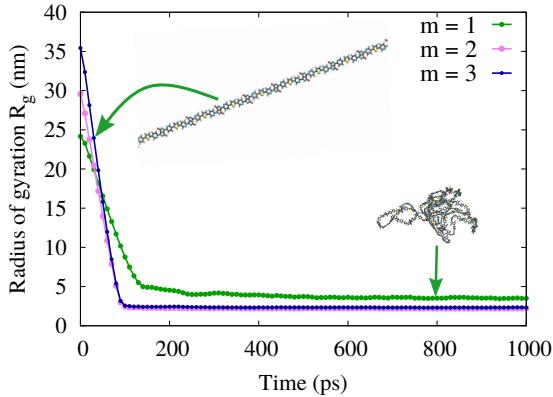
### 3.3.4 Partial charge calculation

For the development of the inter-fragment bonded parameters, partial charges are assigned in Py-T and Py-DPP dimers as obtained with the CHELPG method [130] based on underlying DFT calculations. With all bonded parameters now at hand on the level of FF<sub>3</sub>, we turn to refining the partial charges for the fully assembled molecule shown as donor in Figure 1.2.

Once all the individual fragments are linked to form the DPP2PymT monomer, this step is necessary to account for changes in the charge distribution compared to one observed in the individual dimer structures. Special care needs to be taken regarding the terminating hydrogens of the monomer residue: for a DPP2PymT that itself is located at the termini of the polymer, one hydrogen needs to be excluded from the CHELPG procedure, while any other residue does not contain hydrogens at all at the inter-residue bond. To account for this, we perform three partial charge fits with the charge of either and both of the terminating hydrogens constrained to zero, respectively. This will ensure charge neutrality of the polymer independent of the degree of polymerization. For simulations with addition of side chains, standard partial charge values as described in OPLS-AA were used [138].

Finally, with the definition of all bonded and non-bonded parameters for the DPP2PymT monomers at hand, the complete force field can be assembled for an arbitrary topology, to perform MD simulations, as we will show for  $N_p$ -DPP2PymT polymer in solutions and as melts in the following sections.

### 3.4 $N_p$ -DPP2PymT in solution



**Figure 3.6:** Radius of gyration (in nm) of 50-DPP2PymT in vacuum, starting from all-stretched chain conformations. Each chain collapses into a globule (see inset in the bottom right) within the first 200 ps when simulated in the NVT ensemble at 300 K. Side chains have not been shown for visual clarity.

We verify the quality of the force field using a series of polymer simulations, beginning in this section with a single chain polymer simulation in vacuum, then proceed to solvate the polymer chain in water and in chloroform to cross-check solvent effects as seen in experiments.

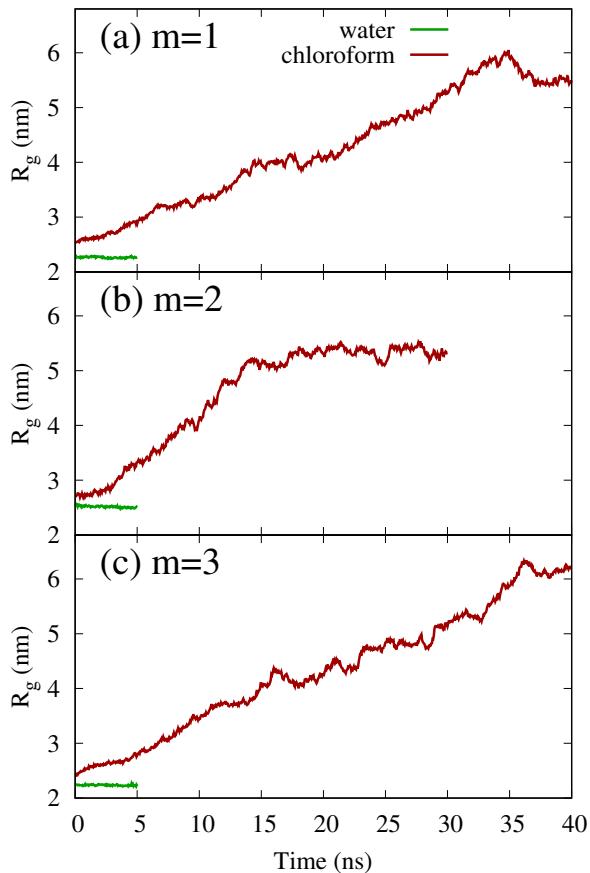
To begin with, a single polymer chain with the degree of polymerization  $N_p = 50$  and 2-hexyldecyl side chains was simulated for each of the three polymer systems of  $m = 1, 2, 3$ . The simulation was started with all-stretched chain conformation as shown in Figure 3.6, where the side chains are removed for visual clarity. This was put in a box of  $105 \times 30 \times 30 \text{ nm}^3$  so that any artifacts in energy calculations arising from the polymer chain interacting with a mirror image of itself can be avoided. It was energy minimized until the forces on every atom were below 100 kJ/mol/nm. The integration time step used was 1 fs for 50-DPP2Py1T and 50-DPP2Py2T and 0.2 fs for 50-DPP2Py3T. Van der Waals interactions were cut off at 1.2 nm. Electrostatic interactions were evaluated using the Particle-Mesh Ewald [88] technique with a real

	$m = 1$	$m = 2$	$m = 3$
$W_f$ (water)	$8.8 \cdot 10^{-3}$	$9.6 \cdot 10^{-3}$	$10.4 \cdot 10^{-3}$
$W_f$ (chloroform)	$5.8 \cdot 10^{-3}$	$6.3 \cdot 10^{-3}$	$6.9 \cdot 10^{-3}$

**Table 3.4:** Weight fraction ( $W_f$ ) of the polymer in water and chloroform solution.

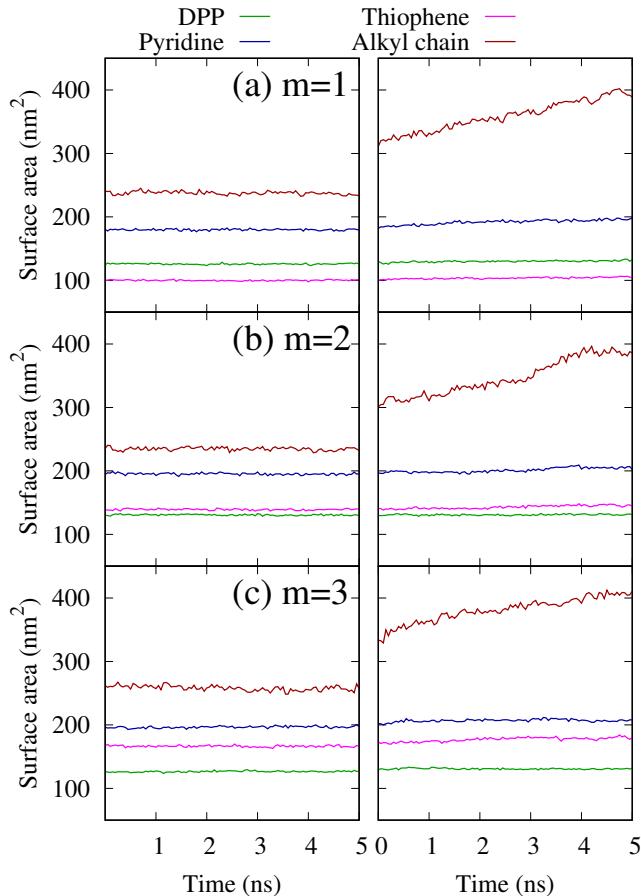
space cut-off of 1.2 nm, a fourth order interpolation and a Fourier grid spacing of 0.16 nm. Upon simulating the structures in vacuum at 300 K in the NVT ensemble, using the velocity rescaling thermostat [139] with a time constant for the temperature coupling of 0.5 ps, the polymer chain collapsed within the first 200 ps into a globule as shown in Figure 3.6. The systems are considered collapsed when the  $R_g$  value has stabilized.

The resulting collapsed polymer structure was then put into two different kinds of solvents to check for the solubility. Chloroform was used as the organic solvent (good) and water as the polar solvent (poor) for the simulations. We used the SPC/E water model [140] and the GROMOS force field for the chloroform solvent, respectively. The single polymer globule was solvated in 259888 (59495) molecules of water (chloroform) and equilibrated in a cubic box of  $20 \times 20 \times 20 \text{ nm}^3$  with periodic boundary conditions. After energy minimization, simulations in the NVT ensemble were performed for 800 ps at room temperature (300 K). The solution was then simulated for 1 ns in the NPT ensemble maintaining the pressure at 1 atm using the Berendsen barostat [141] with a time constant for pressure coupling of 2 ps. These equilibration steps were followed by production runs for 5 ns in water and for up to 40 ns in chloroform. The calculations in water were stopped at 5 ns because there was no significant change in the values of  $R_g$ , as seen (green curve) in Figure 3.7. The calculations in chloroform yielded an increase in the values of  $R_g$  even after 30 ns for  $m = 1,3$ . The resulting weight fraction ( $W_f$ ) in the solution for different values of  $m$  are shown in Table 3.4.



**Figure 3.7:** Radius of gyration of a single 50-DPP2PymT chain for  $m = 1, 2, 3$  in water (green) and chloroform (red). Water is a poor solvent and the polymer remains coiled, while in chloroform (a good solvent) swelling is clearly visible.

As is apparent from Figure 3.7, the simulations of the solvated polymer chains starting from the coiled conformations exhibit qualitatively contrasting behavior depending on the solvent: while the chain remains collapsed into the globule in water, it swells and eventually opens in chloroform. It has been well established that the intramolecular hydrogen bonding in the DPP compounds hinders its solubility in any kind of solvent. In order to have the DPP compounds solvated in water their polarity



**Figure 3.8:** SASA (in  $\text{nm}^2$ ) for the individual segment types (DPP, thiophene, Py, and alkyl side chains) of single DPP2PymT polymer chains with  $m = 1, 2, 3$  in water (left) and chloroform (right).

needs to be enhanced by the addition of ionic functional groups [142], or in some cases also by carboxylic acid functional groups [143]. The absence of any such functional groups in the simulated polymers leads to the low solubility in water. In contrast, the addition of alkyl side chains, as has been implemented in the present simulations enhances the solubility in organic solvents [21]. This trend of solubility is well represented by the radius of gyration,  $R_g$ , in water and chloroform. We can see from Figure 3.7 that the values of  $R_g$  at  $t = 0$  are different for the

green and for the red curves. This is because the polymer shows swelling characteristics in chloroform during NVT and NPT equilibration.

In order to analyze the solubility trend in more detail, we determine the solvent accessible surface area (SASA) for the segment groups (DPP, thiophene, pyridine, and alkyl side chains) of the DPP2PymT polymer chains for  $m = 1, 2, 3$ , respectively. Figure 3.8 shows the SASA in water (left) and chloroform (right). We see that surface areas for DPP, pyridine and thiophenes, which form the backbone of the polymer chain remain fairly constant for the duration of 5 ns in both water and chloroform. However, significant differences can be seen for the SASA of the alkyl chain. It remains constant in water while it increases rapidly in chloroform. This is in line with the experimental findings suggesting that the purpose of the side chains in DPP polymers is to aid in solubility in organic solvents which results in the swelling of the polymer chain in chloroform [21].

The above shows that MD simulations of  $N_p$ -DPP2PymT in solutions yield qualitative results as expected from the literature, supporting the validity of the developed force field.

## 3.5 Melt simulations

Conjugated polymers exhibit a significant change in structural characteristics at high temperatures [144]. As such structural changes often govern the performance of the polymer in opto-electronic devices [145], we also carried out simulations of 50-DPP2PymT melts to study glass-transition, as well as the segmental mobility, by analyzing torsional relaxations of the conjugated substituents.

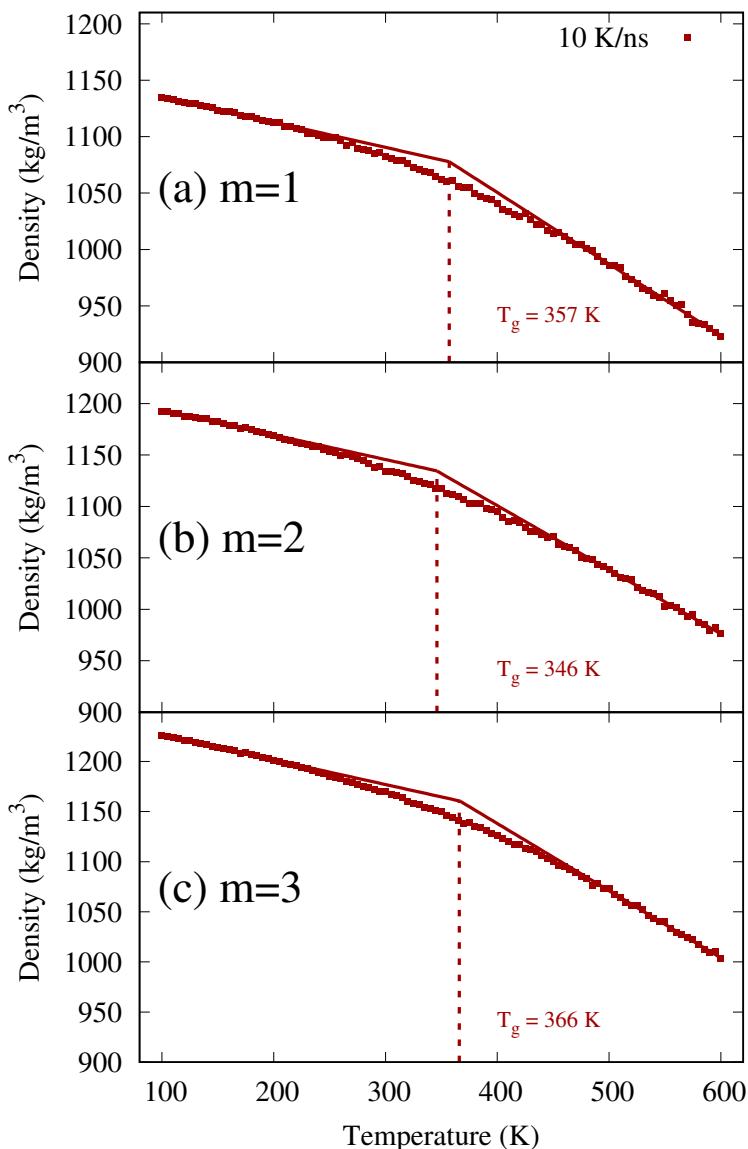
### 3.5.1 Glass transition

Initial configuration with the same degree of polymerization ( $N_p = 50$ ), for polymer melt simulations, has been created with 15 chains in their

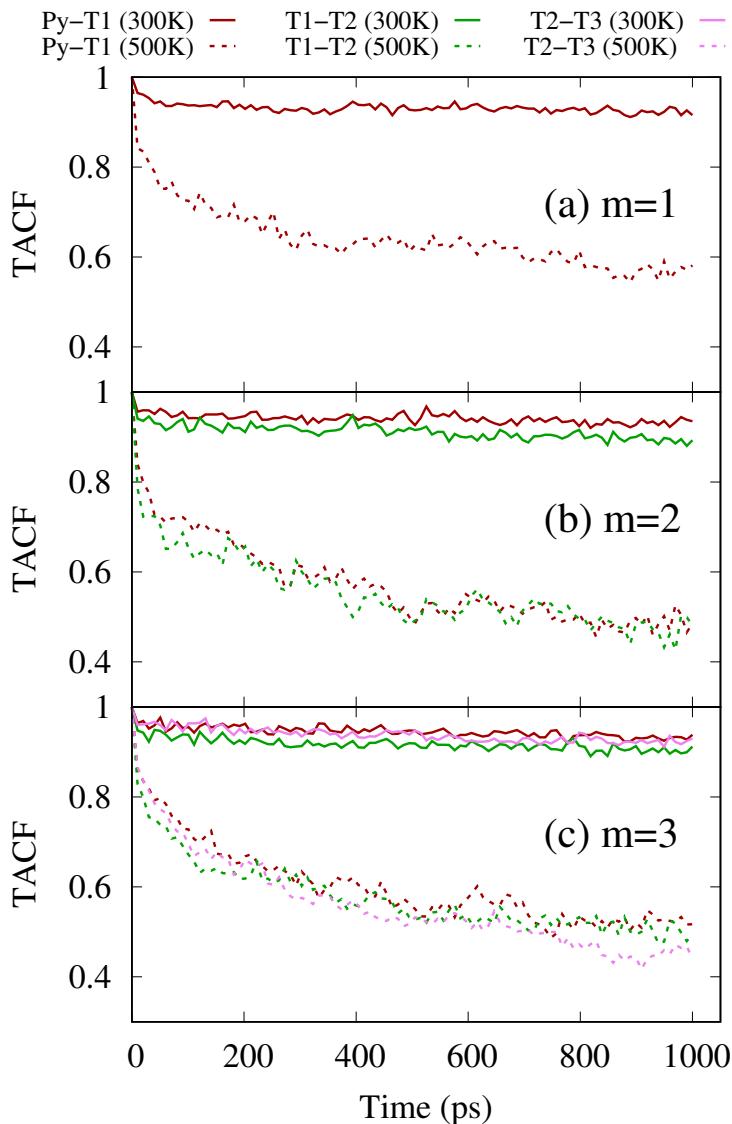
extended state. One single extended chain was aligned along the  $x$ -axis and duplicated 3 times along the  $y$ -axis and 5 times along the  $z$ -axis. The space between each chain was 5 nm. These chains were energy minimized in a  $105 \times 30 \times 30$  nm $^3$  box until the forces on each atom converged to less than 100 kJ/mol/nm.

For similar DPP polymers without aromatic Py substituents a melting temperature of about 550 K is known from experiment [146]. We know as an empirical estimate that the glass transition temperature ( $T_g$ ) is approximately two-thirds of the melting temperature [147], leading to an expected  $T_g$  of around 360 K. We therefore assume that a temperature of 600 K is sufficiently high to simulate the 50-DPP2PymT melts initially above  $T_g$ . In particular, the melts are first simulated for 2 ns in the NVT ensemble at 600 K using the velocity rescaling [139] thermostat with a time constant of 0.5 ps. The resulting structure from the NVT run was simulated at 1 atm in the NPT ensemble. The pressure was maintained using the Berendsen barostat [141] with a constant for pressure coupling as 2.0 ps until we observed a stable density. These equilibrated systems were then cooled from 600 K to 100 K at a cooling velocity of 10 K/ns to determine the glass-transition temperatures.

Figure 3.9 shows the cooling curves for the three polymers. The melt densities at 600 K increase significantly with the number of thiophene units, amounting to 923 kg/m $^3$ , 973 kg/m $^3$  and 1005 kg/m $^3$  for  $m = 1, 2, 3$ , respectively. Glass-transition temperatures are extracted as the intersection of linear temperature dependence of the density fitted in the regions clearly below (from 100 K to 200 K) and above (from 500 K to 600 K)  $T_g$ . It is clear from Figure 3.9 that the transition is not very definite and occurs gradually, and the glass transition temperatures are obtained as  $(357 \pm 30)$  K for  $m = 1$ ,  $(346 \pm 30)$  K for  $m = 2$ , and  $(366 \pm 20)$  K for  $m = 3$ , respectively.



**Figure 3.9:** Temperature dependence of the density of 50-DPP2PymT melt ( $m = 1, 2, 3$ ) for a cooling rate of 10 K/ns. Linear fits to the data at low (from 100 K to 200 K) and high (from 500 K to 600 K) temperatures are used to obtain glass transition temperatures of  $(357 \pm 30)$  K for  $m = 1$ ,  $(346 \pm 30)$  K for  $m = 2$ , and  $(366 \pm 20)$  K for  $m = 3$ , respectively.



**Figure 3.10:** TACF plots between Py-T1 (red), T1-T2 (green) and T2-T3 (violet) rings for DPP2PymT polymer is shown. The solid lines depict the values at 300 K while the dashed lines represent values at 500 K. The TCAF values are the average of all the 750 dihedrals of one kind formed with 15 polymer chains of 50-DPP2PymT.

### 3.5.2 Segmental mobility

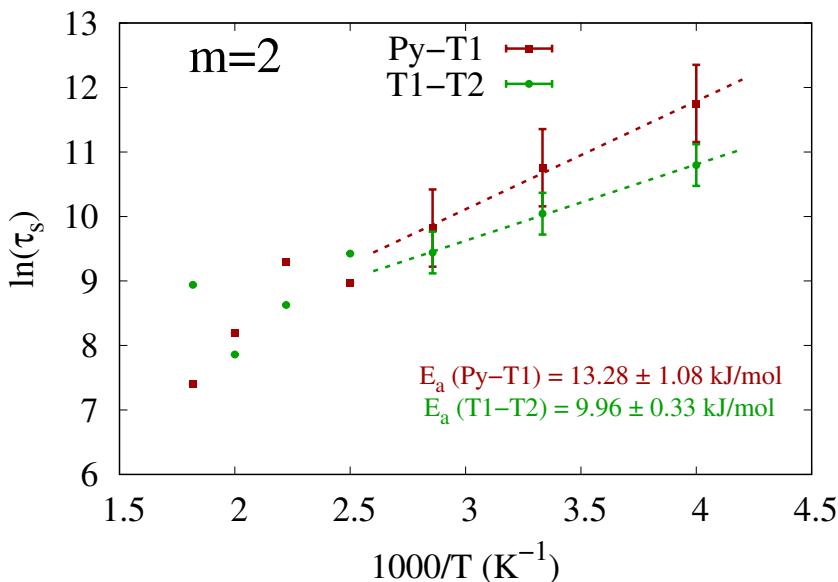
With the details about the glass transition at hand, we study the conformational mobility and isomerization rates of the substituents in the polymer backbone at various temperatures, both above and below  $T_g$ . To this end, we determine the torsional auto-correlation function (TCAF) and analyze its characteristic time decay. We used snapshots from the 10 K/ns cooling simulation for the starting structures. The structures were taken between 600 K and 250 K at an interval of 50 K. Energy minimization, NVT and NPT equilibration for 200 ps was carried out on the starting structures at each temperature. This was followed by a production run for 1 ns in the NPT ensemble. We study the segmental relaxation using a normalized auto-correlation function

$$R_\phi(t) = \frac{\langle \cos \phi_i(t) \cos \phi_i(0) \rangle - \langle \cos \phi_i(0) \rangle^2}{\langle \cos \phi_i(0) \cos \phi_i(0) \rangle - \langle \cos \phi_i(0) \rangle^2} \quad (3.8)$$

where  $\cos \phi_i(t)$  and  $\cos \phi_i(0)$  are the torsional angles at time  $t$  and at start of the simulation, respectively.

The auto-correlation for pyridine-thiophene (Py-T1) and thiophene-thiophene (T1-T2 & T2-T3) dihedrals at 300 K and 500 K are shown in Figure 3.10 for  $m = 1, 2, 3$ . At 300 K, the TCAFs for all dihedrals relax only a little remaining at values above 0.9 at large times, indicating very little change in the relative arrangement of the involved conjugated segments during 1 ns. The observation that the T-T dihedrals show slightly more relaxation than Py-T can be attributed to the minimally softer potential and a barrier-free minimum in the former case [121]. In contrast, at 500 K, the differences between Py-T1 and T1-T2 are not very prominent. Both show, after an initial fast decay, a second much slower component, which is not discernible below the glass transition temperature.

The TACFs clearly point to two characteristic relaxation processes: a rapid one with small relaxation time  $\tau_r$  corresponding to the torsional vibrations around the minima of the dihedral potential, and a slow one



**Figure 3.11:** Logarithm of the relaxation times vs inverse of temperature is shown for DPP2Py2T. Linear fitting has been done at temperatures above the glass transition to extract the activation energy for segmental relaxations. The activation energies is shown on the bottom right for Py-T1 and T1-T2 segmental relaxations.

with a large  $\tau_s$  corresponding to *cis-trans* transitions with energy barriers as discussed in Section 3.3.3. Hence, we determine both relaxation times and the activation energy for torsional relaxations by fitting a sum of two exponentials to the TCAFs. The  $\tau_r$  increases with temperature and typically ranges between 5-20 ps below and 30-100 ps above glass transition temperature for both Py-T (pyridine-thiophene) dihedral and T-T (thiophene-thiophene) dihedrals.

For the slow relaxation process,  $\tau_s$  ranges between  $10^3 - 10^4$  ps at higher temperatures while it is well above  $10^5$  ps for the lower end of the temperature scale. In Figure 4.5 we see that at low temperatures the logarithm of  $\tau_s$  is inversely proportional to the temperature as expected by the Arrhenius equation  $\tau = \tau_0 \exp\left(\frac{E_a}{RT}\right)$ , which holds for polymers below

their glass-transition temperature. From linear fits in this temperature range, we obtain activation energies for the segmental relaxation of  $(13.3 \pm 1.1)$  kJ/mol for the Py-T1 dihedral and  $(9.7 \pm 0.3)$  kJ/mol for T1-T2 dihedral.

These values are qualitatively in correspondence to the dihedral barriers we obtain from force field parametrization and indicates that the torsional motion is the driving mechanism for polymer relaxation and segmental mobility.

### 3.6 Summary

In this chapter we develop an all-atom force field for DPP2PymT for  $m = 1, 2, 3$  by linking the individual fragments that form the building blocks for the polymer. We determine bonded parameters using DFT methods and electrostatic parameters using the CHELPG scheme. This developed force field is validated using a series of solution and melt simulations. Qualitative conclusions drawn from the solution simulations were in line with the well-established fact that chloroform acts a good solvent for DPP polymers with alkyl side chains while water shows poor solvent characteristics. The polymer melt provided reasonable values for the glass-transition temperature ranging between 340 K to 370 K. As a final check we also looked at the auto-correlation function for torsional decay which corroborated well with our force field parameters and softer dihedrals relaxed with lower relaxation times.

# Chapter 4

## Effect of Solvent Removal Rate and Annealing on the Interface Properties in a Blend of a Diketopyrrolopyrrole-Based Polymer with Fullerene

Based on the paper: Vivek Sundaram, Alexey V. Lyulin, Björn Baumeier, J. Phys. Chem. B **126**, 7445 (2022).

### 4.1 Selection of system size

Once we have the recipe (force-field) to perform the necessary classical MD calculations we need to be weary of the system size. It may be noticed that the force-field was developed fragment by fragment. Hence, it is safe to assume that the degree of polymerization for the DPP2PymT system would be independent of the force-field parameters. Now, with the challenge to select the right system size we need to find a balance between how far do the excitations localize on the system the maximum system size that can be computationally evaluated. Hence, in order to select the right system size we will start with the smallest size which is

	$N_{\text{RPA}}$	$N_{\text{occ}}$	$N_{\text{unocc}}$	$N_{\text{BSE}}$	$\varepsilon_{\text{HOMO}}^{\text{KS}}$	$\varepsilon_{\text{LUMO}}^{\text{KS}}$	$\varepsilon_{\text{HOMO}}^{\text{QP}}$	$\varepsilon_{\text{LUMO}}^{\text{QP}}$	$\Omega_{\text{opt}}^{\text{TDA}}$	$\Omega_{\text{opt}}^{\text{full}}$
1-DPP2Py1T	982	96	191	18336	-5.73	-2.93	-6.76	-1.31	2.79	2.44
2-DPP2Py1T	1935	191	381	72771	-5.66	-3.19	-6.57	-1.73	2.45	2.22
3-DPP2Py1T <sup>1</sup>	2890	286	571	163306	-5.65	-3.30	-6.52	-1.90	2.31	2.11
4-DPP2Py1T <sup>1</sup>	3844	381	761	289941	-5.64	-3.36	-6.50	-1.97	2.25	2.06
1-DPP2Py2T	1164	117	233	27261	-5.64	-2.95	-6.65	-1.39	2.68	2.38
2-DPP2Py2T	2300	233	465	108345	-5.55	-3.15	-6.42	-1.73	2.35	2.15
3-DPP2Py2T <sup>1</sup>	3436	349	697	243253	-5.53	-3.23	-6.38	-1.84	2.25	2.07
4-DPP2Py2T <sup>3</sup>	4572	465	929	431985	-5.53	-3.26	-6.36	-1.90	2.19	2.03
Exp. <sup>2</sup>							-5.96	-3.80		1.73
1-DPP2Py3T	1346	138	275	37950	-5.55	-2.96	-6.53	-1.45	2.60	2.34
2-DPP2Py3T	2664	275	549	150975	-5.46	-3.11	-6.32	-1.72	2.29	2.11
3-DPP2Py3T <sup>1</sup>	3982	412	823	339076	-5.45	-3.17	-6.28	-1.81	2.20	2.05
Exp. <sup>2</sup>							-5.77	-3.72		1.74

**Table 4.1:** Number of RPA levels ( $N_{\text{RPA}}$ ), number of occupied ( $N_{\text{occ}}$ ) and unoccupied levels ( $N_{\text{unocc}} = 2N_{\text{occ}}$ ), the product basis set size for BSE calculations ( $N_{\text{BSE}}$ ), Kohn-Sham HOMO ( $\varepsilon_{\text{HOMO}}^{\text{KS}}$ ) and LUMO ( $\varepsilon_{\text{LUMO}}^{\text{KS}}$ ) energies, quasi-particle HOMO ( $\varepsilon_{\text{HOMO}}^{\text{QP}}$ ) and LUMO ( $\varepsilon_{\text{LUMO}}^{\text{QP}}$ ) energies and optical excitation energies for TDA ( $\Omega_{\text{opt}}^{\text{TDA}}$ ) and full BSE ( $\Omega_{\text{opt}}^{\text{full}}$ ) for the  $n$ -DPP2PymT oligomers. All energies are given in eV.

the monomer unit and increase the degree of polymerization in each step until our excitation energies have stabilized while making sure that we do not go beyond the computationally feasible bounds.

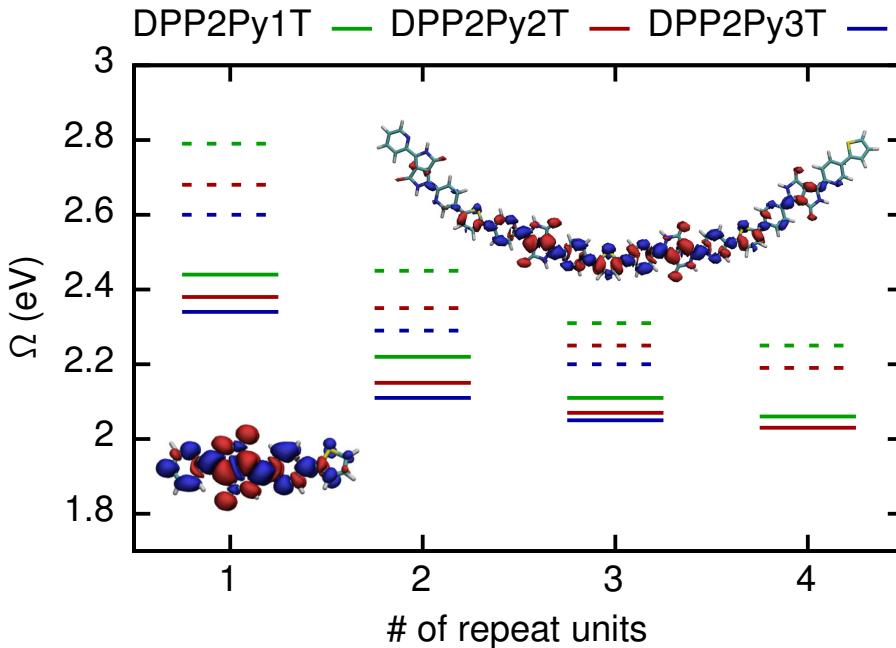
The next step in the process after figuring out the suitable system size is to build a donor-acceptor interface with one or more of the polymers and study their morphologies. For this we begin with GW-BSE calculations on a monomer unit, then increase the repeat units until the optical excitation energy stabilizes.

Geometry optimizations for the  $n$ -DPP2PymT oligomers have been carried out within DFT based on the cc-pVTZ basis set [148] with the PBE0 hybrid functional [149] using the ORCA package [131]. Initial structures for all oligomers have been prepared as suggested by experiments with the thiophenes in trans orientation to each other and the final obtained geometries are nearly planar. For the subsequent GW-BSE calculations, we chose the evGW approach with the PPM. An

optimized auxiliary basis set for cc-pVTZ [46] taken from the Basis Set Exchange [150] has been used in the resolution-of-identity steps. Aiming at a consistent benchmark, we use in each case the full spectrum of single-particle states in the RPA ( $N_{\text{RPA}}$ ), and determine QP corrections to all occupied ( $N_{\text{occ}}$ ) and  $N_{\text{unocc}} = 2N_{\text{occ}}$  unoccupied states. The latter are also taken into account in the product basis for the BSE, i.e.,  $N_{\text{BSE}} = N_{\text{occ}} \cdot N_{\text{unocc}}$ . The explicit values as listed in Table 4.1 for all oligomers. Due to computational limitations, only systems with  $n = 1, 2, 3$  could be treated for  $n$ -DPP2Py3T. The HOMO and LUMO energies as obtained in the KS and QP approaches, as well as the optical excitation energies resulting from the TDA and the full BSE are also given in Table 4.1.

In its ground state, the DPP core is highly electron-withdrawing. Hence, electron-donating aromatic substituents are added to the DPP core [21]. In the compounds studied, pyridine is used as an aromatic substituent which lowers the highest occupied molecular orbital (HOMO) and the lowest occupied molecular orbital (LUMO) levels due to the presence of  $sp^2$ -hybridized nitrogen [5]. Further addition of thiophene (a stronger nelectron donor) increases the number of aromatic substituents in conjugation with DPP. The donor strength increases consequently as evident from the HOMO energy levels determined by cyclic voltammetry measurements [27]. This addition of thiophene also minimally influences the LUMO and the optical excitation energy is lowered from 2.22 eV for  $m = 1$  to 2.05 eV for  $m = 3$  [27].

The quasi-particle energies for HOMO and LUMO follow the expected trend with increasing number of thiophene units  $m$ , as discussed above. We observe that for all oligomers increasing the number of repeat units  $n$  initially, i.e., from monomer to dimer, affects the HOMO energies by about 0.2 eV. Upon further additions of repeat units the values tend to saturate, indicating that the electronic excitations remain localized. Table 4.1 also includes experimental values for the HOMO and LUMO levels of polymers of the three respective structures, as determined by cyclic voltammetry measurements in thin films [27]. The *GW* calculated



**Figure 4.1:** Excitation energy  $\Omega$  (in eV) for the DPP2PymT oligomers as a function of the number of repeat units. The solid lines represent the energies obtained using the full BSE, while the respective dashed ones result from the TDA. The blue and red colour around the molecule shows the increased electron density and increased hole density (Isovalue  $\pm 2 \cdot 10^{-4} \text{ e}/\text{a}_B^3$ ) respectively after exciton formation. The molecules shown are 4-DPP2PyT (top) and 1-DPP2PyT (bottom).

and experimentally observed  $\varepsilon^{QP}$  values are found to demonstrate the same behavior as the experiment with increasing number of thiophenes. Quantitative differences can be attributed to the lack environment effects in the vacuum calculations [151].

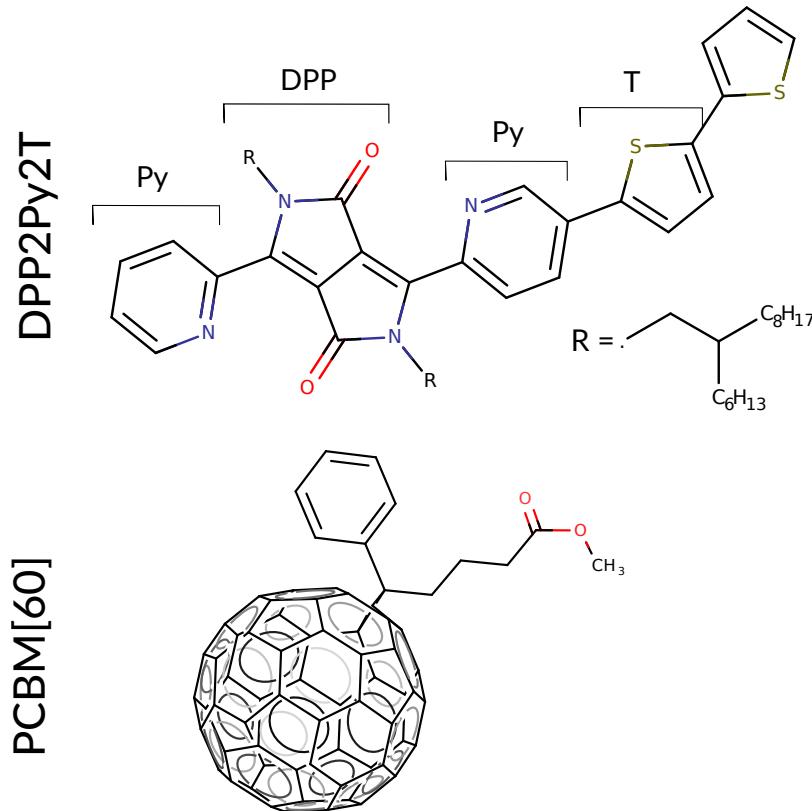
A similar trend is apparent for the energy of the optically active singlet excitations as shown in Figure 4.1 for calculations based on the full BSE (solid lines) and using the TDA (dashed lines), respectively. The values for  $\Omega_{\text{opt}}^{\text{full}}$  obtained for the largest systems are close to those obtained in solution [27] (1.81 eV for DPP2Py1T, 1.73 eV in DPP2Py2T,

1.74 eV in DPP2Py3T). It is also noteworthy that even for the biggest oligomers the difference between the TDA and the full BSE is about 0.2 eV. This is a further indication of the fact that the electronic states and excitations do not delocalize completely over the whole oligomer but remain localized on a few repeat units [152].

This notion is corroborated by the visualizations of the electron density difference upon excitation shown as insets in Figure 4.1. Although the DPP bicyclic core is electron withdrawing, upon exciton generation the electron density is reduced there. In the case of 1-DPP2Py1T the change in electronic density is visible over the entire small molecule. In contrast, for 4-DPP2Py1T we notice that the termini of the oligomer are nearly free of either hole or electron density, with the excitation appearing to be localized on about three repeat units.

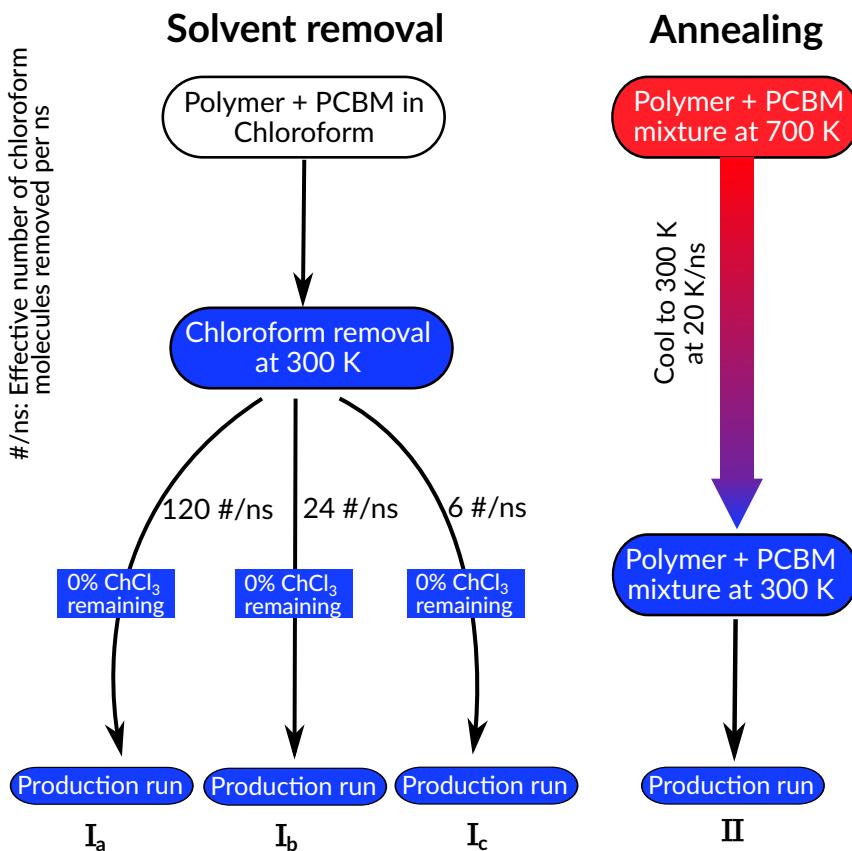
Turning now towards the computational effort required to obtain these results, we first observe that, as also noted in Table 4.1, systems with up to roughly  $N_b = 2500$  basis functions (the same as  $N_{RPA}$ ) can be treated on machines with less than 256 GB of memory. More than 512 GB are needed for the biggest oligomer (4-DPP2Py2T). These requirements can be considered typical for most organic compounds with similar composition and arise from the aforementioned storage of the three-center integrals, containing  $(N_{occ} + N_{unocc}) \cdot N_{RPA} \cdot N_{aux}$  doubles. It should be emphasized that this apparent limitation of the system sizes accessible on regular workstations can be overcome by choosing a lower  $N_{RPA}$ , fewer states in the QP and BSE steps, or a smaller basis set than the large cc-pVTZ set used here for demonstration purposes [134]. We also stress that the peak memory required is reached in the RPA of the *GW* step, and obtaining the solution of the BSE does not increase it as we are making use of the matrix-free Davidson (DPR-MF) method.

Given how the optical excitation delocalizes and the idea of staying in the feasible memory regime, it was decided to proceed with four repeat units for the DPP based polymer. For further in-depth analysis we only consider the polymer DPP2Py2T (see Figure 4.2) in this work but it can



**Figure 4.2:** Chemical structures of DPP2Py2T (top) and PCBM[60] (bottom). For DPP2Py2T, we show the monomer building block consisting of DPP, Pyridine (Py) and Thiophene (T) units. R indicates the position of the 2-hexyldecyl side chain

inferred that these processes can be repeated likewise for other polymer systems as well.



**Figure 4.3:** A brief summary of the two simulation protocols followed in the present study for obtaining the final structures. The first protocol envisages solvent removal at 300 K using three different removal rates 120 #/ns, 24 #/ns and 6 #/ns followed by a production run as shown in the left half of the figure. The second protocol involves first equilibrating the polymer-PCBM[60] mixture at 700 K and then cooling it to 300 K at 20 K/ns and then a production run at the end as seen in the right half of the figure. The four kinds of structures obtained at the end of the protocol are used for structural analysis.

## 4.2 Methodology

Having determined the degree of polymerization for the DPP2Py2T polymer the next challenge is to obtain morphologies that, as realistically as possible, reflect specific structural factors important for the charge separation and transport dynamics [118–120], such as the segmental mobility (dihedral rotations between polymer backbone units which make or break  $\pi$ -conjugation) and the relative orientation of the electron-accepting fullerene with respect to the electron donating fragments of the polymer. However, this is hardly obtainable on the same time (and length) scales involved in the experimental structure processing, in which typically a solvent is removed from a solution with donor and acceptor molecules in minutes over a hot plate. Creating a solvent-vapor interface and removing molecules that cross the interface at regular intervals [153–155] closely emulates experimental procedures but it can result in very long simulations, as the solvent molecules in the middle of the structure may not diffuse to the interface easily. As an alternative, we follow a procedure in which solvent is removed randomly from the solution in steps, followed by equilibration runs after each removal.[153, 156] Effective solvent removal solvent rates  $k_r^{\text{eff}}$  have to be small enough to avoid the generation of artificial structures.

Starting from an initially solvated system in chloroform, we study the structural relaxations and final morphologies upon solvent evaporation, performed using several classical MD simulation protocols depicted in Fig. 4.3: iterative, incremental solvent removal with different effective rates at room temperature and a solvent-free annealing from a liquid melt. The objective for this study is to scrutinize the influence of the simulation protocol on intra- and inter-molecular microstructural properties. We investigate, in particular, the glass-transition temperature of the binary mixture when compared to the pure DPP2Py2T melt, the torsional mobility of the conjugated segments within DPP2Py2T with and without the presence of PCBM[60], as well as the inter-molecular arrangements between DPP2Py2T and PCBM[60] via individual and

combined radial and angular distribution functions, as well as 3D relative density distributions.

Classical all-atom molecular dynamics simulations are performed with the Gromacs 2020.1 simulation package [136]. The Gromos 54A7 force-field [157] is used for PCBM[60] and chloroform molecules while a custom-made force-field, also built upon Gromos 54A7 is used to for the DPP2Py2T polymer [158]. Gromos 54A7 parameters were obtained from The Automated force field Topology Builder [127] where bonded parameters are extracted from the Hessian of the optimised structure and partial charges through the Kollmann-Singh scheme.[128] Starting structures were initialized with 6 polymer chains of DPP2Py2T comprising 4 monomers and 24 molecules of randomly oriented PCBM[60] in a simulation box of size  $12 \times 12 \times 12 \text{ nm}^3$ . This ensured 1:1 w/w ratio of polymer:PCBM[60] to emulate experimental concentrations [159, 160]. The idea behind choosing 4 monomer units in a single chain lies in the delocalization of the electronic density saturating at 4 repeat units when the polymer is pushed to the excited state.[135] Two processes were subsequently followed to ensure a homogeneous blend of the polymer and PCBM[60] mixture. The first involved the solvation of the system in chloroform and then subsequent removal of chloroform at different rates. The second method involved heating the polymer PCBM[60] mixture to its melt temperature and then annealing it back to room temperature. The two methods of preparation have been shown in Fig. 4.3. In the NVT/NPT simulations, pressure was kept constant at 1 atm using the Berendsen barostat[141] while the temperature was maintained using the velocity-rescaling thermostat.[139] Electrostatic interactions were calculated using particle mesh Ewald[88] with a real space cut-off of 1.2 nm. Neighbour lists were updated every 100 time steps using a list cut-off radius of 1.0 nm. Leap frog algorithm as implemented in the md-integrator in Gromacs was used.

In the first simulation protocol, the mixture was solvated with 12000 molecules of chloroform and equilibrated first in the NVT ensemble at and then in the NPT ensemble for  $200 \mu\text{s}$  at 300 K. Due to the high

**Table 4.2:** Parameters for the three different solvent evaporation simulations: number of removed chloroform molecules per step  $N_r$  in stages (1) and (2), the respective lengths of the time intervals  $T_r$ , total time of the simulated evaporation  $T_r^{\text{total}}$ , as well as effective evaporation rate  $k_r^{\text{eff}}$  in molecules/ns.

	$N_r^{(1)}$	$T_r^{(1)}$	$N_r^{(2)}$	$T_r^{(2)}$	$T_r^{\text{total}}$	$k_r^{\text{eff}}$
I <sub>a</sub>	1200	10 ns	120	1 ns	100 ns	120 ns <sup>-1</sup>
I <sub>b</sub>	1200	50 ns	120	5 ns	500 ns	24 ns <sup>-1</sup>
I <sub>c</sub>	1200	200 ns	120	20 ns	2 $\mu$ s	6 ns <sup>-1</sup>

solvent concentration the density saturated to chloroform density of 1.517 g/cm<sup>3</sup>. The next step is solvent evaporation, which we model by removing chloroform from the solution in batches containing  $N_r$  molecules with a time interval  $T_r$  between removal steps. During the whole process, the system is simulated in the *NPT* ensemble described above. As the concentration of DPP2Py2T and PCBM[60] in the solution is very low initially, the step-wise procedure is performed in two stages: In the first stage, nine removal steps are executed with  $N_r^{(1)} = 1200$ , i.e., 10 % of the initially present solvent molecules are removed at each step, until at the end only 1200 chloroform molecules are left. This is followed by stage two, comprising 10 steps with  $N_r^{(2)} = 120$ . As also listed in Table 4.2, we perform simulations at 300 K with different lengths of the time intervals  $T_r$  during the two stages. They are chosen to yield three different total simulation times of the evaporation procedure of 100 ns, 500 ns, 2  $\mu$ s, corresponding to constant effective evaporation rates  $k_r^{\text{eff}}$  of 120 #/ns (I<sub>a</sub>), 24 #/ns (I<sub>b</sub>) and 6 #/ns (I<sub>c</sub>), respectively. Three different  $k_r^{\text{eff}}$  have been chosen to investigate the influence of the simulated speed of evaporation on the structural features in the final room-temperature morphologies.

As an alternative to the explicit solvent evaporation simulations (I<sub>a</sub> - I<sub>c</sub>), we also perform a solvent-free simulation of the DPP2Py2T-PCBM[60] blend. In this annealing approach, the initial mixture is first

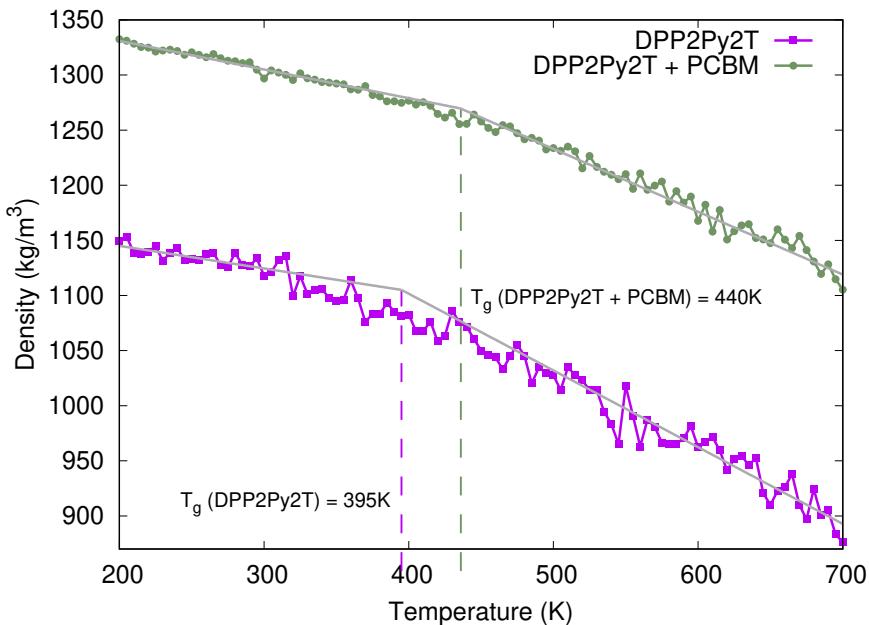
equilibrated at 700 K (which, as will be confirmed later, is about 250 K above the glass transition temperature) and then cooled to 300 K with a cooling rate of 20 K/ns (structure II).

For the structural analysis, *NPT* production runs of 20 ns are performed for the systems obtained via protocols I<sub>a,b,c</sub> and II. The density of the final production runs stabilized at 1285 kg/m<sup>3</sup> for Type I<sub>a,b,c</sub> structures and 1305 kg/m<sup>3</sup> for Type II structures. Note that due to the random selection in the removal of the chloroform molecules, results from five independent simulations runs are averaged in the analysis of systems I<sub>a</sub> - I<sub>c</sub>. Over the last 10 ns of the respective runs, the intermolecular radial distribution function (RDF), angular distribution (ADF) function and the combined distribution function (CDF) between each unit of DPP2Py2T (DPP, Py, T) and C<sub>60</sub> part of PCBM[60] are determined. Also, the preferred side-chain orientation with respect to the DPP block of DPP2Py2T is visualized. All analysis and post-processing was carried out using the Trajectory Analyzer and Visualizer (TRAVIS) code. [161, 162]

## 4.3 Results

### 4.3.1 Glass Transition

Before we turn to the analysis of the inter- and intramolecular structural features of the four DPP2Py2T-PCBM[60] blends from I<sub>a,b,c</sub> and II, we briefly discuss the dependence of the blend density on temperature during the annealing process (II), as shown in Figure 4.4. Initially, at 700 K the DPP2Py2T melt has a density of 870 kg/m<sup>3</sup> which increases to 1150 kg/m<sup>3</sup> at 200 K while the blend density goes from 1110 kg/m<sup>3</sup> at 700 K to 1330 kg/m<sup>3</sup> at 200 K which is considerably higher than the melt density due to the high PCBM[60] density. One can clearly see the glass transition, whose temperature  $T_g$  is estimated from linear fits in the temperature intervals 200 K to 300 K and 550 K to 700 K as  $440 \pm 20$  K. This is significantly higher than the glass-transition temperature of

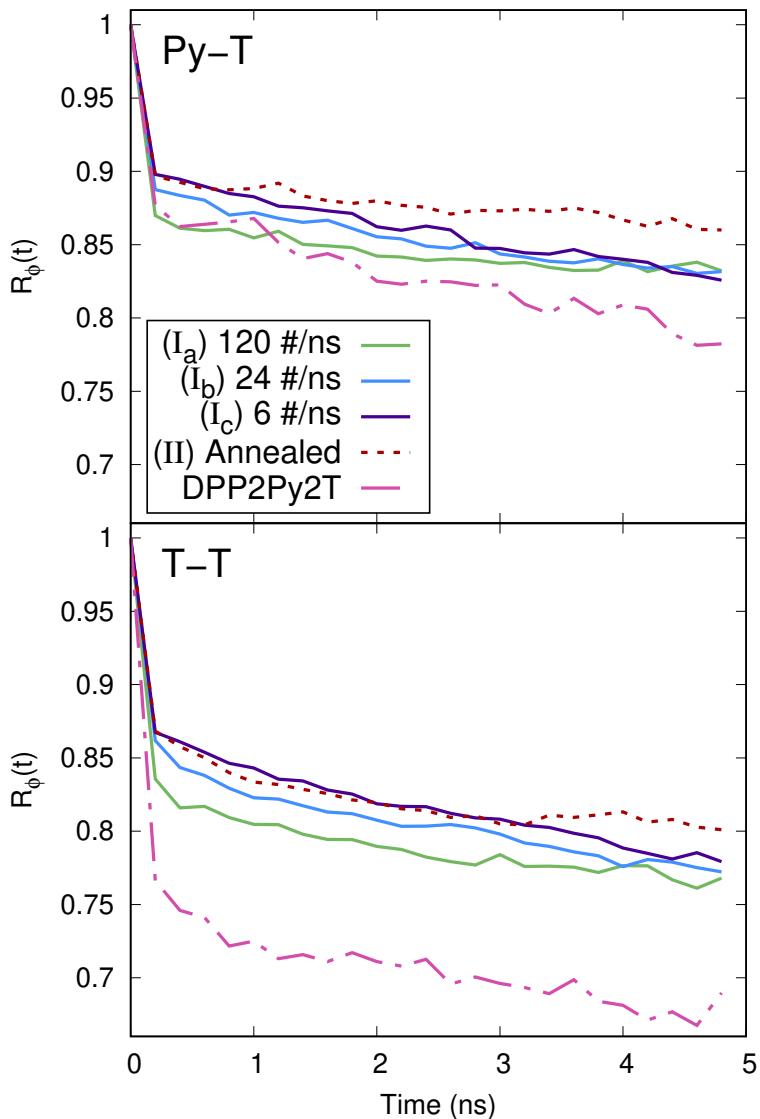


**Figure 4.4:** Density variation of pure DPP2Py2T melt and the DPP2Py2T-PCBM[60] mixture with temperature. Straight lines show linear fitting of the simulated data at high and low temperatures, with the glass-transition temperature of DPP2Py2T melt as  $T_g = 395 \pm 20$  K and for mixture as  $T_g = 440 \pm 20$  K obtained from their intersection.

$395 \pm 20$  K for a pure DPP2Py2T polymer melt due to non-bonded interactions between DPP2Py2T and PCBM[60]. Such an increase of the glass-transition temperature by 40-50 K as a result of the addition of PCBM[60] is consistent with previous studies of blends with similar weight fraction of fullerene and a P3HT polymer [163, 164].

### 4.3.2 Segmental mobility

Torsional auto-correlation functions (TACFs) for the torsional angle between pyridine-thiophene (Py-T) and thiophene-thiophene (T-T) units in the backbone of DPP2Py2T provide insight into the intramolecular



**Figure 4.5:** Torsional auto-correlation function for the Py-T (top panel) and T-T (bottom panel) units in the DPP2Py2T-PCBM[60] blend, obtained by solvent evaporation (solid lines) and solvent-free annealing (dashed line). A reference TACF for the pure DPP2Py2T system is shown by a dash-dotted line.

**Table 4.3:** Relaxation times  $\mathcal{T}_r$  for the Py-T and T-T connections as obtained from the variable solvent removal rates and annealing procedures along with only DPP2Py2T structures at 300 K

	Py-T (ns)	T-T (ns)
I <sub>a</sub> (120 ns <sup>-1</sup> )	100	50
I <sub>b</sub> (24 ns <sup>-1</sup> )	60	40
I <sub>c</sub> (6 ns <sup>-1</sup> )	60	40
II (Annealing)	110	60
DPP2Py2T	40	30

dynamics in the four final morphologies of the DPP2Py2T-PCBM[60] mixture. Specifically, we study the conformational mobility using the auto-correlation function for dihedrals.[165]

$$R_\phi(t) = \langle \cos(\phi(\tau)) \cos(\phi(\tau + t)) + \sin(\phi(\tau)) \sin(\phi(\tau + t)) \rangle_\tau, \quad (4.1)$$

where  $\phi(t)$  is the angle between the normal vectors of the Py and T molecular planes at time  $t$ . The TACF is computed for a time span of 5 ns, averaged over 40 different starting times ( $\tau$ ). To understand the impact of the presence of PCBM[60] on the mobility of the different segments of the polymer, a reference simulation was performed with polymer only at 300 K. Figure 4.5 shows the TACF for the Py-T and T-T units for both simulation protocols and DPP2Py2T polymer only. Firstly, one notices that the relaxations can be classified into two distinct parts-a rapid relaxation that occurs within the first 200 ps characterized by torsional vibrations around the minimum and then later a second decay with a longer timescale corresponding to *cis-trans* flipping between the respective units. The time-scales for the slower relaxation times are obtained from exponential fits to the TACF data and shown in Table 4.3.

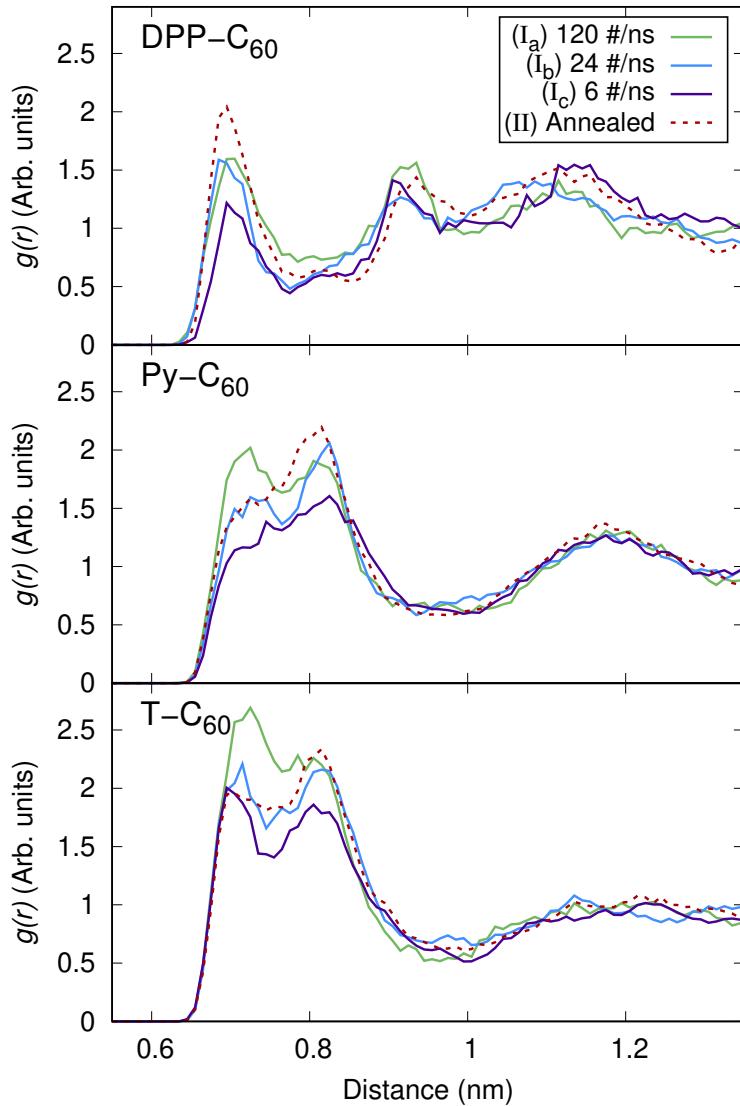
The structures without the presence of PCBM[60] (dash-dotted lines) relax to a greater extent within the first 200 ps and this phenomena is more prominent for T-T link. The relaxation times for both Py-T

and T-T units are much smaller in the absence of PCBM[60] as seen from Table 4.3. We notice that the presence of PCBM[60] increases the relaxation times from 50 ns to 60-110 ns for Py-T units and 30 ns to 40-60 ns for the T-T units. In a nutshell, it can be concluded that the non-bonded interactions between the PCBM[60] and DPP2Py2T in the system acts as a binder to restrict the intra-molecular segmental mobility. This, in turn can be beneficial for charge transport as it helps maintain conjugation within the polymer backbone. Also, structure (II) that was obtained from annealing (red dotted lines), the TACF attains a saturation point faster than structures I<sub>a-c</sub> thereby revealing a higher relaxation time.

Py-T correlation decays, as seen in Fig. 4.5 at slower rate as compared to T-T correlation which can be attributed to the lower torsional barrier for T-T connection compared to Py-T connection.

#### 4.3.3 Fullerene distribution around polymer

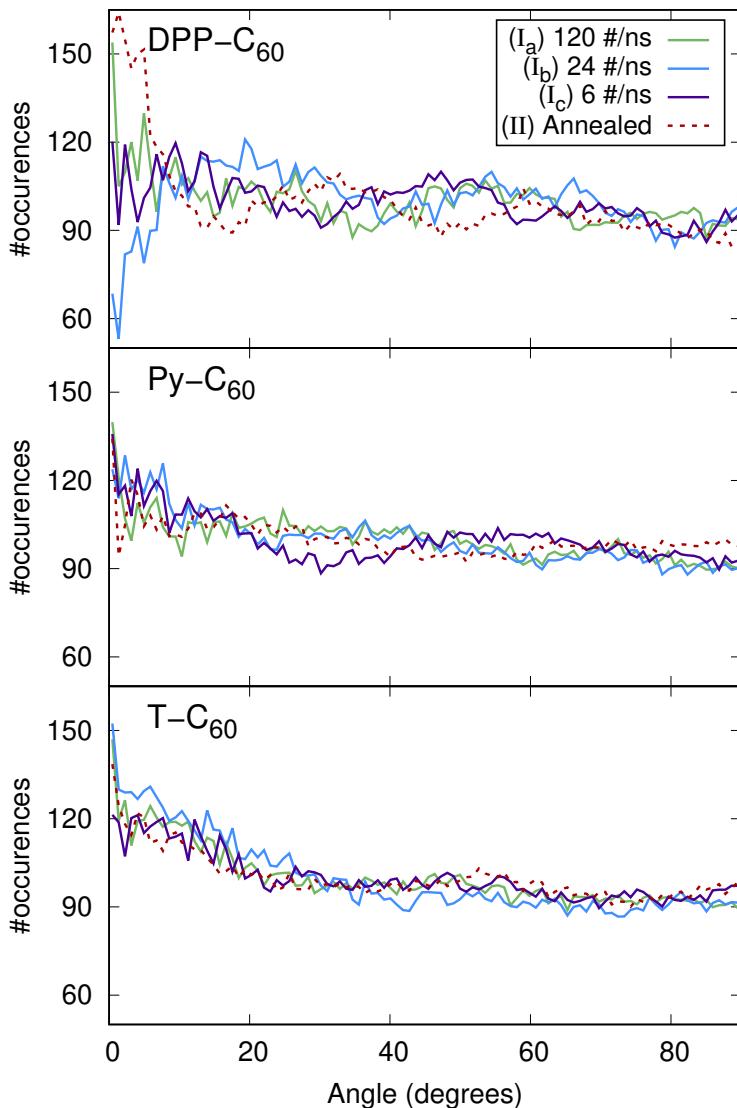
Having seen the intra-molecular structural correlation between pyridine and a thiophene unit and two thiophene units in the polymer system we intend to explore the inter-molecular orientation between the individual units in the polymer and the C<sub>60</sub> unit of PCBM[60]. As stated earlier the relative orientation of the PCBM[60] with the individual units of the polymer is, in principle important for charge transfer phenomena as discussed in the "Introduction" section earlier. To characterize the orientation we use three distribution functions, radial  $g(r)$ , angular  $g(\theta)$  and a combined distribution function  $g(r, \theta)$ . In these calculations we do not consider the Hydrogen atoms in the system and also ignore the side chain on the polymer and PCBM[60]. Also for point to point calculations we will only consider the center of geometry for all units, so the distances in  $g(r)$  are the distance between the center of geometry of the respective units. The results for the radial distribution function can be seen in Fig. 4.6. The first coordination for the C<sub>60</sub> unit w.r.t each unit of the polymer is roughly at about the 0.8 nm. The system obtained after



**Figure 4.6:** Radial distribution function  $g(r)$  between the centers-of-geometry of C<sub>60</sub> and DPP (top panel), pyridine (middle panel) and thiophene (bottom panel) obtained using three different effective solvent evaporation rates (solid lines) and solvent-free annealing (dashed line), respectively.

annealing showed more aggregation of C<sub>60</sub> around the DPP unit of the polymer as visible from a higher peak height. For other simulation pathways, I<sub>a</sub> - I<sub>c</sub>, the DPP unit shows dispersed peaks indicating no preferred distance between the DPP-C<sub>60</sub> pairs. For the Py-C<sub>60</sub> pair two peaks are visible at the first coordination distance of 0.8 nm and at the second coordination distance of 1.2 nm. The first coordination distance between the thiophene and C<sub>60</sub> units are split into two peaks. The system obtained from annealing shows a larger peak at 0.8 nm while the systems obtained from solvent removal show a more prominent peak at 0.75 nm. A general observation is that the C<sub>60</sub> molecule showed a preferential alignment to the thiophene, pyridine and the DPP units in decreasing order as seen from the number of observable peaks and their respective heights. However, in the case when the system was cooled from 700 K the relative height of the all peaks (red curves) are similar indicating that PCBM[60] in these structures does not have a clear preference unlike the structures attained by solvent removal where PCBM[60] shows aggregation around the electron donating units of the polymer. Hence, we can conclude that different removal rates for solvent molecules do not explicitly affect the distance distribution of fullerenes around the polymer backbone. However, annealing from higher temperature has a more prominent effect of C<sub>60</sub> aggregation around the DPP fragment of the polymer.

For the angular and cumulative distribution functions a vector was defined normal to the plane of the polymer unit under study similar to the one we defined in the "Segmental mobility" section. A second vector is defined by connecting the centre of geometries of the polymer unit and C<sub>60</sub>, which is the same as we have used to compute the radial distribution function  $g(r)$ . We calculate the angle between this normal vector and the connecting vector. For instance, 0° represents a fullerene molecule exactly above the plane of the unit of the polymer and 90° represents the fullerene molecule lying within the fragment plane. Cone correction is also applied to maintain consistency of the distribution. The angular distribution in Fig. 4.7 shows an almost flat line with a little monotonic decrease for Py-C<sub>60</sub> and T-C<sub>60</sub> pairs as the angle increases from 0° to

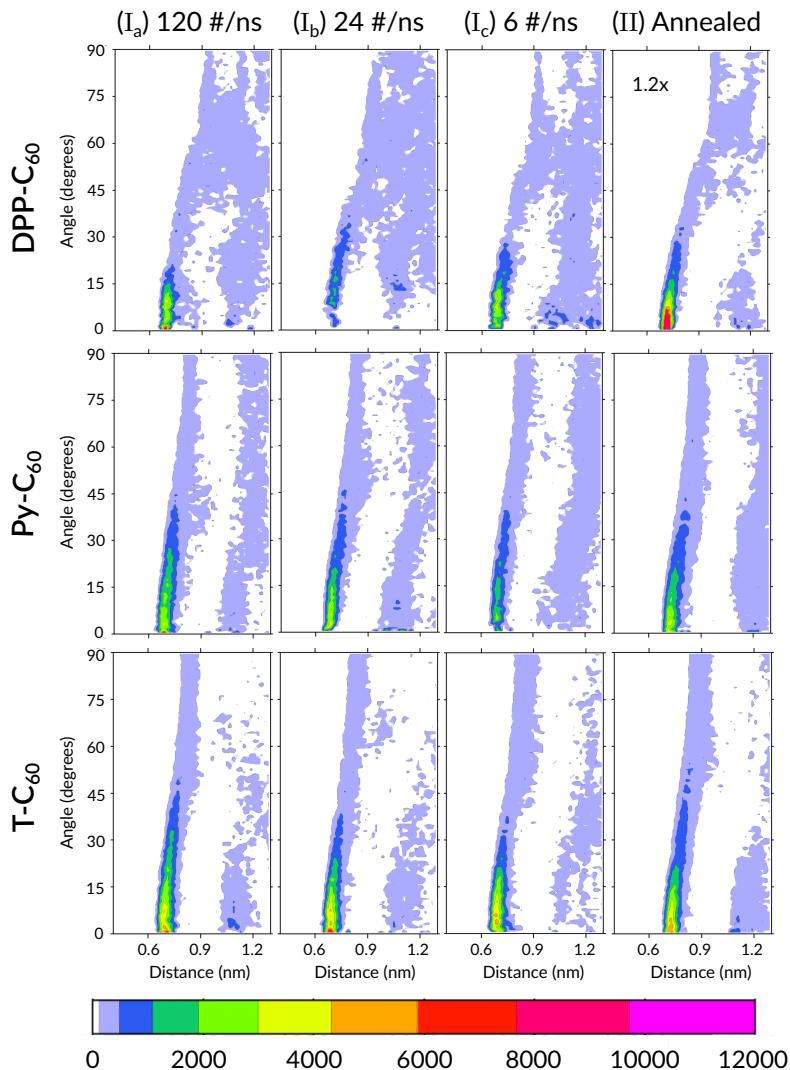


**Figure 4.7:** Angular distribution function defined between normals of the molecular plane for DPP (top panel), pyridine (middle panel) and thiophene (bottom panel) and a vector connecting centers-of-geometry of the polymer unit and C<sub>60</sub> obtained from using three different effective solvent evaporation rates (solid lines) and solvent-free annealing (dashed line), respectively.

90°. Most fluctuations are noticed for the DPP-C<sub>60</sub> pair. It is worthy to note that having a C<sub>60</sub> unit on either side of the polymer plane is considered similar which is why the study was done only upto 90°.

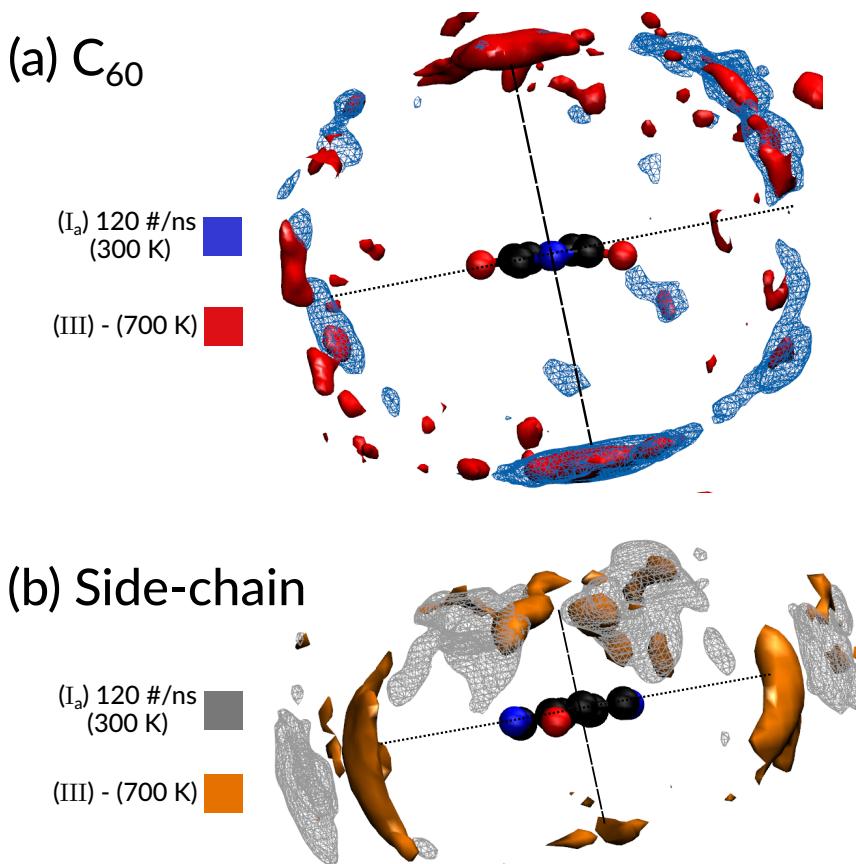
In Fig 4.8 we show the results for a combined distribution function of different units within the polymer and C<sub>60</sub> molecule. The radial and angular distributions present an idea about the relative distance and orientation of the polymer units and the C<sub>60</sub> molecule separately, however, a combined study of the two tells us which inter-molecular positions are most favorable. As a guide to Fig. 4.8 we need to note that each column of the figure contains plots for different simulation protocols while each row contains the plots for arrangement of different units of the DPP2Py2T polymer w.r.t C<sub>60</sub> unit of PCBM[60]. It is also to be noted that the color scale of the plot for the DPP-C<sub>60</sub> pair obtained from the annealing procedure has been scaled down by a factor of 1.2 to match the color scheme of the plots for the four other simulation protocols. The polymer units were taken as the reference while the C<sub>60</sub> units around it were observed.

Observing the first row of plots in Fig. 4.8 we immediately see a distinct bright red spot between 0.6 nm and 0.9 nm at 0° for the system cooled from 700 K. This clearly indicates an accumulation of C<sub>60</sub> units around the DPP units. In the same row the system cooled from 700 K also shows patches of orange at the same spot which led us to conclude that higher temperatures lead to C<sub>60</sub> accumulation around the DPP unit which is detrimental to the type of arrangement preferred for solar cell functioning. Now looking at Py-C<sub>60</sub> and T-C<sub>60</sub> combinations we notice a large band of white space between 0.8 nm to 1 nm. This shows the absence of C<sub>60</sub> units at those distances for any angular orientation. This is consistent with the dip in the peaks for their respective  $g(r)$  as seen in Fig. 4.6. The low density region (white spaces) is credited to the excluded volume for PCBM[60]. That is why it can be seen that higher the concentration of C<sub>60</sub>, more the excluded volume thereby indicating more white spaces as seen in Fig 4.8.



**Figure 4.8:** Combined distribution function involving the radial distribution on the horizontal axes and angular distribution on the vertical axes show the corresponding angle distributions. The colour coding represents the chances of finding a particular combination of the two distribution functions. Note that for the DPP-C<sub>60</sub> pair obtained from annealing, the colour scale has been scaled down by a factor of 1.2 to match with the colour scheme of the other cases.

#### 4.3.4 Influence of side-chain alignment



**Figure 4.9:** The relative densities of (a) C<sub>60</sub> and (b) side-chain around the DPP unit of the polymer is shown when visualized from the polymer plane. The filled bubbles (red and orange indicate density around structure (II) that was cooled from 700 K while the wired mesh (blue and grey) indicate density around structure (I<sub>a</sub>) that was maintained at 300 K and had a chloroform removal rate of 120 #/ns.

In this section we investigate the relative density of the C<sub>60</sub> unit of PCBM[60] and the side-chain (2-hexyldecyl) around the DPP unit of the polymer. The center of geometry of the DPP unit is positioned at the

origin and the center of geometries of C<sub>60</sub> and individual branches of the side-chain are observed in this scenario.

As we see in Fig. 4.9(a) the C<sub>60</sub> molecules show a larger density exactly above and below the plane of the DPP unit. However, one can notice that for structure (I<sub>a</sub>) (blue mesh) the density of C<sub>60</sub> is limited to only one side of the DPP plane while for structure (II) (solid red) there is presence of C<sub>60</sub> density on both sides of the DPP plane. As we do not distinguish between 0° and 180° while calculating angle distribution it can intuitively be understood why this leads to higher concentration of C<sub>60</sub> around DPP for structure (II), as observed in Fig. 4.8 as well. This brings about the question as to how the side-chains present on the DPP would be oriented to facilitate such an alignment.

The relative density of the side-chain around the DPP unit of the polymer can be seen in Fig. Figure 4.9(b) as viewed from the molecular plane. It is noteworthy to remember that the side-chains are attached to the Nitrogen atom of the DPP unit as seen in Figure 4.2. Also, the side-chain contains two branches containing 6 and 8 carbon atoms in each branch respectively as shown in Fig. Figure 4.2. Hence, the two branches are treated separately by considering the center of geometry of each branch individually. In the end, the relative density of both branches are added to reveal the total side chain density around the DPP unit. We can see in Fig. figsidechain(b) that in case of structure (II) (orange solid) the side-chains pack closer to the DPP unit within the plane and very little density is observed above and below the DPP plane indicating a constricted packing which facilitates the positioning of C<sub>60</sub> molecule above and below the DPP plane. However, for structure (I<sub>a</sub>) (gray mesh) the side chains are farther away from the DPP unit indicating a more open structure and also the two branches of the side-chain pack distinctly away from each other in perpendicular planes. This inherently blocks a substantial region above the DPP unit which makes it inaccessible to the C<sub>60</sub> unit.

The solvent concentration and orientation around the DPP unit influences the packing of side-chains. The purpose of side-chains in

these polymers is enhancing solubility to facilitate ease of processing. In simulations, the longer the polymer is exposed to a good solvent (chloroform in this case) the greater is the extent of side-chain expansion which in-turn reduces the available space around the DPP unit (see Fig.S1 of the Supporting Information).

Finally, we note that a local arrangement of PCBM[60] near the electron-donating substituents of the polymer is also commonly considered as the most favorable arrangement, inferred from the fact that DPP2Py2T/PCBM[60] mixtures have been used in active layers of polymer-based solar cells [20]. Here, (efficient) charge generation requires the easy formation of charge transfer states after photon absorption, with the hole on the internally electron-donating substituents of the polymer and the electron on fullerene. The formation of such states can be considered energetically unlikely between the internally electron-accepting DPP and PCBM[60]. While this observation provides merely a very indirect qualitative link between the simulations and real systems, it suggests that the simulation protocol (II) based on annealing of a melt above the glass transition yields qualitatively wrong local arrangements and that explicit solvent effects need to be accounted for in a reliable simulation protocol. Such information is relevant for multiscale modeling approaches that aim at gaining microscopic insight into the charge separation/generation processes in the DPP (and possibly other push-pull architecture) polymer heterojunctions with PCBM [5, 134, 135, 166–168] which rely on an accurate atomistic model of its morphology.

#### 4.4 Summary

In this chapter we analyse the relative orientation of C<sub>60</sub> unit of PCBM[60] around different units of DPP2Py2T polymer for two simulation protocols involving a variation in solvent removal rate and annealing from melt. The glass-transition temperature for the polymer-PCBM[60] mixture was found to be 440 ± 20 K which was expectedly higher than  $T_g$  for DPP2Py2T only. The intra-molecular segmental mobility was lowered

upon addition of PCBM[60] as seen from the higher relaxation times owing to the non-bonded interaction between the PCBM[60] and DPP2PymT. As for local orientation of C<sub>60</sub> around the polymer, solvent removal rate did not influence as much as the temperature variation did. We see that the C<sub>60</sub> unit preferred closer vicinity to the DPP unit of the polymer when cooled from 700 K. This was assisted by the closer packing of the side-chains in the plane of the DPP unit thereby leaving available space for C<sub>60</sub> to come close to DPP.

## Chapter 5

# Quantum-quantum and quantum-classical schemes with projection-based-embedded ***GW*-BSE**

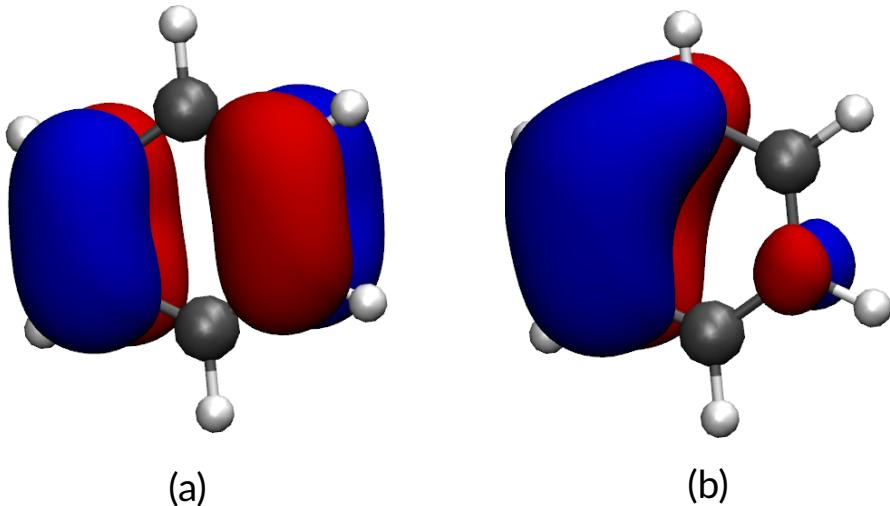
In the previous chapter, we have established from classical MD simulations morphologies of DPP2Py2T-PCBM[60] blends with different local *intra/inter*-molecular configurations [169]. To get insight into the energetics, specifically the energy level alignment of local excitations on donor and acceptor, respectively, and donor-acceptor charge-transfer excitations, a quantum-mechanical theory is required. Treating even a single dimer of DPP2Py2T and PCBM[60] (262 atoms) on the level of the *GW*-BSE method, introduced in Section 2.2, poses a formidable challenge to current state-of-the-art implementations. Such a treatment needs to be embedded in a model for the environment such as the *GW*-BSE/MM model outlined in Chapter 2.4. To reduce the computational costs of the explicit *GW*-BSE calculation on a dimer, we consider in the following an adaptation of projection-based-embedded (PbE) DFT, a quantum-quantum embedding technique, to *GW*-BSE [170–175]. Such an approach allows defining active and inactive subsystems, with only the smaller active subsystem being explicitly treated by *GW*-BSE, while interaction among the subsystems is on DFT level [176–178].

In this chapter, we describe the theoretical background and the implementation of projector-based-embedded *GW*-BSE in VOTCA-XTP, as well as its coupling to a classical environment in a quantum-quantum-classical embedding scheme. As prototypical systems, we consider a DPP bicyclic ring including side-chains, a lipophilic dye in aqueous environment, and a  $\pi$ -stacked dimer of benzene and tetracyanoethylene (TCNE) in water, respectively. In particular, we scrutinize the individual and combined effects of different choices of active regions, the influence of screening from the inactive region, and strategies for basis set truncation on energies of local and charge-transfer excitations, respectively.

## 5.1 Projection-based-embedded (PbE) Density-Functional Theory

Projection-based embedding was introduced by Manby, Miller, and co-workers [170] as a formally exact DFT-in-DFT embedding scheme. Its idea is to partition the density of a full reference system  $n(\mathbf{r})$  into the densities of two subsystems  $n^A(\mathbf{r})$  and  $n^B(\mathbf{r})$ , such that  $n(\mathbf{r}) = n^A(\mathbf{r}) + n^B(\mathbf{r})$ , and where one of them (from now on  $n^A(\mathbf{r})$ ) is considered active, and the other (from now on  $n^B(\mathbf{r})$ ) an inactive embedding density. In this scenario, it is possible to derive a set of effective Kohn–Sham equations for the (orbitals forming) active density embedded in the inactive density, and which result in an exact reproduction of the total energy of the full system. Here, we summarize briefly the main elements of such a DFT-in-DFT embedding approach.

The scheme begins with a self-consistent DFT calculation on the complete reference system in the full-molecule AO basis as described in Chapter 2.1.4. The  $N$  occupied canonical molecular orbitals from this calculation,  $\phi_i(\mathbf{r})$  for  $i = 1, \dots, N$ , are then first transformed into localized orbitals,  $\phi_i^{LO}(\mathbf{r})$ , with a unitary transformation that leaves the total density of the system unchanged. In VOTCA-XTP, we employ the Pipek–Mezey (PM) localization scheme [179], which maximizes the



**Figure 5.1:** The localized state of HOMO on a benzene molecule is shown here for illustrative purposes.

atomic Mulliken population subject to the constraint of keeping the total density fixed. For the actual maximization step, we make use of the unitary optimization algorithm as described by Lehtola et al. [180]. In Figure 5.1 we show as an example the HOMO orbital of benzene in the (a) canonical form and (b) after PM localization. It seems intuitive that after the localization, it is more straightforward to specify a set of atoms in the active regions A and to construct an initial active density from those localized orbitals which have a significant Mulliken population  $q_i^A > q_t$  on these atoms ( $q_t$  is typically 0.4):

$$n^A(\mathbf{r}) = 2 \sum_{\substack{i=1 \\ q_i^A > q_t}}^N |\phi_i^{\text{LO}}(\mathbf{r})|^2. \quad (5.1)$$

With  $n^A(\mathbf{r})$  given, one can determine the density of the inactive region as  $n^B(\mathbf{r}) = n(\mathbf{r}) - n^A(\mathbf{r})$ . Up to this point, we have achieved merely a partitioning of the total density in two subsystems. In the next step, we consider the active density variable and seek a set of effective

equations for the embedded calculation on subsystem A. This requires an embedding potential that takes into account the electrostatic and exchange-correlation interactions with the electrons in subsystem B, while ensuring that the Pauli exclusion principle is followed between the orbitals forming the densities of both subsystems. In projector-based embedding, both objectives can effectively be achieved by raising the energies of orbitals associated to subsystem B to very high energies during the calculation for subsystem A.

If we now consider  $n^A(\mathbf{r})$  the initial density of the active region, and denote the variable density as  $\tilde{n}^A(\mathbf{r})$ , the Fock matrix in the full-molecule AO basis for an embedded (A-in-B) calculation on the electrons in subsystem A is given by

$$\underline{\mathbf{F}}^{A\text{-in-}B} = \underline{\mathbf{h}}_{\text{core}}^{A\text{-in-}B}[n^A + n^B] + \underline{\mathbf{J}}[\tilde{n}^A] + \underline{\mathbf{K}}[\tilde{n}^A] + \underline{\mathbf{V}}_{\text{xc}}[\tilde{n}^A], \quad (5.2)$$

where  $\underline{\mathbf{h}}_{\text{core}}^{A\text{-in-}B}$  is the embedded core Hamiltonian based on the partitioned initial densities  $n^A$  and  $n^B$ . The density dependent terms  $\underline{\mathbf{J}}$ ,  $\underline{\mathbf{K}}$ , and  $\underline{\mathbf{V}}_{\text{xc}}$  are updated in each iterative steps and hence depend on the updated active density  $\tilde{n}^A$ . The embedded core Hamiltonian reads

$$\begin{aligned} \underline{\mathbf{h}}_{\text{core}}^{A\text{-in-}B}[n^A + n^B] &= \underline{\mathbf{h}}_{\text{core}} + \underline{\mathbf{J}}[n^A + n^B] - \underline{\mathbf{J}}[n^A] + \underline{\mathbf{K}}[n^A + n^B] - \underline{\mathbf{K}}[n^A] \\ &\quad + \underline{\mathbf{V}}_{\text{xc}}[n^A + n^B] - \underline{\mathbf{V}}_{\text{xc}}[n^A] + \mu \underline{\mathbf{P}}_B \end{aligned} \quad (5.3)$$

and contains the core Hamiltonian of the full system, the difference between the Hartree, exchange, and exchange-correlation terms for the full system and the initially chosen active subsystem, respectively, as well as a projection term  $\mu \underline{\mathbf{P}}_B$  with the projection operator

$$\underline{\mathbf{P}}_B = \underline{\mathbf{S}} \underline{\mathbf{D}}^B \underline{\mathbf{S}} \quad (5.4)$$

based on the density matrix  $\underline{\mathbf{D}}^B$  of the environment and the atomic orbital overlap matrix  $\underline{\mathbf{S}}$  to ensure orthogonality between the occupied states of the environment and the rest of the active subsystem. The level-shift parameter  $\mu$  is typically of the order of  $10^5 - 10^6$  Hartree. In

the limit  $\mu \rightarrow \infty$ , the two subsystems are exactly orthogonal. The total KS-DFT energy of the embedded system is given by

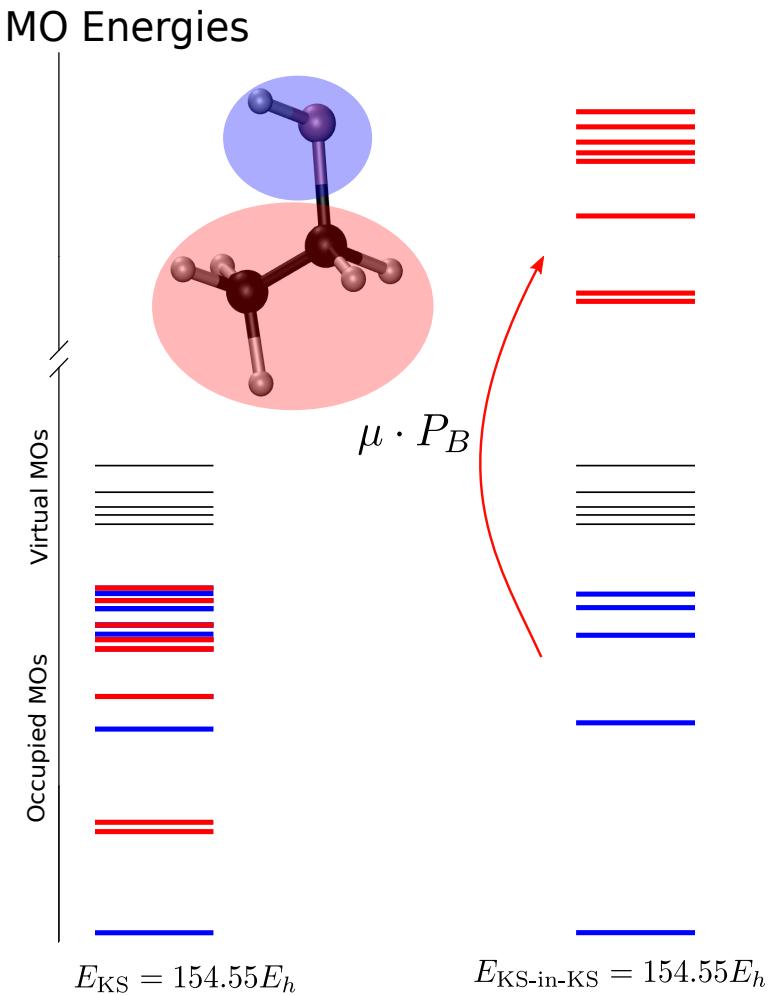
$$\begin{aligned} E^{\text{KS-emb}}[\tilde{n}^A; n^A, n^B] &= E^{\text{KS}}[\tilde{n}^A] + E^{\text{KS}}[n^A + n^B] - E^{\text{KS}}[n^A] \\ &\quad + \text{tr}((\tilde{\underline{D}}^A - \underline{D}^A)(\underline{h}_{\text{core}}^{\text{A-in-B}}[n^A + n^B] - \underline{h}_{\text{core}})), \end{aligned} \quad (5.5)$$

where  $E^{\text{KS}}[n]$  is the Kohn–Sham energy functional from Eq. (2.30), and  $\underline{D}^A$  and  $\tilde{\underline{D}}^A$  the density matrices associated with  $n^A(\mathbf{r})$  and  $\tilde{n}^A(\mathbf{r})$ , respectively.

Figure 5.2 shows as an example the molecular orbital energies for a single ethanol molecule as resulting from the localization of a full KS-DFT calculation and an embedded KS-DFT calculation, respectively. As indicated in the figure inset, the active region for the embedded calculation is chosen to be the blue shaded OH group, while the red shaded rest is assumed to be inactive. In the energy level diagram, the occupied molecule orbital energies are also colored accordingly. The effect of the projection operator in the embedded calculation is indicated as pushing the energies of the occupied molecular orbitals associated with the inactive region (red) to high energies  $\mu$ . Note that the occupied orbitals of the active region do not necessarily have to be exactly identical to the ones from the full calculation. Also note that in the projector-based embedding calculation the virtual orbitals are not affected and therefore identical to the ones obtained in the full KS-DFT reference, i.e., they are not necessarily associated with one of the two regions.

## 5.2 Truncation of the full atomic orbital basis (trPbE)

All matrices composing the Fock matrix  $\underline{F}^{\text{A-in-B}}$  of the embedded A-in-B DFT calculation have been expressed so far in the full molecular atomic orbital basis. As a consequence, the virtual orbitals remain identical to the one of the full reference calculation and also the dimension of the



**Figure 5.2:** Schematic representation of the localized KS orbital energies prior to and after embedding. The blue highlighted region and is the active subsystem while the red highlighted region is the environment. The red orbital levels corresponding to the environment are projected to a very high energy level in order to ensure orthogonality. The black levels represent the virtual orbitals and they remain untouched during this entire process.

eigenvalue problem is the same. The most noteworthy computational gain of the embedding lies in the fact that only a smaller number of occupied states need to be explicitly determined in the self-consistent procedure. It raises the question if it is not possible to reduce the atomic orbital basis for the actual embedding step. Not only would that offer computational savings by decreasing the dimension of the eigensystem, but also by construction affect the virtual orbital space and, in an extreme case, localize the orbitals intrinsically in the active region.

Miller *et al.* have shown that a reduced atomic orbital basis can be constructed by truncating the full basis of the reference calculation via manipulation of the projection term [174, 175]. The procedure begins with an additional classification of atoms in the inactive region into *border atoms* and *distant atoms*. This distinction is based on whether any of the atomic basis functions centered at an atom of the inactive region contributes beyond a threshold to the density of the active region. In practice, this is determined based on the net Mulliken population of an atomic orbital  $\alpha$  in the active density matrix  $\underline{\mathbf{D}}^A$

$$q_\alpha = D_{\alpha\alpha}^A S_{\alpha\alpha}, \quad (5.6)$$

where  $\underline{\mathbf{S}}$  is the overlap matrix. If any of the  $q_\alpha$  exceeds a threshold value (typically  $10^{-4}$ ), the atom associated with basis function  $\alpha$  is added to the list of border atoms. All remaining atoms are the distant atoms. Subsequently, the originally assigned inactive molecular orbitals are also split into border and distant MOs. Border MOs are those inactive molecular orbitals which have a Mulliken population larger than a threshold (again, typically 0.4) on any of the border atoms. All remaining MOs are distant molecular orbitals.

This splitting into border and distant molecular orbitals also allows a similar splitting of the projection operator via the respective density matrices  $\underline{\mathbf{D}}^{\text{border}}$  and  $\underline{\mathbf{D}}^{\text{distant}}$

$$\mathbf{P}_B = \mathbf{P}_B^{\text{border}} + \mathbf{P}_B^{\text{distant}} = \underline{\mathbf{S}} \underline{\mathbf{D}}^{\text{border}} \underline{\mathbf{S}} + \underline{\mathbf{S}} \underline{\mathbf{D}}^{\text{distant}} \underline{\mathbf{S}}. \quad (5.7)$$

The split projector is now used in Eq. (5.3), however, with different values for the level-shift,  $\mu^{\text{border}} \sim 100 \text{ Ha}$  and  $\mu^{\text{border}} \sim 10^5 \text{ Ha}$ . These steps now allow evaluating the Fock matrix in Eq. (5.2) in a reduced, truncated, basis which only includes the basis functions centered at the active and border atoms. Further reductions are possible by not including all basis functions of the border atoms but only those shells for which a single function  $\alpha$  exceeds  $q_t$  used with Eq. (5.6).

Besides setting up the truncated basis, there are several aspects about the practical implementation of truncated-basis A-in-B calculations that are noteworthy: First, setting up the variable parts of the Fock matrix in Eq. (5.2) is straightforward using the truncated basis. However, the embedded core Hamiltonian Eq. (5.3) including the split projection requires special consideration, as it contains terms depending on the initial active density  $n^A$  and total density  $n^A + n^B$  of the full reference calculation. As a consequence, the embedded core Hamiltonian  $\mathbf{h}_{\text{core}}^{\text{A-in-B}}$  is first setup in the full basis, taken over from the preceding full reference calculation. From this full matrix, only those matrix elements are kept that correspond to elements involving the truncated basis functions. A second crucial factor to consider is creating an initial guess for the embedding SCF after truncation of the basis. One can start with the same approach as for the embedded core Hamiltonian to transform the initial active density matrix  $\underline{\mathbf{D}}^A$  from the full to the truncated basis, denoted as  $\bar{\underline{\mathbf{D}}}^A$ . However, this does not guarantee that the molecular orbitals in the two subsystems are fully orthogonal to each other, leading initially to a significant contribution of the projection operator(s) to the Fock matrix. Additionally, the full subsystem density matrix's idempotency is not preserved using  $\bar{\underline{\mathbf{D}}}^A$ , which causes the orthonormality of the localized molecular orbitals to break down when eliminating atomic orbitals with small coefficients. This can be mitigated, for instance, by canonical purification of the truncated initial active density matrix [181].

### 5.3 Projection-based-embedded *GW*-BSE calculations

On its own, performing a DFT-in-DFT projector-based embedding calculation offers no computational gain as it requires first the self-consistent determination of a full system reference. Such a calculation allows, however, to limit a *GW*-BSE calculation for the electronic excitations on the active region only. As the embedding potential is already included in the preceding subsystem-DFT calculation, there are no changes to the procedure of the *GW*-BSE steps, except that all quantities involved use the embedded Kohn–Sham molecular orbitals (in the full or truncated basis) and their energies as starting point. While the DFT-in-DFT calculation can be shown to reproduce the full reference total energy exactly, we have seen that there are by construction changes in the molecular orbitals, and we therefore cannot in general expect a *GW*-BSE calculation after subsystem-DFT (from now on referred to for short as subsystem-*GW*-BSE) to yield the same excitation energies as a full *GW*-BSE calculation. One can get an indication of what the general changes are by considering, e.g., the expression for the self-energy Eq. (2.88) split into the exchange part  $\Sigma^x$  and correlation part  $\Sigma^c$ . Expressing the former in terms of molecular orbitals it reads

$$\Sigma^x(\mathbf{r}, \mathbf{r}') = -2 \sum_{i=1}^{N_{\text{occ}}} \phi_i(\mathbf{r})\phi_i(\mathbf{r}')v_C(\mathbf{r}, \mathbf{r}'). \quad (5.8)$$

The exchange part of the self-energy is therefore affected by (i) summing over fewer occupied states in the subsystem-*GW*-BSE calculation and (ii) the changes in the molecular orbitals themselves. Note that while  $\Sigma^x$  itself only depends on the occupied orbitals, it enters the quasiparticle energies of both occupied and unoccupied states as evaluated by Eq. (2.85). Therefore, even though the virtual molecular orbitals are unchanged in the subsystem-DFT calculation (using the full basis), their *GW* quasiparticle energies may have different contributions arising from  $\Sigma^x$ .

For the frequency-dependent correlation part, a similar analysis is more complicated. Recall from Chapter 2.2 that it reads

$$\Sigma^c(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int G_1(\mathbf{r}, \mathbf{r}', \omega + \omega') \widetilde{W}(\mathbf{r}, \mathbf{r}', \omega) d\omega', \quad (5.9)$$

where the single-electron Green's function is

$$G_1(\mathbf{r}, \mathbf{r}', \omega) = \sum_m \frac{\phi_m(\mathbf{r})\phi_m(\mathbf{r}')}{\omega - \varepsilon_m - i\eta \text{sign}(E_F - \varepsilon_m)} \quad (5.10)$$

and the determination of the screened Coulomb interaction depends on the irreducible polarizability in the form

$$\chi_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_i^{\text{occ}} \sum_j^{\text{virt}} \frac{\phi_i(\mathbf{r})\phi_j(\mathbf{r})\phi_j(\mathbf{r}')\phi_i(\mathbf{r}')}{\omega - (\varepsilon_j - \varepsilon_i) + i\eta} - \frac{\phi_i(\mathbf{r})\phi_j(\mathbf{r})\phi_j(\mathbf{r}')\phi_i(\mathbf{r}')}{\omega + (\varepsilon_j - \varepsilon_i) - i\eta}. \quad (5.11)$$

Subsystem embedding changes both  $G_1$  (which also leads to the discussed changes in  $\Sigma^x$ ) and  $\chi_0$  via the different orbitals and their energies. For the polarizability, embedding implies several noteworthy modifications. Even if the  $\phi(\mathbf{r})$  and energies  $\varepsilon$  were unchanged, the sum over occupied orbitals is limited to the active occupied orbitals (the ones from the inactive one are found in the virtual space at high energy, and should be excluded from the sum over virtual orbitals). As a result, the screening only has contributions from transitions between occupied orbitals in the active subsystem and virtual orbitals of the combined system (in the full basis calculation), while contributions from transitions from occupied orbitals in the inactive region to all virtual orbitals are removed. The inactive region therefore can be considered static from the perspective of the screened Coulomb interaction, similar to a QM/MM embedding with only static moments in the MM region, as discussed in Chapter 2.4. If additionally a truncated basis restricted ideally to the active region is used, this will also affect the virtual orbitals and essentially limit the transitions to those within the active region, removing charge-transfer like transitions between the subsystems from the response. The effect on the calculated contribution of these CT-like transitions to  $\Sigma^c$  is expected

to be small, however. Both considerations regarding the modifications type of transitions excluded in the screening in subsystem-*GW* are in general then combined with additional effects of changed orbitals and their energies in the active region. From the lack of screening from the now inactive region, one can generally expect the contributions of  $\Sigma^c$  to the quasiparticle energies to be smaller (in absolute values) in the subsystem-*GW* calculation compared to the full *GW* case. In other words, even when the orbitals themselves are only minimally affected (for weakly interacting, non-bonded molecular structures, for instance), one can expect to find the occupied (virtual) quasiparticle energies from the embedded calculation to be below (above) the ones from the full calculations. In particular, the HOMO-LUMO gap in subsystem-*GW* is then larger than the respective gap in full-system *GW*.

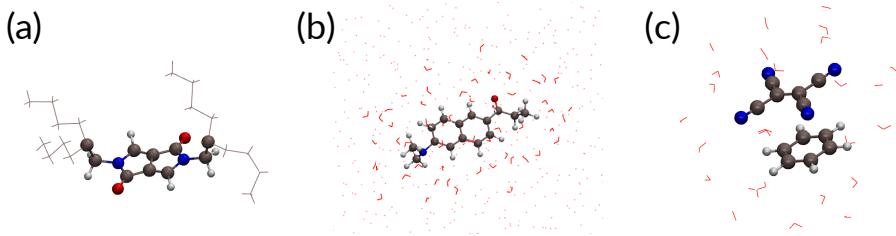
One can make similar examinations on the level of the BSE. Naturally, as the number of occupied orbitals is reduced, the electron-hole transitions used to expand the two-particle wavefunctions are limited to the transitions starting from the active subsystem. Any changes to the quasiparticle energies as a result of the points discussed above will directly impact the free transition term  $D_{vc,v'c'}$  from Eq. (2.122) in the BSE Hamiltonian. The effects on the exchange and direct (screened) terms of the electron-hole interaction kernel  $K_{vc,v'c'}^x$  and  $K_{vc,v'c'}^d$  are similar to those discussed for  $\Sigma^x$  and  $\Sigma^c$ , respectively. Especially, the reduced screening can be expected to result in stronger electron-hole attraction compared to the full-system calculation and might in turn compensate to some degree the larger quasiparticle gap in the free transition.

In Ref. [114, 182], the authors reported subsystem-*GW*-BSE calculations on a benzene-ammonia dimer, aqueous methylenecyclopropane, and a water-solvated adenine-thymine dimer. The embedding scheme differs from the projection-based-embedding in that they employ mainly another subsystem-DFT approach (sDFT), which starts out from Kohn–Sham-like calculations on fragments and determines the full-system density from them. In such an approach the effective potential in the KS-like equations for a fragment contain contributions from non-additive terms

in the kinetic energy and exchange-correlation potential. We point the reader to Ref. [183–187] for more details of this sDFT approach. Starting out from fragment molecular orbitals and ignoring charge-transfer like transitions between the fragments allows for a natural partitioning of Eq. (5.11) into fragment contributions. Then the screening effects can then, in short, be included when evaluating the screened Coulomb interaction for  $\Sigma^c$  in the active region via a self-consistent reaction field similar to alternative approaches to account for classical polarization effects in *GW*-BSE/MM setting [188, 189] to the one used in this work. For the weakly interacting molecular clusters in Ref. [114], it could be shown that environment screening included like this can recover to a large extent the full-system frontier orbital energies, and the transitions between them. It is not immediately obvious how to adapt the same to an approach starting from a supramolecular basis, what effects embedding has on excitations beyond the fundamental gap, and if changes in the self-energy are restricted to the correlation part only in case of strongly interacting subsystems.

### 5.3.1 Projection-based-embedded *GW*-BSE/MM

The PbE-DFT and by extension PbE-*GW*-BSE methods are quantum-quantum embedding techniques. In Chapter 2.4, quantum-classical embedding methods with explicit atomistic detail (both static and polarizable) have been introduced in a *GW*-BSE/MM scheme. As in this the interactions between the quantum and classical regions are purely represented by electrostatic potentials, it is straightforward to combine both embedding approaches into one quantum-quantum-classical embedding scheme. The static moments of the MM regions continue acting as an additional background potential to the quantum-quantum region. Similarly, the electric field acting on the polarizable sites in the MM region is created by the total electron density (plus the nuclei) of the subsystem QM region. It is worth highlighting that when a polarizable model is used in the PbE-*GW*-BSE/MM, the outer SCF coupling the quantum and classical regions implies that during such a calculation also



**Figure 5.3:** Molecular structures used as test systems for PbE-GW-BSE: (a) DPP bicyclic ring with branched alkyl side chains, (b) prodan in close (quantum) and distant (classical) water, (c) a water-solvated benzene-TCNE dimer.

the density of the inactive region can respond to the polarization of the MM region.

## 5.4 Frontier orbital and near-gap electron-hole excitation energies

In this section, we present the results of PbE-GW-BSE calculations without and with classical embedding for the three test systems shown in Figure 5.3: (a) a single DPP bicyclic ring with branched alkyl side-chains, (b) prodan, a polarity-sensitive dye, solvated in water, and (c) a benzene-TCNE donor-acceptor pair in water. For all systems, we study the effects of the PbE with (trPbE) or without basis truncation on the frontier orbitals as well as the selected local or CT excitations. We pay special attention to the differences in contributions of  $\Sigma^x$  and  $\Sigma^c$  to the quasiparticle energies, and of  $K^x$  and  $K^d$  to the BSE energies between full and embedded calculations. In addition to this, each of the three test systems is here chosen with specific objectives:

- (a) For the DPP molecule with branched side chains, we intend to demonstrate a PbE-GW-BSE calculation for a system in which the two regions are connected by a covalent bond. We investigate

the sensitivity of the embedding results on choice the choice of the active region.

- (b) Prodan in water has been chosen to showcase the quantum-quantum-classical PbE-GW-BSE/MM approach. Also here, we consider the influence of the choice of the active regions on the predicted excitation energies.
- (c) The donor-acceptor benzene-TCNE dimer solvated in water is used to evaluate the differences in embedding effects on localized and charge-transfer type excitations.

All calculations have been done using the VOTCA-XTP package which interface to the ORCA software [131] for the full system reference DFT calculations. The def2-TZVP basis-set[148] with an optimized auxiliary basis [133] along with the PBE0 hybrid functional [149] has been used in all DFT and subsequent *GW*-BSE calculations. If not stated otherwise,  $G_0W_0$  variant has been chosen with the Plasmon-Pole model (PPM) for self-energy calculation in the *GW* step. To maintain consistency, we use in each case the full spectrum of single-particle states in the RPA ( $N_{\text{RPA}}$ ), and consider all occupied ( $N_{\text{occ}}$ ) and the lowest  $N_{\text{virt}} = 2N_{\text{occ}} - 1$  unoccupied states for the calculation of the QP corrections and the expansion of the BSE product basis, i.e.,  $N_{\text{BSE}} = 2N_{\text{occ}}(2N_{\text{occ}} - 1)$ .

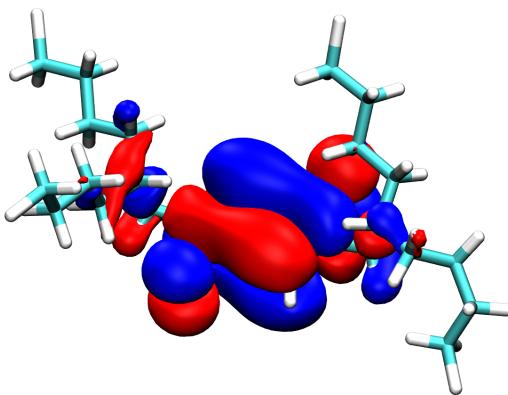
#### 5.4.1 DPP bicyclic ring with branched alkyl side-chains

As a first test system, we consider a single DPP unit from the class of DPP2PymT polymers studied in this work. Alkyl side chains with a branched structure are attached to the respective nitrogen atoms. A short  $\text{C}_2\text{H}_2$  group contained the branching point, and each branch is formed by  $\text{C}_4\text{H}_9$ , as can also be seen in Figure 5.3(a). The geometry of this structure is optimized in vacuum (DFT with the PBE0 functional and def2-TZVP basis). In (conjugated) polymer systems, it is often assumed that the frontier orbitals relevant for charge transport are localized on the

**Table 5.1:** Overview of the number of basis functions ( $N_{\text{basis}}$ ), number of functions in the auxiliary basis ( $N_{\text{aux}}$ ), number of occupied states included in the quasiparticle calculation and the BSE product basis ( $N_{\text{occ}}$ ), idem for virtual states ( $N_{\text{virt}}$ ), number of transitions in the RPA ( $N_{\text{RPA}}$ ), and dimension of the BSE Hamiltonian ( $N_{\text{BSE}}$ ), for the full-GW-BSE, PbE-GW-BSE, and trPbE-GW-BSE calculations on the three testsystems.

	$N_{\text{basis}}$	$N_{\text{aux}}$	$N_{\text{occ}}$	$N_{\text{virt}}$	$N_{\text{RPA}}$	$N_{\text{BSE}}$
<b>DPP+alkyl</b>						
full-GW-BSE	1194	2940	115	229	124085	52670
PbE-GW-BSE	1194	2940	51	101	55029	10302
trPbE-GW-BSE	1152	2835	51	101	52887	10302
<b>aqueous prodan</b>						
full-GW-BSE	2994	7377	336	672	893088	451584
PbE-GW-BSE(dye)	2994	7377	61	121	162138	14762
PbE-GW-BSE	2994	7377	96	191	255168	36672
trPbE-GW-BSE(dye)	2453	6035	61	191	129137	14762
<b>aqueous benzene-TCNE</b>						
full-GW-BSE	1650	4062	183	365	268461	133590
PbE-GW-BSE	1650	4062	53	105	77751	11130
trPbE-GW-BSE( $10^{-4}$ )	1384	3400	53	105	63653	11130
trPbE-GW-BSE( $10^{-5}$ )	1638	4032	53	105	77115	11130

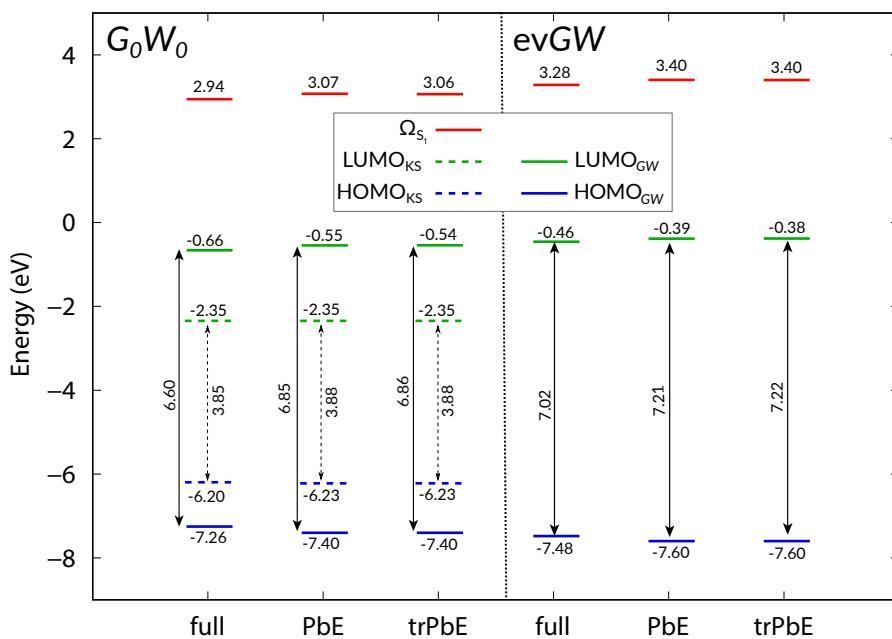
actual functional backbone, and that the side-chains do not participate in the electronic processes. For the testing of the PbE-GW-BSE approach, such chemical intuition suggests to actually select only the DPP unit including the nitrogen atoms as the active region and the complete two branched alkyl side chains into the inactive one. However, as can be seen from the isosurfaces of the HOMO from a full KS calculation in Figure 5.4, the occupied frontier orbital extends further into the side chains, even slightly beyond the branching atom.



**Figure 5.4:** Isosurfaces of the KS HOMO ( $\pm 0.01 \text{ a}_\text{B}^{-3}$ ) of the single DPP bicyclic ring with branch alkyl side chain as obtained from regular KS-DFT.

The choice of only the DPP core as the active region is therefore expected to yield considerable modifications to the occupied electronic states. With the objective to first inspect the different effects in the *GW*-BSE steps upon embedding without significant modification of the respective orbitals, we extended the active region to include the  $\text{C}_2\text{H}_2$  groups until the respective branching carbon atom for all following PbE calculations. Figure 5.5 shows the results of  $G_0\text{W}_0$  (left) and ev $GW$  (right) calculations for the HOMO, LUMO and  $\Omega_{\text{S}1}$  energies for full, PbE, and trPbE calculations, respectively. The results are also collected in Table 5.2, in which also the individual contributions to the quasiparticle energies according to  $\varepsilon^{\text{QP}} = \varepsilon^{\text{KS}} - V^{\text{xc}} + \Sigma^{\text{x}} + \Sigma^{\text{c}}$ , the HOMO-LUMO gap, and contributions of the free transition energy ( $D$ ), the exchange ( $K^{\text{x}}$ ) and direct ( $K^{\text{d}}$ ) terms of the electron-hole interaction to the electron-hole excitation energy  $\Omega_{\text{S}1}$  are listed. The quantity  $\Delta_{\text{full}}^{\text{PbE}}$  is the difference of the PbE calculation to the full one, and  $\Delta_{\text{PbE}}^{\text{trPbE}}$  measures the additional change of the basis set truncation with respect to the PbE calculation.

We first start the discussion of the KS HOMO and LUMO energies on DFT level only. They are shown in Figure 5.5 as dashed lines (blue:



**Figure 5.5:** Energy level (in eV) diagram for a DPP bicyclic ring with branched alkly side-chains showing HOMO (blue), LUMO (green), and lowest electron-hole excitation (red) energies on KS level (dashed) or *GW*-BSE level (solid), as resulting from full, PbE, and trPBE calculations, respectively. Left panel shows results from  $G_0W_0$  calculations, right panel from ev $GW$  calculations. Double-headed arrows additionally indicate the respective HOMO-LUMO gaps. See also Table 5.2 for details.

HOMO, green: LUMO) in the  $G_0W_0$  panel (note that they are identical in the ev $GW$  case). With the choice of the active region as discussed above, the change in the energy of the HOMO level upon embedding is small (-0.03 eV), and the LUMO is unaffected as expected from the theoretical basis given in Section 5.1. Using a truncated atomic orbital basis does not yield any changes of the KS electronic structure at the shown accuracy. As one can see from the values  $N_{\text{basis}}$  in Table 5.1, the basis truncation procedure as described in Chapter 5.2 removes only 42 of the 1194 functions, or about 3.5 %.

**Table 5.2:** Results of full- $G_0W_0$ -BSE, PbE- $G_0W_0$ -BSE and trPbE- $G_0W_0$ -BSE calculations for a DPP bicyclic ring with branched alkyl side-chains: KS and QP HOMO and LUMO energies together with the individual contributions from the exchange-correlation potential  $V^{xc}$ , and the self-energy split in  $\Sigma^x$ , and  $\Sigma^c$ , according to  $\varepsilon^{QP} = \varepsilon^{KS} - V^{xc} + \Sigma^x + \Sigma^c$ , as well as KS and QP HOMO-LUMO gaps. The lowest electron-hole excitation energy from the respective BSE calculation  $\Omega_{S_1}$  is also split into the free transition energy ( $D$ ) and contributions from the exchange ( $K^x$ ) and direct ( $K^d$ ) terms of the electron-hole interaction. All energies in eV.

	full	PbE	$\Delta_{\text{full}}^{\text{PbE}}$	trPbE	$\Delta_{\text{PbE}}^{\text{trPbE}}$
<b>DFT</b>	$\varepsilon_H^{\text{KS}}$	-6.20	-6.23	-0.03	-6.23
	$\varepsilon_L^{\text{KS}}$	-2.35	-2.35	0.00	-2.35
	$E_{\text{gap}}^{\text{KS}}$	3.85	3.88	0.03	3.88
<b><math>G_0W_0</math></b>	$V_H^{xc}$	-12.52	-12.47	0.05	-12.47
	$\Sigma_H^x$	-14.08	-14.02	0.06	-14.02
	$\Sigma_H^c$	0.50	0.38	-0.12	0.38
	$V_L^{xc}$	-11.25	-11.23	0.02	-11.23
	$\Sigma_L^x$	-7.52	-7.50	0.02	-7.50
	$\Sigma_L^c$	-2.04	-1.92	0.12	-1.92
	$\varepsilon_H^{QP}$	-7.26	-7.40	-0.14	-7.40
	$\varepsilon_L^{QP}$	-0.66	-0.55	0.11	-0.54
<b>BSE</b>	$E_{\text{gap}}^{QP}$	6.60	6.85	0.25	6.86
	$D$	7.63	8.00	0.37	8.00
	$K^x$	0.54	0.56	0.02	0.56
	$K^d$	-5.23	-5.49	-0.26	-5.50
	$\Omega_{S_1}$	2.94	3.07	0.13	3.06
					-0.01

Considering from now on the actual  $G_0W_0$  results, one can first identify the typical effects of quasiparticle corrections on the HOMO and LUMO energies. In the full approach, the HOMO energy is lowered by 1.06 eV, and the LUMO energy raised by 1.69 eV, such that the HOMO-LUMO gap increases by 2.75 eV. Upon PbE, however, the respective

shifts are more pronounced, by 0.14 eV (0.11 eV) for the HOMO (LUMO), cf. Table 5.2. The quasiparticle gap  $E_{\text{gap}}^{\text{QP}}$  is hence increased by 0.25 eV. Within the  $GW$  formalism, a larger gap can often be associated with reduced screening. This notion is corroborated by the data provided for the contributions from  $V^{\text{xc}}$ ,  $\Sigma^{\text{x}}$ , and  $\Sigma^{\text{c}}$  to  $\varepsilon^{\text{QP}}$  in Table 5.2. From the respective  $\Delta_{\text{full}}^{\text{PbE}}$  for, e.g., the HOMO, one can see that the differences from the KS exchange-correlation potential and the exchange part of the self-energy almost compensate (note that  $V^{\text{xc}}$  is used with a negative sign in Eq. (2.85)), and that consequently the differences in quasiparticle energies between full- and PbE- $G_0W_0$  are practically determined by effects in the correlation part  $\Sigma^{\text{c}}$  alone. For the LUMO, very similar observations can be made. Again, due to the minimal reduction of the basis set, no significant changes are noted in trPbE- $G_0W_0$  compared to PbE- $G_0W_0$ .

Turning now towards the respective BSE results, we first note in the full calculation that the excitation energy results as 2.94 eV, which consists of the effective free-transition energy  $D = 7.63$  eV, and the exchange ( $K^{\text{x}} = 0.54$  eV) and direct ( $K^{\text{d}} = -5.23$  eV) parts of the electron-hole interaction kernel. The fact that  $D$  exceeds  $E_{\text{gap}}^{\text{QP}}$  by about one eV indicates that the electron-hole excitation is not exclusively given by a HOMO-LUMO transition. Qualitatively, the same holds also in the PbE (and trPbE) calculations. The  $S_1$  energy in PbE- $G_0W_0$ -BSE results with 3.07 eV only 0.13 eV higher than in the full reference. This is noteworthy because as discussed above  $E_{\text{gap}}^{\text{QP}}$  is larger by almost twice this value. Upon inspection of the individual contributions on the BSE level given in Table 5.2, one first observes that the free interlevel contribution  $D$  is larger by 0.37 eV and exceeds the relative increase on  $E_{\text{gap}}^{\text{QP}}$ , indicating that the lack of screening in  $\Sigma^{\text{c}}$  is larger for states outside the fundamental gap. That  $\Omega_{S_1}$  as obtained in the PbE approach is close compared to the full one despite these observations is due to the effect of the direct electron-hole interaction in the BSE Hamiltonian. This contribution, which is solely responsible for effective electron-hole binding, is with  $K^{\text{d}} = -5.49$  eV in the PbE case stronger by 0.26 eV. As discussed in Chapter 2.2.3, the direct terms containes the screened

Coulomb interaction  $W$ . Lack of screening from the inactive region then implies that the electron-hole pair is subject to a stronger, more bare-Coulomb like, electron-hole attraction. Such a stronger binding compensates at least to some extent (by  $\sim 50\%$ ) the relatively larger energy differences of the free transitions.

In the right panel of Figure 5.5, the energy level diagram is given for results in which the ev $GW$  method is used in all three cases. Generally ev $GW$  leads here to a larger quasiparticle gap (by 0.42 eV) and a larger  $\Omega_{S_1}$  (by 0.34 eV) than in  $G_0W_0$  (full approach). Interestingly, the difference in  $E_{\text{gap}}^{\text{QP}}$  upon PbE is here with 0.13 eV somewhat smaller. Compared with a larger reference value in the full ev $GW$  calculation, this reduces its relative error from 3.8 % to 2.7 %. For the first electron-hole excitation energy, the absolute deviation to the full calculation is with 0.12 eV almost identical to the  $G_0W_0$  case. Due to the larger reference value, we still find a slightly reduced relative deviation of 3.7 % compared to 4.4 %. As before, basis truncation has no noticeable impact on the energy level diagram.

#### 5.4.2 Prodan in water

The study on the DPP bicyclic ring with branched alkyl side chains in the previous section has indicated that the proper choice of the active region is important, that differences in the energy levels between full and PbE calculations are attributable to the lack of screening effects from the inactive region, and that basis truncation had a minimal effect. We will now turn to a different test system, to scrutinize if these findings are specific to the DPP system in which the active and inactive region were connected by a covalent bond. We will also present and analyze the use of PbE in the GW-BSE/MM scenario for quantum-quantum-classical embedding.

The system we have chosen for this study consists of prodan, a polarity-sensitive dye, solvated in bulk water, as shown in Figure 5.3(b). Up excitation of the  $S_1$  state absorption of a photon, the dipole moment

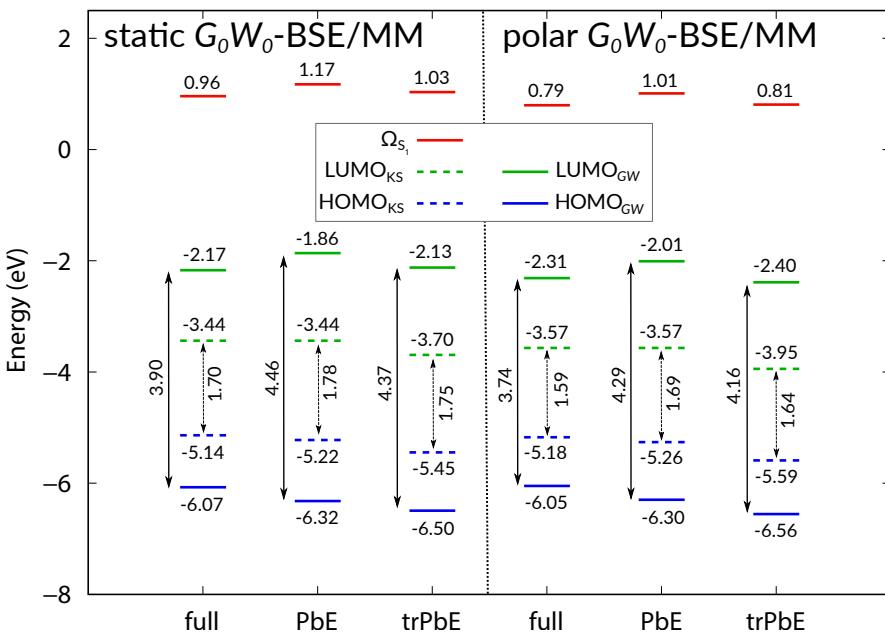
of prodan in the excited state is significantly increased compared to the ground state. In polar solvents, such as water, additional screening effects originate from the structural relaxation of the solvent molecules, which in turn affect the excited state properties of the solute, and lead to a significant reduction of the emission energy. In Ref. [190], this process was simulated by an iterative *GW*-BSE+MD procedure. We take a single snapshot of one of those trajectories, and perform static and polarizable  $G_0W_0$ -BSE/MM calculations on this structure.

To begin with, the system is partitioned in quantum and classical regions. We assign the prodan molecule and any water molecule whose center-of-mass is within 1 nm of the solute center-of-mass to the QM region. As a result, the QM region contains 199 atoms (prodan and 55 water molecules), which can be considerably challenging for full *GW*-BSE calculations. This QM region is then embedded in another 1 nm wide shell, which contains 630 water molecules treated on MM level.

We begin with the results of full  $G_0W_0$ -BSE calculations in a static MM environment, as shown in the left panel of the energy level diagram in Figure 5.6 and data summarized in Table 5.3. From the computational details in Table 5.1 it can be inferred that the computational cost for such a calculation is significant. We will discuss this in more detail in Section 5.5.1 below. In the  $G_0W_0$  step, we observe the typical lowering of the HOMO energy and increase of the LUMO energy with respect to the KS reference, resulting in a HOMO-LUMO gap of  $E_{\text{gap}}^{\text{QP}} = 3.90 \text{ eV}$  as compared to  $E_{\text{gap}}^{\text{KS}} = 1.70 \text{ eV}$ . For the PbE calculations, we now first split the QM region into an active region containing only the dye molecule and an inactive region containing the 55 water molecules (setup indicated as "dye" in Table 5.1 and Table 5.3). One can see from the dashed lines in Figure 5.6 that the PbE lowers the HOMO energy by 0.08 eV and yields a correspondingly increased HOMO-LUMO gap as the LUMO energy remains constant. From the isosurface plot as shown in Figure 5.7, one can see that the full KS HOMO is not exclusively localized on the dye alone but exhibits contributions from some close water molecules. With the restriction of the active region to prodan only, the neglect of these

**Table 5.3:** Results of full- $G_0W_0$ -BSE/MM, PbE- $G_0W_0$ -BSE/MM and trPbE- $G_0W_0$ -BSE/MM calculations (all static) for aqueous prodan: KS and QP HOMO and LUMO energies together with the individual contributions from the exchange-correlation potential  $V^{xc}$ , and the self-energy split in  $\Sigma^x$ , and  $\Sigma^c$ , according to  $\varepsilon^{QP} = \varepsilon^{KS} - V^{xc} + \Sigma^x + \Sigma^c$ , as well as KS and QP HOMO-LUMO gaps. The lowest electron-hole excitation energy from the respective BSE calculation  $\Omega_{S_1}$  is also split into the free transition energy ( $D$ ) and contributions from the exchange ( $K^x$ ) and direct ( $K^d$ ) terms of the electron-hole interaction. All energies in eV.

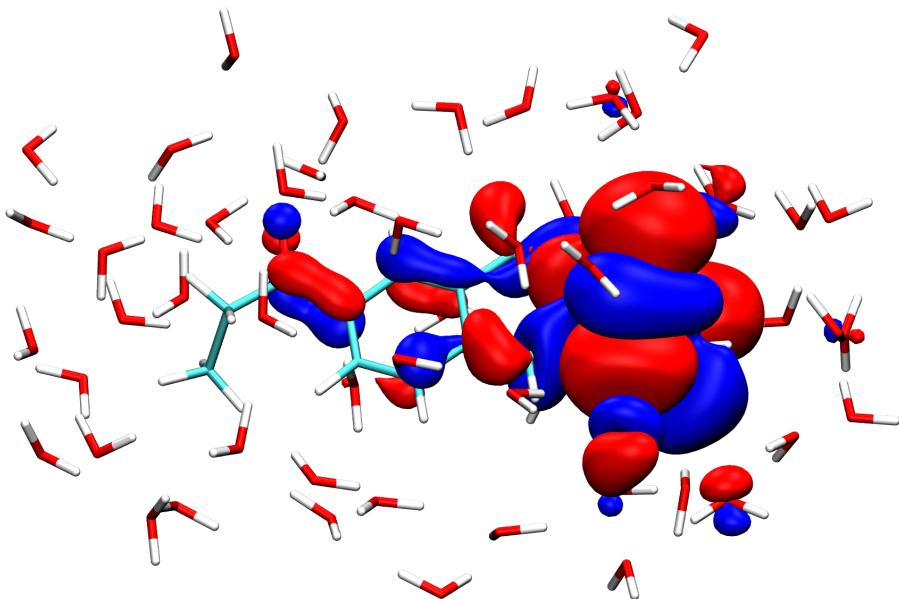
		dye			dye+water			
		full	PbE	$\Delta_{\text{full}}^{\text{PbE}}$	trPbE	$\Delta_{\text{PbE}}^{\text{trPbE}}$	PbE	$\Delta_{\text{full}}^{\text{PbE}}$
<b>DFT</b>	$\varepsilon_H^{\text{KS}}$	-5.14	-5.22	-0.08	-5.45	-0.23	-5.15	-0.01
	$\varepsilon_L^{\text{KS}}$	-3.44	-3.44	0.00	-3.70	-0.26	-3.44	0.00
	$E_{\text{gap}}^{\text{KS}}$	1.70	1.78	0.08	1.75	-0.03	1.71	0.01
	$V_H^{xc}$	-13.02	-12.79	0.23	-12.80	-0.01	-12.97	0.05
	$\Sigma_H^x$	-16.08	-15.86	0.22	-15.85	0.01	-16.04	0.04
	$\Sigma_H^c$	2.13	1.98	-0.15	2.00	0.02	2.03	-0.10
<b><math>G_0W_0</math></b>	$V_L^{xc}$	-11.31	-10.93	0.38	-10.96	-0.03	-11.08	0.23
	$\Sigma_L^x$	-7.61	-7.21	0.40	-7.27	-0.06	-7.37	0.24
	$\Sigma_L^c$	-2.44	-2.15	0.29	-2.12	0.03	-2.23	0.21
	$\varepsilon_H^{\text{QP}}$	-6.07	-6.32	-0.25	-6.50	-0.18	-6.19	-0.12
	$\varepsilon_L^{\text{QP}}$	-2.17	-1.86	-0.31	-2.13	-0.27	-1.96	0.21
	$E_{\text{gap}}^{\text{QP}}$	3.90	4.46	0.56	4.37	-0.09	4.23	0.33
<b>BSE</b>	$D$	4.14	4.76	0.62	4.71	-0.05	4.52	0.38
	$K^x$	0.11	0.18	0.07	0.18	0.00	0.15	0.04
	$K^d$	-3.30	-3.77	-0.47	-3.86	-0.09	-3.64	-0.34
	$\Omega_{S_1}$	0.96	1.17	0.21	1.03	-0.14	1.03	0.07



**Figure 5.6:** Energy level (in eV) diagram for aqueous propan showing HOMO (blue), LUMO (green), and lowest electron-hole excitation (red) energies on KS level (dashed) or  $GW$ -BSE level (solid), as resulting from full, PbE, and trPbE calculations, respectively. Left panel shows results from static  $G_0W_0$ -BSE/MM calculations, right panel from polar  $G_0W_0$ -BSE/MM calculations. Double-headed arrows additionally indicate the respective HOMO-LUMO gaps. See also Table 5.3 for details.

contributions changes the electronic state noticeably. Even though there is no covalent bond formed, this is very similar to the case of the DPP structure in the previous section, in which the HOMO extended over some carbon atoms of the side chain.

The quasiparticle gap in static PbE- $G_0W_0$ -BSE/MM results with 4.46 eV considerably larger than in the full reference (3.90 eV), compatible with the earlier observations about reduced screening effects for the embedding. Indeed, the differences in the respective contributions from  $V^{xc}$  and  $\Sigma^x$  nearly compensate and the difference in quasiparticle cor-



**Figure 5.7:** Isosurfaces of the KS HOMO (isovalue  $\pm 0.01 \text{ a}_\text{B}^{-3}$ ) of aqueous prodan chain as obtained from regular KS-DFT.

rections from full- to PbE- $G_0W_0$  mostly arises from effects in  $\Sigma^c$ . Basis set truncation removes 541 basis functions (see Table 5.1) in aqueous propan and has a more pronounced impact on the electronic levels. As can be seen in Figure 5.6 and Table 5.3, both HOMO and LUMO levels on KS level result lower in energy in trPbE as compared to the reference system (by 0.31 eV and 0.26 eV, respectively) and also to the respective PbE results. As a consequence of the almost constant shift,  $E_{\text{gap}}^{\text{KS}}$  is only 0.05 eV larger. On  $G_0W_0$  level, we note that the differences in the contributions in the quasiparticle corrections are relatively small and lead to only a small reduction of the quasiparticle gap by 0.09 eV. On BSE level, the calculated  $\Omega_{S_1}$  energy varies from 0.96 eV (full) via 1.17 eV (PbE) to 1.03 eV (trPbE). Inspection of the respective contributions to the BSE energy from the data in Table 5.3 reveals the same qualitative behaviour as discussed in Chapter 5.4.1: the lack of screening from the inactive region increases the contribution from free transitions  $D$ , which

is compensated to some extent by the for the same reason also increased electron-hole attraction in  $K^d$  ( $\sim 0.5$  eV).

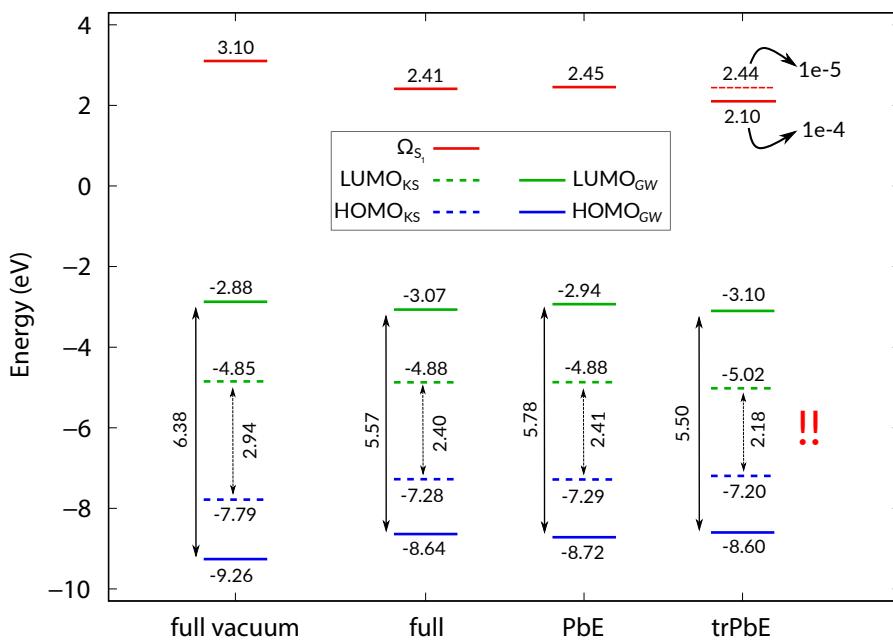
In the last two columns of Table 5.3, we additionally show results of PbE- $G_0W_0$ -BSE results in which the six water molecules with substantial contribution to the full system HOMO orbital are moved from the inactive to the active region. As can be seen, the change in the HOMO energy is very small compared to the full system, indicating that the more extended active region is a suitable choice. In this scenario, the differences in the quasiparticle energies for the HOMO and LUMO are smaller, but still amount to a  $E_{\text{gap}}^{\text{QP}}$  increased by 0.33 eV. Different choices for the active/inactive region splitting do not affect the qualitative observation that this difference is effectively only given by  $\Sigma^c$  contributions. In the BSE calculation, the lowest electron-hole excitation energy is obtained as 1.03 eV, only 0.07 eV higher than in the full  $G_0W_0$ -BSE calculation. Here, the stronger by -0.34 eV electron-hole interaction in the BSE kernel compensates the larger by 0.38 eV free transition contribution.

When polarizable  $GW$ -BSE/MM is employed, the respective calculations of the coupled system in the self-consistent reaction field of the MM environment become state dependent. What is shown as polar  $G_0W_0$ -BSE/MM in the right panel of Figure 5.6 are the energy levels in the final step of a self-consistent procedure to evaluate the total energy of the  $S_1$  excited state according to Eqs. (2.148)-(2.150). Note also that the PbE and trPbE results shown there are for the case in which only the prodan molecule is in the active region, as in the left panel. We refrain from analyzing the shown data in detail because even though the exact numbers are different, there are no fundamental differences in what has been observed for static  $G_0W_0$ -BSE. In other words, while the external potential is different in both cases, the intrinsic effects going from full to PbE or trPbE calculations are the same.

### 5.4.3 Benzene-TCNE dimer in water

The final system under consideration is a dimer of benzene and TCNE, as it is known to exhibit intermolecular charge-transfer type excitations in which the hole is predominantly located on benzene and the electron on TCNE molecule [191]. In a polar solvent, these CT excitations are massively lowered in energy compared to the vacuum case, and a proper treatment of the polarizable environment is essential. Not only is the type of electron-hole excitation different to the ones studied in Section 5.4.1 and Section 5.4.2 but also the localization on the contributing frontier orbitals. We scrutinize in the following if the observations regarding PbE- or trPbE-GW-BSE calculations made for the previous two test systems also hold in case of intermolecular excitations. To this end, we prepared first a dimer of benzene and TCNE molecules stacked with a separation of 3.7 Å. This initial structure was then solvated with water using `packmol`[192, 193]. From this solvated system, the benzene-TCNE dimer and the 26 closest water molecules have been selected for the following calculations. Note that we are only interested in the trends of effects from using PbE or trPbE here, so a more involved procedure to obtain relaxed atomic positions is not required for this purpose.

In Figure 5.8 we show the resulting energy level diagram for the different variants of  $G_0W_0$ -BSE calculations on this benzene-TCNE dimer, including a reference calculation for the dimer in vacuum, see also Table 5.4. In the vacuum reference, the HOMO-LUMO gap is increased in  $G_0W_0$  by 3.44 eV compared to the KS value, and the CT excitation energy is obtained as 3.01 eV. When embedded in water, the full  $G_0W_0$ -BSE calculation exhibits a reduced  $E_{\text{gap}}^{\text{KS}}$  (by 0.54 eV), which is predominantly caused by a shift in the HOMO level. The  $G_0W_0$  gap is even reduced compared to vacuum by 0.81 eV, but interestingly here we also observe a downward shift of the LUMO energy albeit only by 0.19 eV compared to the upward shift by 0.62 eV of the HOMO. From the isosurfaces of the HOMO and LUMO in the full- $G_0W_0$  calculation as shown in Figure 5.9 one can clearly see the donor-acceptor character of the dimer in the distribution of the frontier orbitals on the respective



**Figure 5.8:** Energy level (in eV) diagram for benzene-TCNE dimer solvated in water showing HOMO (blue), LUMO (green), and lowest electron-hole excitation (red) energies on KS level (dashed) or  $G_0W_0$ -BSE level (solid), as resulting from full, PbE, and trPBE calculations, respectively. Double-headed arrows additionally indicate the respective HOMO-LUMO gaps. See also Table 5.4 for details.

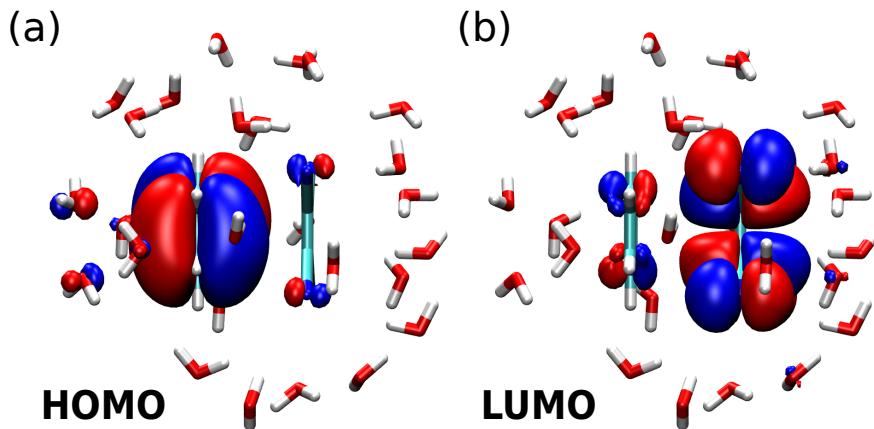
molecules, corroborating the notion that a transition from HOMO to LUMO is of charge-transfer character. Its CT excitation energy is obtained as 2.41 eV, lower by 0.69 eV than in the vacuum case. It is noteworthy that this lowering of the CT energy is caused mainly by the reduction of the contribution from the free transition term  $D$  (6.41 eV in vacuum vs. 5.59 eV in water) while the direct part of the electron-hole interaction is just reduced by 0.12 eV due to the additional screening from the environment.

From Figure 5.9 it can also be seen that both frontier orbitals are not completely localized on the respective molecules, as the HOMO has

**Table 5.4:** Results of full- $G_0W_0$ -BSE, PbE- $G_0W_0$ -BSE and trPbE- $G_0W_0$ -BSE calculations for a benzene-TCNE dimer in water: KS and QP HOMO and LUMO energies together with the individual contributions from the exchange-correlation potential  $V^{xc}$ , and the self-energy split in  $\Sigma^x$ , and  $\Sigma^c$ , according to  $\varepsilon^{QP} = \varepsilon^{KS} - V^{xc} + \Sigma^x + \Sigma^c$ , as well as KS and QP HOMO-LUMO gaps. The lowest electron-hole excitation energy from the respective BSE calculation  $\Omega_{S_1}$  is also split into the free transition energy ( $D$ ) and contributions from the exchange ( $K^x$ ) and direct ( $K^d$ ) terms of the electron-hole interaction. All energies in eV.

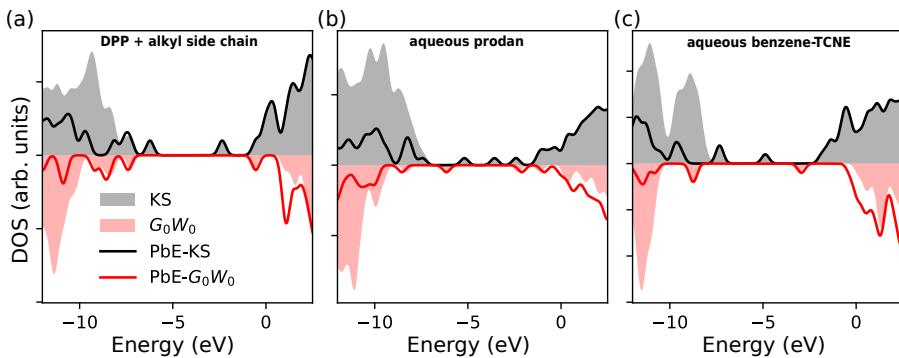
	vacuum	full	PbE	$\Delta_{\text{full}}^{\text{PbE}}$	trPbE	$\Delta_{\text{PbE}}^{\text{trPbE}}$	
<b>DFT</b>	$\varepsilon_H^{\text{KS}}$	-7.79	-7.28	-7.29	-0.01	-7.20	0.09
	$\varepsilon_L^{\text{KS}}$	-4.85	-4.88	-4.88	0.00	-5.02	-0.14
	$E_{\text{gap}}^{\text{KS}}$	2.94	2.40	2.41	0.01	2.18	-0.23
$G_0W_0$	$V_H^{xc}$	-10.40	-10.46	-10.40	0.06	-10.43	-0.03
	$\Sigma_H^x$	-11.97	-12.02	-11.95	0.07	-11.95	0.00
	$\Sigma_H^c$	0.10	0.20	0.12	-0.08	0.11	-0.01
	$V_L^{xc}$	-11.54	-11.56	-11.52	0.04	-11.53	-0.01
	$\Sigma_L^x$	-7.88	-7.93	-7.89	0.04	-7.92	-0.03
	$\Sigma_L^c$	-1.68	-1.82	-1.69	0.13	-1.70	-0.01
	$\varepsilon_H^{QP}$	-9.26	-8.64	-8.72	-0.08	-8.61	0.11
	$\varepsilon_L^{QP}$	-2.88	-3.07	-2.94	0.13	-3.10	-0.16
<b>BSE</b>	$E_{\text{gap}}^{QP}$	6.38	5.57	5.78	0.21	5.51	-0.27
	$D$	6.40	5.59	5.80	0.21	5.52	-0.28
	$K^x$	0.01	0.01	0.01	0.00	0.01	0.00
	$K^d$	-3.31	-3.18	-3.36	-0.18	-3.43	-0.07
	$\Omega_{S_1}$	3.10	2.42	2.45	0.03	2.10	-0.35

minimal contributions at the TCNE and some close water molecules. The LUMO shows a similar pattern. For the PbE and trPbE calculations, we choose only the dimer as the active region. As one can see in Figure 5.8 and Table 5.4, the associated neglect of the contributions of the water molecules to the HOMO has very little effect on the KS HOMO energy in



**Figure 5.9:** Isosurfaces of the (a) HOMO and (b) LUMO (isovalue  $\pm 0.01 \text{ a}_\text{B}^{-3}$ ) of the aqueous benzene-TCNE dimer as obtained from regular KS-DFT.

PbE, and concomitantly, the KS HOMO-LUMO gap. The quasiparticle energies in  $\text{PbE}-G_0W_0$  are 0.08 eV lower for the HOMO and 0.13 eV higher for the LUMO, compared to the full- $G_0W_0$  result, and as a consequence the gap is larger by 0.21 eV. As observed for the DPP system and aqueous prodan, the differences in the respective contributions to the quasiparticle corrections from  $V^{\text{xc}}$  and  $\Sigma^{\text{x}}$  mostly cancel out. The CT excitation energy in  $\text{PbE}-G_0W_0$ -BSE is with 2.45 eV only 0.03 eV larger than in the full calculation, as the comparatively stronger electron-hole interaction almost completely compensates the larger free quasiparticle transition energies. Upon truncation of the basis set, we find unexpectedly larger deviations in  $\text{trPbE}-G_0W_0$ . Already on Kohn-Sham level, we find that the HOMO energy is increased to -7.20 eV and the LUMO lowered to 5.02 eV, leading to  $E_{\text{gap}}$  being reduced by 0.23 eV. At quasiparticle level, the gap energy is reduced by 0.27 eV as compared to the untruncated PbE case, so that the quasiparticle gap is very close to the full- $G_0W_0$  result. This should be considered coincidental. Indeed, the energy of the CT excitation is also lower by 0.35 eV and is as a consequence 0.32 eV smaller than in the full calculation reference, as the electron-hole interaction remains under-screened.



**Figure 5.10:** Density of states (DOS) for the three systems studied in this work: (a) DPP with branched alkyl side chains, (b) aqueous propan, and (c) aqueous benzene-TCNE. Grey (red) shaded areas show the DOS as obtained by full KS ( $G_0W_0$ ) calculations, while the black (red) lines indicate the respective PbE-KS (PbE- $G_0W_0$ ) DOS. A Gaussian broadening with standard deviation 0.2 eV is used in all cases.

## 5.5 Discussion

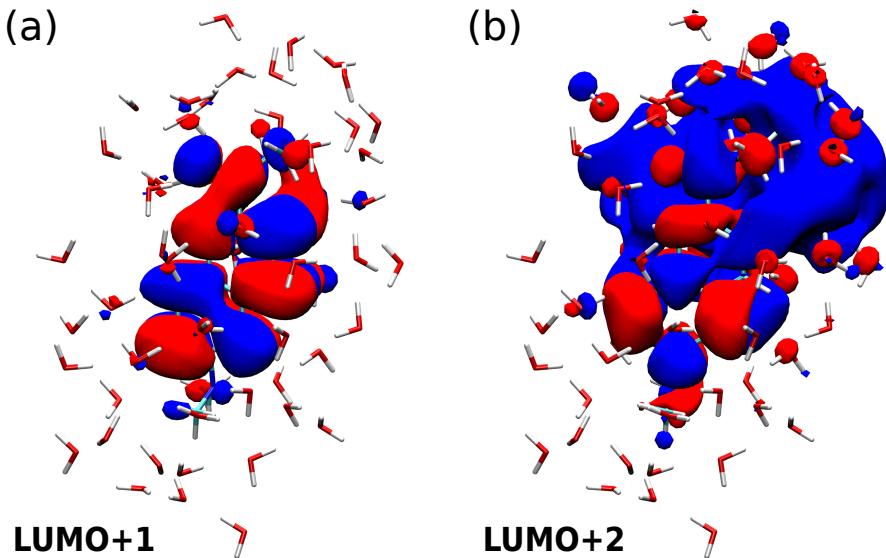
We begin the discussion of the analysis of the PbE and trPbE techniques in application to the *GW*-BSE methodology by taking a broader view on the obtained quasiparticle energies than just the frontier orbital and fundamental gap energies. To this end we show in Figure 5.10 a comparison between the density of states (DOS) for the three systems studied in this work as obtained from full and PbE calculations on KS and  $G_0W_0$  levels. On KS level (grey shaded area: full; black line: PbE) one can see in all three cases the frontier orbital peaks are well reproduced, and by construction also the full range derived from unoccupied orbitals. For the energy region lower than the respective HOMO energy, one can generally observe the presence of fewer states, as expected. Especially for the two water-solvated systems a significant part of the full DOS is removed by the embedding. When comparing the same data obtained on  $G_0W_0$  level of theory, one can see the small deviations of the frontier orbital energies in PbE- $G_0W_0$  with respect to the full calculation. For lower energy occupied and higher energy unoccupied levels, the comparison

**Table 5.5:** Results of full- $G_0W_0$ -BSE/MM, PbE- $G_0W_0$ -BSE/MM calculations (all polar) for aqueous prodan: KS and QP LUMO+1 and LUMO+2 energies together with the individual contributions from the exchange-correlation potential  $V^{xc}$ , and the self-energy split in  $\Sigma^x$ , and  $\Sigma^c$ , according to  $\varepsilon^{QP} = \varepsilon^{KS} - V^{xc} + \Sigma^x + \Sigma^c$ . All energies in eV.

	LUMO+1			LUMO+2		
	full	PbE	$\Delta_{full}^{PbE}$	full	PbE	$\Delta_{full}^{PbE}$
$\varepsilon^{KS}$	-2.32	-2.32	0.00	-1.06	-1.06	0.00
$V^{xc}$	-10.72	-10.47	0.25	-8.19	-5.91	2.29
$\Sigma^x$	-6.76	-6.52	0.24	-4.53	-3.11	1.43
$\Sigma^c$	-2.66	-1.58	0.24	-2.23	-1.58	0.65
$\varepsilon^{QP}$	-1.02	-0.79	0.23	0.37	0.16	-0.21

is not so straightforward. When one compares PbE- $G_0W_0$  to PbE-KS, one can see similarities in the broad shape of the DOS, but also that not all orbital energies experience the same quasiparticle corrections. This seems to affect the virtual DOS above the LUMO more significantly. As one can see in PbE- $G_0W_0$  results in Figure 5.10(a) and (c) in particular, there are peaks in the DOS below those the full calculation DOS. This is an indication that the QP corrections for these levels, whose KS reference energy is the same in full- and PbE-KS, are less pronounced when an embedded calculation is performed. This is different to the observation that quasiparticle corrections are generally stronger for the frontier orbitals due to the lack of screening.

To elucidate we consider in more detail the differences between the PbE and full calculation split among the different contributions to the QP energies as done for HOMO and LUMO in Table 5.3, now for LUMO+1 and LUMO+2 from the final step in the polarizable  $G_0W_0$ -BSE/MM calculations. The results are summarized in Table 5.5. For the LUMO+1, we find qualitatively the same behavior as for the LUMO as discussed in Section 5.4.2: the contributions from  $V^{xc}$  and  $\Sigma^x$  nearly cancel out, and



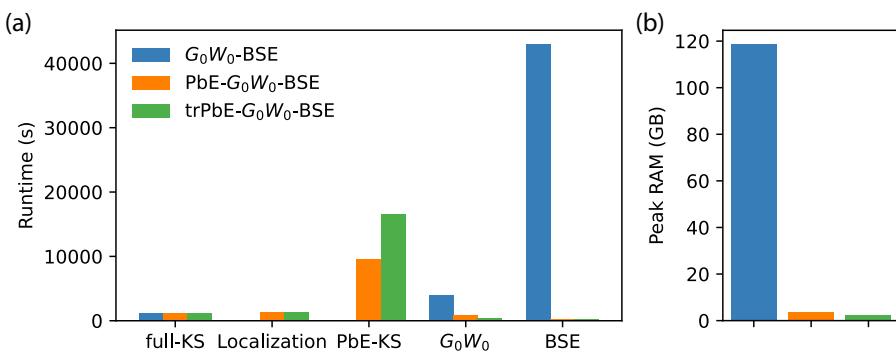
**Figure 5.11:** Isosurfaces of the (a) LUMO+1 and (b) LUMO+2 (isovalue  $\pm 0.01 \text{ a}_\text{B}^{-3}$ ) of the aqueous prodan as obtained from regular KS-DFT.

the too weak screening in PbE leads to a 0.23 eV higher quasiparticle energy as compared to the full calculation. For LUMO+2, the same does not hold: the difference in contributions from  $V^{xc}$  and  $\Sigma^x$  is significant, and contributes with -0.86 eV to the difference in quasiparticle energies. The contribution from the difference  $\Sigma^x$  is with 0.65 eV positive, consistent with the argument of too weak screening. Its magnitude is however much bigger than for LUMO and LUMO+1. In total, we find a quasiparticle energy that is *lower* by 0.21 eV in PbE as compared to the full calculation.

Inspecting the nature of the respective states might provide an indication of the origin of this different behavior. In Figure 5.11 we shows isosurfaces of LUMO+1 and LUMO+2, respectively. LUMO+1 (like LUMO) is predominantly localized on the dye molecule itself, with only small contributions from some close water molecules. In contrast, LUMO+2 is markedly different and extends to a large amount onto water molecules, that is into the inactive region from the point of view of

projection-based-embedding. It stands to reason that in such a situation the removal of the occupied states of the inactive region from the expression for the exchange part of the self-energy in Eq. (5.8) introduces additional deviations.

### 5.5.1 Embedding cost



**Figure 5.12:** Computational costs of the different calculation steps in full-, PbE-, and trPbE- $G_0W_0$ -BSE calculations for the aqueous prodan system from Section 5.4.2. (a) Runtime (in s) on 28 threads of an Intel(R) Xeon(R) Gold 5120 CPU @ 2.20GHz. (b) Peak RAM consumption in GB.

In addition to the quality of the results of the different PbE- $G_0W_0$ -BSE calculations compared to the full calculations, we consider the respective computational costs using the aqueous prodan system from Section 5.4.2 as an example. As can be seen from Table 5.1, large savings can be expected in the RPA steps (reduction of the number of transitions to less than 20 %), and in the BSE solution (reduction of the product basis size to about 3 %). Figure 5.12(a) shows the runtime of the respective calculation steps. For all of the three variants, the underlying DFT calculation on the full system is performed with ORCA in about 1100 s. Localizing all 336 occupied orbitals with the unitary optimization requires in VOTCA-XTP around 1300 s, and performing the PbE-KS around 9600 s. Note that the internal DFT implementation

in VOTCA-XTP is not optimized for performance and intended for development purposes instead. As expected, the most significant saving in computational time is the actual  $G_0W_0$  and BSE steps of the procedure. For the former, the reduced number of transitions in the RPA, combined with a lower number of states for which quasiparticle corrections have to be determined, reduced the time from 3900 s to around 800 s. Note that the absolute cost of the  $G_0W_0$  step is, in fact, reasonably small (only a factor  $\sim 3.5$  in the full calculation) compared to the DFT step, due to the use of the Plasmon-Pole Model for the frequency-dependence of the self-energy, which requires the explicit evaluation of the microscopic dielectric function only for two frequencies, see also Section 2.2.2. The most dramatic compute time reduction is observed for the BSE step. Here PbE leads to a reduction from 43000 s to just 270 s due to the massively reduced dimension of the product basis. Basis set truncation allows for another reduction of runtimes in  $G_0W_0$  and BSE by a factor of 2, respectively. Apart from runtimes, the size of accessible systems for the GW-BSE approach is often also limited by the peak memory consumption, at least in the implementation in VOTCA-XTP. From the data shown in Figure 5.12(b), it is clear that the full and PbE approaches differ vastly in peak memory consumption. Embedding and the subsequent reduction in the in-memory storage of three-center Coulomb integrals after contraction with molecular orbitals requires only 3.8 GB of RAM compared to 118.7 GB. The truncated basis has a smaller effect on top of this (2.5 GB). This clearly shows that PbE techniques can remove some of the computational bottlenecks of GW-BSE calculations, at the price of some deviations in the obtained results due to the lack of screening contributions from the inactive region.

## 5.6 Summary

In this chapter, we have introduced and scrutinized projection-based-embedding techniques of GW-BSE calculations. Based on the analysis of the three test systems DPP ring with branched alkyl side chains,

aqueous prodan, and an aqueous benzene-TCNE dimer, we could see that PbE can offer significant computational gains, making larger systems accessible to the many-body Green's functions based methodology. We have demonstrated that it can also be directly incorporated in quantum-classical embedding (*GW*-BSE/MM) schemes. We also found that the agreement with full calculations depends on the choice of the active region and is subject to effects from the neglect of screening contributions from the inactive electrons in the *GW* steps, which leads generally to an increased quasiparticle HOMO-LUMO gap. It was also noted that the lack of screening is in part compensated in the BSE as it manifests itself in an increased electron-hole attraction, so that deviations from full results for the electron-hole excitation energy are on the order of 0.1 eV for the different types of excitations studied here. It should be noted that all excitations are near-gap excitations, and it can not be guaranteed that the same quantitative agreement will hold for higher-energy excitations. Finally, we have seen that additional truncation of the basis set can reduce the computational costs by a factor of two with respect to full-basis PbE, but results appear to be sensitive to the chosen threshold values for removing basis functions.

# Chapter 6

## Excitonics at a DPP2Py2T-PCBM interface

With the embedding strategies as developed in Chapter 5, we now turn towards the investigation of the excitonic properties at the interfaces between donor and acceptor in the mixed DPP2Py2T-PCBM[60] morphologies introduced in Chapter 4.

### 6.1 Introduction

For any organic solar-cell active layer, its ability to efficiently generate free charges from an initial photoexcitation of the absorbing material is one of the key effects to control and to optimize. As described earlier, an important part of the process is the formation of an intermediate bimolecular charge-transfer excited state. While an important factor, the conversion from an either DPP2Py2T or PCBM[60] localized excitation (LE) to a CT excitation depends not only on the alignment of the energies alone (which in turn can be very sensitive to structural details). Instead, the conversion can be considered an activated tunneling process, and can be described by an effective transition rate. In this chapter, we will work towards the estimation of different conversion rates based on PbE-GW-BSE calculations, for some representative structures. For this specific system, this investigation is motivated by the experimental observation

that the population of CT states at the interface is sensitive to whether the donor or acceptor material is initially excited, and the energy of the CT excitation.

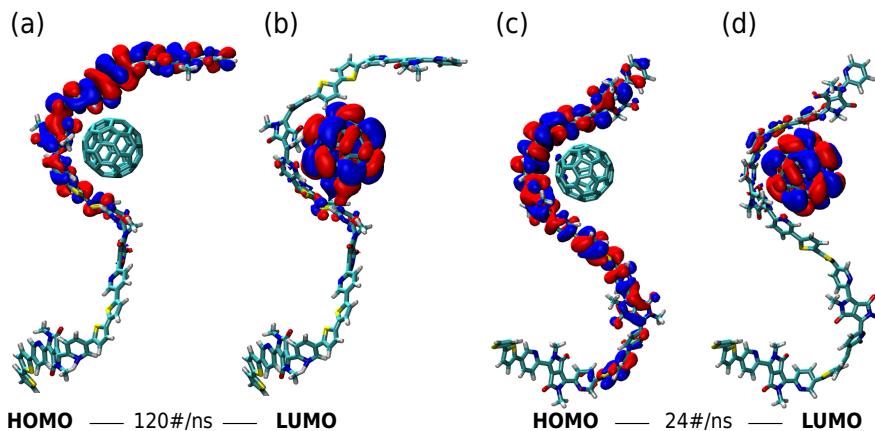
Dynamics at the DPP2Py2T-PCBM[60] interface cannot only be understood from an analysis of the vertical energy levels alone. Focusing in particular on the conversion from initially excited localized excitations on either DPP2Py2T or PCBM[60] to bi-molecular charge-transfer excitations, it is attractive to consider them by an effective transfer rate [167, 168, 194]. In the non-adiabatic high temperature or activated crossing limit, Marcus theory [195, 196] yields the rate of electronic excitation transfer between an initial LE and final CT state as

$$\omega_{\text{LE-CT}} = \frac{2\pi}{\hbar} \frac{|J_{\text{LE-CT}}|^2}{\sqrt{4\pi\lambda_{\text{LE-CT}}k_B T}} \exp\left[-\frac{(\Delta E_{\text{LE-CT}} - \lambda_{\text{LE-CT}})^2}{4\lambda_{\text{LE-CT}}k_B T}\right], \quad (6.1)$$

where  $\Delta E_{\text{LE-CT}}$  is the adiabatic energy difference, and  $\lambda_{\text{LE-CT}}$  the reorganization energy. The expression also contains the *non-adiabatic coupling* element,  $J_{\text{LE-CT}}$ . In principle, it should be possible to evaluate all three quantities that enter the Marcus rate from electronic structure methods, including the (PbE-)GW-BSE methods developed in this work. In the following, this is being considered for two representative DPP2Py2T-PCBM[60] structures. In Chapter 4 we discussed how different processing techniques could yield different morphologies and briefly stated that these morphologies could have an impact on the excitation energies of the DPP2Py2T-PCBM blend. We pick two of those structures: the structure after solvent removal at a rate of 120 #/ns and after a solvent removal rate of 24 #/ns, c.f. structures I<sub>a</sub> and I<sub>b</sub> in Figure 4.3.

## 6.2 Computational details

Out of these large-scale structures, we select first the DPP2Py2T and PCBM[60] with closest center-of-mass distance from each fragment to C<sub>60</sub> center-of-mass as they are expected to exhibit the strongest interactions between donor and acceptor molecule and potentially the energetically



**Figure 6.1:** Isosurfaces of the full KS HOMO and LUMO (isovales  $\pm 0.005 \text{ a}_\text{B}^{-3}$ ) for two DPP2Py2T-PCBM[60] structures taken from the final results of I<sub>a</sub> and I<sub>b</sub> solvent evaporation approaches in Chapter 4 with removal rates of 120 #/ns and 24 #/ns, respectively.

lowest lying CT excitation energies. As these dimer structures still contain in total 650 atoms, we remove in a first postprocessing step the side chains from the DPP2Py2T (barring one methyl group) and replace the PCBM[60] with a plain C<sub>60</sub> at the same position and orientation. This step reduces the number of atoms to 262, and the corresponding structures are shown in Figure 6.1. To keep the subsequent calculations on the electronic structure tractable, we use then Stuttgart-Dresden effective core potentials [197–199] with an uncontracted version of the associated optimized basis set, augmented by polarization functions from the 6-311G\*\* basis [200, 201]. We have shown in [134] that this basis set offers a good balance between computational cost of GW-BSE calculations and accuracy. In all calculations, the PBE0 hybrid functiona [149] has been used in steps involving KS-DFT. With these choices, the number of electrons in the full system is 898, so 449 occupied levels, with a total basis set size of 4692. A full GW-BSE calculation would therefore have in the most extreme case  $N_{\text{RPA}} = 1905107$  transitions in the screening calculations, with the BSE product space  $N_{\text{BSE}} = 996004$ , and the

computational cost is enormous. Inspecting the orbital distributions of HOMO and LUMO for the two selected structures in Figure 6.1 obtained from full KS-DFT calculations reveals that in both cases the HOMO is not delocalized over the whole DPP2Py2T chain. Particularly in the 120 #/ns conformation, the HOMO remains localized on the part close to the PCBM[60] molecule. In the 24 #/ns structure, the HOMO extends somewhat further, but note that the isovalue chosen for the plot is relatively low. Based on these observations, we proceed with a PbE setup, in which the active region comprises the end 2-mer of DPP2Py2T and the C<sub>60</sub>, in total 262 atoms. In this setting, there are 644 active electrons, so 322 occupied states. Still, the number of transitions in the RPA is 140714 and the BSE product space dimension is 207368. All PbE-G<sub>0</sub>W<sub>0</sub>-BSE calculations were performed on the Snellius supercomputer of the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-4183.

### 6.2.1 Non-adiabatic LE-CT couplings

While these calculations will provide crucial information about the different vertical excitation energies, more quantities need to be calculated for the determination of the rates in Eq. (6.1). The first of these quantities are the non-adiabatic coupling elements between LE and CT excitons. Electronic states such as the excitations obtained from the solutions of the BSE are adiabatic states |Φ<sub>i</sub><sup>ad</sup>⟩ of some electronic Hamiltonian H<sub>el</sub>. Corresponding diabatic states |Φ<sub>a</sub><sup>diabatic</sup>⟩ can be obtained via a unitary transformation

$$|\Phi_a^{\text{diabatic}}\rangle = \sum_{j=1}^{N_{\text{states}}} U_{aj} |\Phi_j\rangle. \quad (6.2)$$

With this, the adiabatic form of the electronic Hamiltonian with adiabatic energies ε<sub>i</sub>, i.e., ⟨Φ<sub>i</sub>|H<sub>el</sub>|Φ<sub>j</sub>⟩ = ε<sub>j</sub>δ<sub>ij</sub> is transformed into the diabatic form

$$H_{ab}^{\text{diabatic}} = \langle \Phi_a^{\text{diabatic}} | H_{\text{el}} | \Phi_b^{\text{diabatic}} \rangle = \sum_{ij} U_{ia} \langle \Phi_i | H_{\text{el}} | \Phi_j \rangle U_{bj}. \quad (6.3)$$

Considering specifically a two-state problem ( $N_{\text{states}} = 2$ ), the transformation can be written explicitly as a rotation

$$\underline{\mathbf{U}} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (6.4)$$

and the diabatic Hamiltonian as

$$\underline{\mathbf{H}}^{\text{diabatic}} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (6.5)$$

Its off-diagonal elements

$$J_{ab} = \frac{1}{2} \sin(2\theta)(\varepsilon_2 - \varepsilon_1) \quad (6.6)$$

are then the non-adiabatic couplings between the two diabatic states. In practice, the optimal  $\theta$  of this specific unitary transformation matrix  $\underline{\mathbf{U}}$  is determined by extremalizing some function  $f(\underline{\mathbf{U}})$ . Here, we use the *Fragment Charge Difference* diabatization, in which the diabatic states are eigenstates of the so-called fragment charge difference (FCD) matrix [202], based on the definition of donor (D) and acceptor (A) fragments, as  $\Delta Q_{ij} = Q_{ij}(D) - Q_{ij}(A)$ . The fragment charges are typically obtained from a population analysis of the individual adiabatic densities and the transition density between them. Again, for a two-state model, the coupling is given explicit by

$$J_{ab}^{\text{FCD}} = \frac{|\Delta Q_{12}|(\varepsilon_2 - \varepsilon_1)}{((\Delta Q_{11} - \Delta Q_{22})^2 + 4\Delta Q_{12}^2)^{1/2}}. \quad (6.7)$$

### 6.2.2 Reorganization energy estimates

Evaluation of the Marcus rates further requires calculation on the respective reorganization energies  $\lambda_{\text{LE-CT}}$ . Within the Marcus picture,  $\lambda_{\text{LE}_x-\text{CT}} = E_{\text{CT}}(\text{LE}_x) - E_{\text{CT}}(\text{CT})$ , where  $x = \text{P}, \text{F}$  (P: polymer, F:

fullerene) and  $E_a(G)$  represents the total energy of state  $a$  at geometry of state  $G$ .

As such, this would require the cumbersome optimization of the dimer structures in the respective CT and LE states. Instead, we approximate the energies from monomer calculations, such that

$$E_{\text{CT}}(\text{LE}_P) = E_P^+(\text{P}^*) + E_F^-(\text{F}^0) \quad (6.8)$$

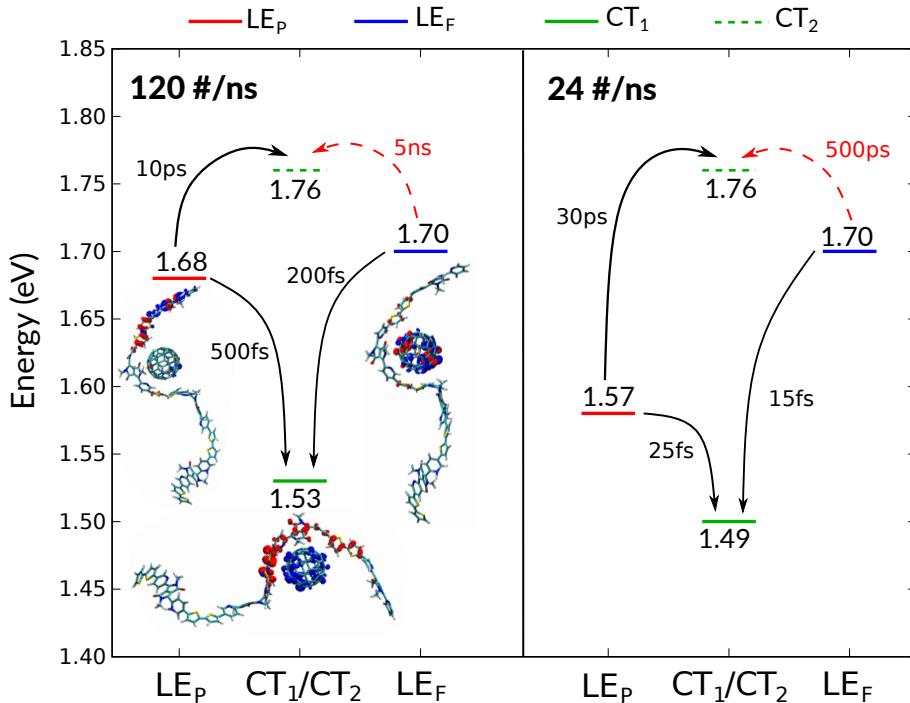
$$E_{\text{CT}}(\text{LE}_F) = E_P^+(\text{P}^0) + E_F^-(\text{F}^*) \quad (6.9)$$

$$E_{\text{CT}}(\text{CT}) = E_P^+(\text{P}^+) + E_F^-(\text{F}^-) \quad (6.10)$$

where the superscripts refer to the state of the monomers (0: ground state, +: cation, -: anion, \*: excited). The total energy calculations and geometry optimizations in this step are performed using (time-dependent) DFT with the def2-tzvp basis set and PBE0 functional. For DPP2Py2T we perform these calculations on both a single repeat unit and two repeat units but not longer due to the large computational cost involved in optimizations for larger structures and the fact that the hole part of the excited states remains localized on a smaller subpart of the full structure as in Figure 6.1. In similar spirit, we determine the vertical to adiabatic energy relaxations of the excited states,  $\Lambda_a = E_a(0) - E_a(A)$ .

### 6.3 Interface excitonics

The vertical excitation energies as resulting from PbE- $G_0W_0$ -BSE calculations on the two chosen DPP2Py2T-PCBM[60] interface structures are shown in Figure 6.2 and listed in Table 6.1. We focus specifically on the two lowest excitations localized on DPP2Py2T ( $\text{LE}_P$ ) and PCBM[60] ( $\text{LE}_F$ ), respectively, and the two lowest energy excitations with near integer CT character ( $\text{CT}_1$  and  $\text{CT}_2$ ). There are some noteworthy similarities between both structures. The energies of the localized excitation on the PCBM[60] as well as that of the second CT state  $\text{CT}_2$  are identical with 1.70 eV and 1.76 eV, respectively. Even the lowest-lying  $\text{CT}_1$  energy is only different by 0.04 eV. It is possible that this is due to fact that both



**Figure 6.2:** Energy levels for LE<sub>P</sub>, LE<sub>F</sub> along with CT type excitations CT<sub>1</sub> and CT<sub>2</sub> in the two DPP2Py2T-PCBM[60] interface structures as obtained from PbE-G<sub>0</sub>W<sub>0</sub>-BSE calculations. Estimates for the timescales of the transitions are indicated with the respective arrows. For the 120 #/ns structure, isosurfaces ( $\pm 0.0008 \text{ ea}_\text{B}^{-3}$ ) of the electron difference densities are shown as insets. Red color indicates hole (lack of electron) density while blue indicates electron density.

structures, as shown in Figure 6.1 exhibit very similar local arrangements of the polymer and fullerene. The fullerene part itself is also not subject to significant distortions of the internal molecular geometry, much in contrast to the DPP2Py2T chain. For the LE<sub>P</sub> energy, we find therefore slightly larger difference of 0.11 eV between the 120 #/ns and 24 #/ns structures.

**Table 6.1:** System specific parameters of the kinetic model for LE to CT excitation conversion at the DPP2Py2T-PCBM[60] interfaces from snapshots out of the two solvent evaporation simulations from Chapter 4.

	120#/ns	24#/ns		
$\Omega_P/eV$	1.68	1.57		
$\Omega_F/eV$	1.70	1.70		
$\Omega_{CT_1}/eV$	1.53	1.49		
$\Omega_{CT_2}/eV$	1.76	1.76		
$J_{LE_P-CT_1}^2/(eV)^2$	$4.8 \cdot 10^{-5}$	$7.9 \cdot 10^{-4}$		
$J_{LE_P-CT_2}^2/(eV)^2$	$3.0 \cdot 10^{-5}$	$2.8 \cdot 10^{-3}$		
$J_{LE_F-CT_1}^2/(eV)^2$	$1.8 \cdot 10^{-4}$	$2.1 \cdot 10^{-3}$		
$J_{LE_F-CT_2}^2/(eV)^2$	$1.5 \cdot 10^{-7}$	$2.6 \cdot 10^{-6}$		
	1-mer	2-mer		
$\lambda_{LE_P-CT}/eV$	0.13	0.10		
$\lambda_{LE_F-CT}/eV$	0.27	0.14		
$\Lambda_{LE_P}/eV$	0.09	0.18		
$\Lambda_{LE_F}/eV$	0.18	0.18		
$\Lambda_{LE_{CT}}/eV$	0.22	0.19		
$\omega_{LE_P-CT_1}/s^{-1}$	$4.2 \cdot 10^{11}$	$2.0 \cdot 10^{12}$	$1.9 \cdot 10^{13}$	$4.2 \cdot 10^{13}$
$\omega_{LE_P-CT_2}/s^{-1}$	$9.0 \cdot 10^{11}$	$7.4 \cdot 10^{10}$	$8.2 \cdot 10^{12}$	$3.2 \cdot 10^{10}$
$\omega_{LE_F-CT_1}/s^{-1}$	$5.1 \cdot 10^{12}$	$5.3 \cdot 10^{12}$	$6.7 \cdot 10^{13}$	$7.1 \cdot 10^{13}$
$\omega_{LE_F-CT_2}/s^{-1}$	$2.7 \cdot 10^8$	$1.8 \cdot 10^8$	$3.2 \cdot 10^9$	$2.1 \cdot 10^9$
$\frac{\omega_{LE_F-CT_2}}{\omega_{LE_F-CT_1}}$	$3.4 \cdot 10^{-5}$	$5.2 \cdot 10^{-5}$	$4.8 \cdot 10^{-5}$	$2.9 \cdot 10^{-5}$
$\frac{\omega_{LE_F-CT_2}}{\omega_{LE_P-CT_2}}$	$3.0 \cdot 10^{-4}$	$2.5 \cdot 10^{-3}$	$3.9 \cdot 10^{-4}$	$6.6 \cdot 10^{-2}$

It is tempting to infer from these two alignments of the excitation energy levels something regarding the respective driving forces of dynamical processes, such as the conversion from the LE to CT states. However, as mentioned above, the vertical excitation energies alone do not provide an

exhaustive description of such processes. In particular, while the energies are the same or similar in the two structures, it does not necessarily imply that the excited states themselves are. In Table 6.1, we also list the squared excitonic LE-CT coupling elements  $J^2$  as determined from the FCD method in Section 6.2.1. Here, some more differences between the 120 #/ns and 24 #/ns structures are apparent. The squared LE-CT couplings are systematically smaller in the former case, somewhere between one and two orders of magnitude. However, in both structures, the squared coupling between the LE on the fullerene to the higher of the two CT excited states is three orders of magnitude smaller than that to CT<sub>1</sub>. As the Marcus rate in Eq. (6.1) is linear in  $J^2$ , this is a remarkable difference. Still, it is not clear if this is also a *significant* difference, as the effects of the exponential term can be dominant.

To model the missing data in the Marcus rate, we need to determine the respective reorganization energies, as well as adiabatic excitation energies for the energy difference  $\Delta E_{\text{LE-CT}}$ . As explained in Section 6.2, this would require, in principle, the optimization of the DPP2Py2T-PCBM[60] geometry in the environment for the four different excited states based on PbE- $G_0W_0$ -BSE. As this is practically impossible, we resort to optimizations of isolated DPP2Py2T and PCBM[60], with a mix of DFT and TD-DFT, and the resulting reorganization energies  $\lambda$  and corrections to the vertical energies  $\Lambda$  are given in Table 6.1. Two approximations are considered: one in which the respective optimization of the polymer are performed on a single repeat unit (1-mer) and one for two repeat units (2-mer). From a visual inspection of the HOMO distribution in Figure 6.1(a) and (c), the latter seems to be the more appropriate one. In this case, we find similar reorganization energies for the transition from LE<sub>P</sub> to CT (0.10 eV) and from LE<sub>F</sub> to CT (0.14 eV). In comparison, the 1-mer calculations predict about a factor of two difference. Similarly, we find that in the 2-mer approximation the adiabatic corrections to the vertical excitation energies are almost identical for all three types of excitations amounting to 0.18 eV for the two LE, and 0.19 eV for the CT excitations. As a consequence, the adiabatic corrections hardly influence the driving forces  $\Delta E_{\text{LE-CT}}$  in Eq. (6.1).

Bringing everything together yields then values for the rates (evaluated at room temperature) as given in Table 6.1 and the corresponding time scales as indicated in Figure 6.2 for the 2-mer models. We find that the rates for transition from the two LE to the CT<sub>1</sub> are somewhat higher in the 24 #/ns structure. What both structures have in common is that the rate for transfer from LE<sub>F</sub> to CT<sub>2</sub>,  $\omega_{\text{LE}_F-\text{CT}_2}$ , is consistently smaller than any other rates. We also show the ratio of the rates for transfer from the LE<sub>F</sub> exciton to CT<sub>2</sub> and CT<sub>1</sub> in Table 6.1, finding that there is a 4-5 orders of magnitude difference. Notably, there is also roughly three orders of magnitude difference compared to the rate from LE<sub>P</sub> to CT<sub>2</sub>. In terms of rough timescales, the results indicate that the transition to CT<sub>1</sub> occurs on a scale of tens (hundreds) of fs in the 24 #/ns (120 #/ns) structure, and the transition from LE<sub>P</sub> to CT<sub>2</sub> on a scale of tens of ps. The transition to CT<sub>2</sub> from LE<sub>F</sub> in comparison roughly takes place on ns timescale., much slower than any other process. This difference is, in the data we have obtained, mostly determined from the low value of  $J_{\text{LE}_F-\text{CT}_2}^2$ .

## 6.4 Summary & Discussion

While the observation from our rate calculations based on PbE- $G_0W_0$ -BSE calculations appear to correspond qualitatively to what has been observed experimentally for similar DPP2Py2T-PCBM[60] morphologies (slow or suppressed population of higher CT states from fullerene as compared to from the polymer excitation), our results should be considered as somewhat preliminary. Related to the PbE-GW-BSE methodology it is positive to note that its use allowed us in this chapter to tackle localized and charge-transfer type excitations in large-scale complex molecular systems with these many-body Green's function based methods. This allows not only the calculation of the vertical excitation energies but also in a straightforward manner the determination of the non-adiabatic coupling elements. In the evaluation of the Marcus rates, the comparatively drastic approximations required for the computation of reorganization

energies and adiabatic energy corrections should be considered with care, as already discussed.

Besides these points, several additional aspects need to be addressed for obtaining a fully reliable picture of the interface excitonics:

- Frontier orbitals of fullerene in vacuum are highly degenerate, leading to a rather dense spectrum of LE in the low-energy range. While the molecular environment in DPP2Py2T-PCBM[60] lifts some of that degeneracy, it is possible that the assumption using a single  $LE_F$  in the above is not sufficient and instead at least some near-degenerate expression for the non-adiabatic couplings might have to be used.
- The PbE-*GW*-BSE calculations presented in this chapter were performed on selected dimer structures in vacuum. Previous studies on *GW*-BSE/MM level [134, 135, 203] have shown that in particular the energy of CT excitations is very sensitive to polarization effects in the molecular environment, with energy stabilizations reported in the range of 0.5-1.0 eV. Performing polarizable PbE-*GW*-BSE/MM calculations as discussed in Chapter 5 are computationally demanding because of the required self-consistency procedures. Also, the proper parametrization of the Thole model for the classical response is far from trivial for polymeric materials, due to the strong conformational effects. Instead, we have performed some preliminary classical embedding calculations, in which the materials around the DPP2Py2T-PCBM[60] dimer was filled with a polarizable lattice [152, 204], similar to a polarizable continuum model with dielectric constant  $\epsilon_r = 3.5$ . However, early findings do not point towards energetic stabilizations of the CT states of the kind observed before, but much smaller on the order of 0.1 eV. This small effect could be a result of the details of the lattice calculation, or the fact that the CT excited states in the DPP2Py2T-PCBM[60] morphology with the local structures as shown in Figure 6.1 are significantly less polarizable than mixed small molecule donor and fullerene morphologies. It should also be noted that the energies of the localized

excitons are close to the experimentally measured ones [27] and a large polarization-induced downshift of the CT excitation energies would result in very low CT energies at around 0.5 eV, which may seem unreasonable. Smaller stabilization of 0.1-0.2 eV appear more reasonable, instead, as they would also bring down the  $\text{CT}_2$  energy below the two LE energies.

- In this chapter, we only considered the two DPP2Py2T and PCBM[60] arrangements from the final structures obtained by the morphology prediction presented in Chapter 4 with the closest distance between the donor and acceptor. As especially the CT energies are known to be sensitive to donor-acceptor distance and orientation, other arrangements of DPP2Py2T and PCBM[60] should also be investigated. Such a study would potentially reveal effects of structural disorder on the interface exciton energies, which may have consequences on the calculated driving forces.
- While the above points can, at least in principle, be addressed by performing more, computationally demanding, PbE-*GW*-BSE/MM calculations, one aspect is more fundamentally difficult. The timescales extracted from our calculations have revealed fast processes occurring in tens or hundreds of fs, and slow processes on the scale of ns. While it can be argued for the fast processes that they are faster than most of the changes in the molecular geometries, this is certainly not the case for the slow processes. To elucidate, at least calculations for different frames along the MD trajectories have to be analyzed, for instance with a view of time correlations in the excitation energies and coupling elements, to deduce a proper approximation.

# Chapter 7

## Summary & Conclusion

This thesis documents in general development and application of different molecular simulation methods to the study of excitons and their conversion dynamics in bulk polymer-fullerene morphologies.

Specifically, several advancements on both classical and quantum levels have been made to arrive at the results which are described through Chapter 3 to Chapter 6. The overall objective was to scrutinize the existing methodologies and develop them further to gain a nanoscale understanding of dynamical processes involving excitons in such complex composite systems. As an interesting prototypical system, we have chosen a series of DPP-based polymers with variable number of thiophene substituents as donor material combined with a fullerene-based acceptor, as experimental evidence has shown that these combinations allow for efficient formation of charge-transfer excitations across the interface. Investigations of this kind typically require multiscale modeling approaches which can not only describe but predict the structure of the material at atomistic resolution on the one hand, and the quantum-mechanical processes involving electrons on the other hand. In most cases, this puts extremely high demands on the accuracy of the respective methods as well as their applicability to realistic system sizes. It has become evident that, even when enormous computational resources are available, abstractions of the problem and approximations are inevitable.

From the point of view of the simulation of the atomistically-resolved morphology, the objective was to obtain a *representative* structure of a DPP2PymT-PCBM[60] bulk heterojunction as it is prepared from solution as in experiment. As generic force-fields do not cover the intricate details of interactions in particular of the complex polymer material with different molecular architectures, the first step was to develop and refine the parametrization of custom-made fully-atomistic force-field for the polymer DPP2PymT using a combination of quantum and classical methods, as presented in Chapter 3. The quality of the resulting force-field was validated by studying the behavior of polymers with different side chains in good and poor solvents, as well as by an analysis of intra-molecular dynamics such as the segmental mobility of the units connected by softer dihedrals.

With the custom-made force-field for DPP2PymT at hand, we turned to the actual simulation of a representative bulk heterojunction morphology in Chapter 4. Here, choices had to be made regarding the details of system already with an eye on the subsequently planned investigations of the excitonics: To be close to the typical experiments, one should consider roughly 50 repeat units, a size which is completely infeasible to treat with electronic structure methods. An inspection of the dependence of the calculated electron-hole excitation in vacuum optimized oligomers of DPP2PymT with variable number of substituents has revealed that the excitation in these extended structures is typically restricted to 2-3 repeat units and also energetically stabilized at  $\sim 2\text{ eV}$ . With disorder likely to lead to further localization, it was decided to use oligomers with 4 repeat units of DPP2PymT for a Molecular Dynamics simulation of a mixture with PCBM[60] mixture. Here, we put particular emphasis on investigating the effects of different simulation protocols on the interfacial properties, such as local orientations of polymer and fullerene. To mimic experimental procedures, an iterative procedure for solvent removal was employed, and the final structures obtained for different evaporation rates were shown to be different from one resulting from annealing. Annealing the system resulted in larger PCBM[60] concentrations around

the DPP unit of the polymer backbone while solvent removal allowed larger concentrations around the thiophene units of the polymer.

Even with the choice of the relatively small scale DPP2Py2T-PCBM[60] mixtures, the typical system sizes needed to study charge-transfer type excitations in a dimer of a single polymer and fullerene molecule exceed what is being able to be treated at the level of excited state electronic structure theory methods. In this thesis, we have investigated the use of many-body Green's functions theory in the *GW* approximation and the Bethe–Salpeter equation (BSE) as method of choice, motivated by the fact that it has provided capable of describing both localized and charge-transfer type excitations well in smaller molecular dimers. To avoid the computational bottlenecks involved in the full calculation of complex dimer configurations of DPP2Py2T-PCBM[60] in the atomistic environment, we have implemented and scrutinized quantum-quantum and quantum-quantum-classical embedding methods as discussed in Chapter 5. We paid special attention to the role the selection of active atoms played in projection-based-embedded DFT-*GW*-BSE calculations, and how missing (with respect to the full calculation) screening effects in the electronic self-energy affect the frontier orbitals and lowest excitation energies in different prototypical systems. It was found that for the lowest ( $\text{HOMO} \rightarrow \text{LUMO}$ ) transitions, the increased HOMO-LUMO gap due to the lack of screening contribution in the *GW* step is partially compensated by a stronger electron-hole interaction in the BSE. With a view on the application to DPP2Py2T-PCBM[60] systems, this error seemed similar enough between localized and CT excitons to allow an assessment of relative excitation energies, as the computational gain achieved using the embedding technique was significant. This gain was seen not only in the runtime but also the memory consumption of the *GW* step of the calculation, which typically is the biggest bottleneck in the VOTCA-XTP implementation of *GW*-BSE. Development of this method allowed for the computation of large DPP2Py2T-PCBM[60] systems on a level of desired theory which was seemingly impossible otherwise.

For the actual treatment of selected DPP2Py2T-PCBM[60] dimers from the evaporation results in Chapter 4, first a suitable choice for the active region for the embedding calculation was based on an inspection of the respective frontier orbitals in a full DFT reference calculations. Then, it was seen in Chapter 6 that the energies of the respective monomer localized excitations in the dimers were hardly affected by the presence of the other molecules. For the newly formed CT excitations, we observe the lowest one ( $CT_1$ ) energetically below the localized excitons, while a second one ( $CT_2$ ) is obtained at slightly higher energy. To gain an understanding of the dynamical conversion processes at the interface beyond simplified energy level arguments, we have also studied the calculation of the rates for converting the localized excitations to the two CT type excitations within the Marcus picture. This required estimates of reorganization and adiabatic relaxation energies, as well as the determination of the excitonic coupling between the two respective states. This exploratory work indicates that while the conversion of the localized excitations to the lowest CT exciton is fast (on the order of femtoseconds) and mostly independent of initial excitation of donor or acceptor. The conversion to  $CT_2$  however, is much slower in comparison, on the order of nanoseconds, even slower than typical timescales for structural relaxations. Further investigation of this with different MD snapshots will be required to get an in-depth understanding. While the charge transfer dynamics for such systems have been computed for the first time, there is a qualitative agreement with experimental which reported similar slow dynamics to from localized exciton on PCBM[60] to  $CT_2$ . A look at the electron and hole densities in the CT state it was evident that the electron density was concentrated on the PCBM[60] while the hole density was concentrated primarily on the thiophene units which validates our earlier assumption in Chapter 4 where we stated that having PCBM[60] units closer to thiophene units is desirable as they are the significant contributors to the CT state.

At this point, we can revisit the questions asked in Chapter 1, answering them based on the results obtained in this thesis.

1. **Can we model the different excited states, in particular the charge transfer states, in bulk heterojunction solar cells?** Yes, it is possible to use a simple model(s) that incorporates the effect of various morphologies on charge-transfer rates of a bulk heterojunction solar cell. However, this requires careful consideration and interpretation of multiscale structural and electronic dynamics to couple the two simulation scales.
2. **How do processing techniques have an influence on the local morphology of the blend?** The processing techniques do influence the local morphology of the blend, in particular as to how the electron rich thiophenes are oriented with respect to the fullerene acceptors. This is seen in Chapter 4 where systems generated from annealing have entirely different orientation and morphology when compared to systems generated from solvent removal. In Chapter 6, one can see that when CT states are formed the area of electron deficiency is concentrated primarily on the thiophenes indicating them as the hotspots for charge transfer. Hence, a variation in morphology will have an impact on the charge transfer states. An in-depth analysis of this is proposed as a future study.
3. **Are current methods to model electronically excited states sufficiently scalable to systems of size which can accurately represent realistic materials?** This does not entail a straightforward answer as we do have methods in place that can accurately represent morphologies, or electronic band gaps and excitation energies. The challenge is that while these methods exist it is very difficult to couple them to gain accurate representations given that these methods operate on very different length and time scales. However, as described in this thesis there are ways to couple excitonic calculations with electronic density based theory embedded in a classical envelope. There have been approximations made along the way to be able to achieve this, nevertheless such a combination of models yield results that qualitatively and to a great extent

quantitatively match experimental outcomes and can securely be used for further studies of large molecules.

While the work reported in this thesis provides crucial developments of methods towards an in principle complete picture for studying dynamical processes of excitons in large-scale molecular systems, such as bulk-heterojunction solar cells, there is still plenty of room for future work. This concerns both the application of the current state-of-the-art to the particular (type of) material system at the focus of the present research and further methodological developments.

Among the work on the excitonics at the DPP2Py2T-PCBM[60] interface, the points mentioned at the end of Chapter 6 are worth some attention. These include, among other, the extension of the present approach to more frames from the MD trajectories and to other structural motifs (larger distances, different (not  $\pi$ - $\pi$  stacked) orientations). It would also be interesting to perform a larger set of hybrid quantum-classical QM/QM/MM calculations, to assess in more depth the preliminary findings that environment effects do not as significantly affect the CT excitation energies in this system than in other small molecule donor-acceptor type systems.

From the point of view of methodological developments and implementations in VOTCA-XTP, some attention should be paid to automatize some of the steps in the setup of the embedding calculations, e.g., the automated selection of the active region based on some analysis of the frontier orbitals in a reference calculation. Further computational gains can be expected from a more fine-grained truncation scheme, in which the actual reduced basis is constructed only from individual shells of contributing basis functions, not all functions of the atoms. Finally, the most important aspect is the work on the inclusion of screening effects of the inactive density region to the correlation part of the self-energy in the embedded *GW*-BSE methods. Besides an adaptation of the work in Ref. [114], the polarizable *GW*-BSE/MM framework as presented in Chapters 2 and 5 might provide an effective means to account for these screening effects in the classical polarizable force-field. This would

require the atoms of the inactive region to be treated at *hybrid resolution* with both quantum (the electron density) and classical (the polarization response) degrees of freedom. Developing an approach like this would also entail careful considerations of how the two not spatially separated degrees of freedom interact, and how to consistently parametrize the classical model for the specific chemistry. This is left as future work.

# Acknowledgement

I owe a gratitude to a lot people who have been a part of this journey for putting me where I am today. I have tried my best to mention as many names as possible. Sincere apologies if I have forgotten someone.

First and foremost I express my gratitude to my supervisors Dr. A Lyulin and Dr. B. Baumeier for giving me the opportunity to be a part of this wonderful journey. Without your guidance and mentorship this project would never reached this stage. Special mention to Bjoern for being my daily supervisor and more than just a mentor. You have been a great teacher and I saw myself grow from a young researcher to a matured skilled person under your guidance. The experiences in daily interaction with you were enlightening not only on a academic level but also on a personal level as well which helped me grow into the person I am today.

I express my appreciation to the members of the thesis committee for their valuable time and thoughtful contributions. Your insights have greatly enhanced the quality of this work.

My deepest gratitude and love for my family who have been my support system throughout the process and will continue to be for the rest of my life. Mumma Papa you have been a great motivator paving the path for success for me. My brother, Piyush who is a friend, confidant and partner in crime in many ways. Also my uncle who has never failed to amuse us with his lively and cheerful spirit. We never miss to discuss your antiques pulled off at various airports across the world. Finally not forgetting Mansi for being a part of my life for almost a decade now and walking this path alongside me to help me grow as a human being.

During my years at CASA and SMB I met quite some people who have been with me partly or completely in the journey. I would like to thank the secretaries of both groups, Enna and Helmi who have stood up to assist whenever any sort of bureaucratic help was needed. My friends and colleagues from the Baumeier Research Group who have been involved in valuable discussions and interactions along the way. Gianluca, Onur, Ruben, Wouter, Zhong, Haoxiang, Jens, Pranav and Youri, all of you have taught me something valuable and have shared memorable moments both on an intellectual and a casual level. I will cherish them for a very long long time. Of course my other colleagues who I have interacted and shared evening drinks with. Antonio, Vi, Fransesco, Nicky, Harshit, René, Koondi, Sanjana, Willem, Maikel, Robert. Also not to forget colleagues from the SMB group, Nikos and Marteen who have been together in a few conferences. Your camaraderie has made this academic experience richer and more memorable.

Finally my Acknowledgement is incomplete without mentioning my friends who have been a pillar and my family away from home. Abhijeet, Arnab and Sambit thank you for all the trips, discussions and a treasure trove of lifelong memories. Our cooking and crypto sessions together will be a part of my memory forever. Getting through the covid times would have been a whole challenge altogether without you guys. The special discussion of railport and sea station will be etched in my memory forever. Gratitude to my friends Rachna and Pulkit who have been a part of many evening sessions and card games along the way. I am sure there are more beautiful memories and moments coming my way with friends like all of you. I consider myself very fortunate to have lived in a house with hospitable and fun housemates, Adrien, Olivia, Daria and Mijali.

I express my gratitude for all mentioned above and anyone I have missed who have contributed to a beautiful living experience in the Netherlands for me.

## About the author

Vivek Sundaram was born in Patna, a beautiful but largely crowded city in India. Having completed his high school in Siliguri and Kota, he decided to pursue his love for science as a Materials Science Engineer in the Indian Institute of Technology, Kanpur. He completed his Bachelor's in 2017 and enrolled in a dual degree program to complete his Master's in Computational Materials Science in 2018. In August 2018, he had the opportunity to work as a doctoral candidate in the Eindhoven University of Technology alongside dr. B. Baumeier in the Mathematics and Computer Science department and prof. dr. A.V. Lyulin in the Applied Physics department.

Since May 2023, Vivek has been working as a Financial Risk Model Developer at de Volksbank, Utrecht.

# Bibliography

- [1] R. Voss, *Statistics netherlands*, (2023) <https://www.cbs.nl/en-gb/news/2023/10/renewable-electricity-share-up-by-20-percent-in-2022>.
- [2] W. Shockley, H. J. Queisser, and R. Ell, “Detailed balance limit of efficiency of pn junction solar cells”, *J. App. Phys.* **32**, 510 (1961).
- [3] S. Rühle, “Tabulated values of the Shockley-Queisser limit for single junction solar cells”, *Sol Energy* **130**, 139 (2016).
- [4] L. C. Andreani, A. Bozzola, P. Kowalczewski, M. Liscidini, and L. Redorici, “Silicon solar cells: toward the efficiency limits”, *Adv. Phys.: X* **4**, 1548305 (2019).
- [5] W. Li, K. H. Hendriks, M. M. Wienk, and R. A. Janssen, “Diketopyrrolopyrrole Polymers for Organic Solar Cells”, *Acc. Chem. Res.* **49**, 78 (2016).
- [6] J. Zhao, A. Wang, M. A. Green, and F. Ferrazza, “19.8% efficient "honeycomb" textured multicrystalline and 24.4% monocrystalline silicon solar cells”, *Appl. Phys. Lett.* **73**, 1991 (1998).
- [7] A. Shah, P. Torres, R. Tscharner, N. Wyrsch, and H. Keppner, “Photovoltaic technology: The case for thin-film solar cells”, *Science* **285**, 692 (1999).
- [8] G. Macfarlane, T. McLean, J. Quarrington, and V. Roberts, “Exciton and phonon effects in the absorption spectra of germanium and silicon”, *J. Phys. Chem. Solids* **8**, 388 (1959).
- [9] M. A. Green, “Improved value for the silicon free exciton binding energy”, *AIP Adv.* **3**, 112104 (2013).

- [10] H. Sai, T. Matsui, and K. Matsubara, “Stabilized 14.0%-efficient triple-junction thin-film silicon solar cell”, *Appl. Phys. Lett.* **109**, 183506 (2016).
- [11] W. Qarony, M. I. Hossain, M. K. Hossain, U. M. Jalal, A. Haque, A. R. Saad, and Y. H. Tsang, “Efficient amorphous silicon solar cells: characterization, optimization, and optical loss analysis”, *Results Phys.* **7**, 4287 (2017).
- [12] C. Duan, F. Huang, and Y. Cao, “Recent development of push-pull conjugated polymers for bulk-heterojunction photovoltaics: Rational design and fine tailoring of molecular structures”, *J. Mater. Chem.* **22**, 10416 (2012).
- [13] B. C. Thompson and J. M. Fréchet, “Polymer-fullerene composite solar cells”, *Angewandte Chemie - International Edition* **47**, 58 (2008).
- [14] M. Scharber and N. Sariciftci, “Efficiency of bulk-heterojunction organic solar cells”, *Prog. Poly. Sci.* **38**, Topical issue on Conductive Polymers, 1929 (2013).
- [15] A. A. Bakulin, A. Rao, V. G. Pavelyev, P. H. M. van Loosdrecht, M. S. Pshenichnikov, D. Niedzialek, J. Cornil, D. Beljonne, and R. H. Friend, “The role of driving energy and delocalized states for charge separation in organic semiconductors”, *Science* **335**, 1340 (2012).
- [16] A. E. Jailaubekov, A. P. Willard, J. R. Tritsch, W.-L. Chan, N. Sai, R. Gearba, L. G. Kaake, K. J. Williams, K. Leung, P. J. Rossky, and X.-Y. Zhu, “Hot charge-transfer excitons set the time limit for charge separation at donor/acceptor interfaces in organic photovoltaics”, *Nat. Mater.* **12**, 66 (2013).
- [17] D. Caruso and A. Troisi, “Long-range exciton dissociation in organic solar cells”, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13498 (2012).

- [18] Y. Liang, D. Feng, Y. Wu, S. T. Tsai, G. Li, C. Ray, and L. Yu, “Highly efficient solar cell polymers developed via fine-tuning of structural and electronic properties”, *J. Am. Chem. Soc.* **131**, 7792 (2009).
- [19] M. S. Su, C. Y. Kuo, M. C. Yuan, U. S. Jeng, C. J. Su, and K. H. Wei, “Improving device efficiency of polymer/fullerene bulk heterojunction solar cells through enhanced crystallinity and reduced grain boundaries induced by solvent additives”, *Adv. Mater.* **23**, 3315 (2011).
- [20] M. M. Wienk, M. Turbiez, J. Gilot, and R. A. Janssen, “Narrow-bandgap diketo-pyrrolo-pyrrole polymer solar cells: The effect of processing on the performance”, *Adv. Mater.* **20**, 2556 (2008).
- [21] D. Chandran and K.-S. Lee, “Diketopyrrolopyrrole: A Versatile Building Block for Organic Photovoltaic Materials”, *Macromol. Res.* **21**, 272 (2013).
- [22] D. M. Chapin, C. S. Fuller, and G. L. Pearson, “A new silicon p-n junction photocell for converting solar radiation into electrical power [3]”, *J. App. Phys.* **25**, 676 (1954).
- [23] H. Choi, S. J. Ko, T. Kim, P. O. Morin, B. Walker, B. H. Lee, M. Leclerc, J. Y. Kim, and A. J. Heeger, “Small-bandgap polymer solar cells with unprecedented short-circuit current density and high fill factor”, *Adv. Mater.* **27**, 3318 (2015).
- [24] R. S. Ashraf, I. Meager, M. Nikolka, M. Kirkus, M. Planells, B. C. Schroeder, S. Holliday, M. Hurhangee, C. B. Nielsen, H. Sirringhaus, and I. McCulloch, “Chalcogenophene comonomer comparison in small band gap diketopyrrolopyrrole-based conjugated polymers for high-performing field-effect transistors and organic solar cells”, *J. Am. Chem. Soc.* **137**, 1314 (2015).
- [25] K. H. Hendriks, G. H. Heintges, V. S. Gevaerts, M. M. Wienk, and R. A. Janssen, “High-molecular-weight regular alternating diketopyrrolopyrrole-based terpolymers for efficient organic solar cells”, *Angewandte Chemie - International Edition* **52**, 8341 (2013).

- [26] P. Heinrichová, J. Pospíšil, S. Stříteský, M. Vala, M. Weiter, P. Toman, D. Rais, J. Pfleger, M. Vondráček, D. Šimek, L. Fekete, P. Horáková, L. Dokládalová, L. Kubáč, and I. Kratochvílová, “Diketopyrrolopyrrole-Based Organic Solar Cells Functionality: The Role of Orbital Energy and Crystallinity”, *J. Phys. Chem. C* **123**, 11447 (2019).
- [27] K. H. Hendriks, A. S. Wijpkema, J. J. Van Franeker, M. M. Wienk, and R. A. Janssen, “Dichotomous Role of Exciting the Donor or the Acceptor on Charge Generation in Organic Solar Cells”, *J. Am. Chem. Soc.* **138**, 10026 (2016).
- [28] J. Peet, J. Kim, N. Coates, et al., “Efficiency enhancement in low-bandgap polymer solar cells by processing with alkane dithiols”, *Nat. Mater.* **6**, 497 (2007).
- [29] E. Schrödinger, “An Undulatory Theory of the Mechanics of Atoms and Molecules”, en, *Phys. Rev.* **28**, 1049 (1926).
- [30] S. Blinder, in *Introduction to quantum mechanics* (Elsevier, Jan. 2021), pp. 129–149.
- [31] M. Born and R. Oppenheimer, “Zur Quantentheorie der Moleküle”, *Annalen der Physik* **389**, 457 (1927).
- [32] V. May and O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems*, en, 1st ed. (Wiley, Feb. 2011).
- [33] J. C. Slater, “The Self Consistent Field and the Structure of Atoms”, en, *Phys. Rev.* **32**, 339 (1928).
- [34] B. Bransden and C. Joachain, in *Physics of atoms and molecules* (Longman Scientific and Technical, Jan. 1990), pp. 116–123.
- [35] D. R. Hartree, “The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion”, en, *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 111 (1928).
- [36] P. Hohenberg and W. Kohn, “Inhomogeneous Electron Gas”, en, *Phys. Rev.* **136**, B864 (1964).

- [37] W. Kohn and L. J. Sham, “Self-Consistent Equations Including Exchange and Correlation Effects”, en, Phys. Rev. **140**, A1133 (1965).
- [38] W. Kohn, “Nobel lecture: electronic structure of matter—wave functions and density functionals”, Rev. Mod. Phys. **71**, 1253 (1999).
- [39] K. Burke, “Perspective on density functional theory”, en, J. Chem. Phys. **136**, 150901 (2012).
- [40] F. Malet, A. Mirtschink, K. J. H. Giesbertz, L. O. Wagner, and P. Gori-Giorgi, “Exchange–correlation functionals from the strong interaction limit of DFT: applications to model chemical systems”, en, Phys. Chem. Chem. Phys. **16**, 14551 (2014).
- [41] J. P. Perdew, “Density-functional approximation for the correlation energy of the inhomogeneous electron gas”, en, Phys. Rev. B **33**, 8822 (1986).
- [42] J. P. Perdew, “Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas”, en, Phys. Rev. B **34**, 7406 (1986).
- [43] C. Lee, W. Yang, and R. G. Parr, “Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density”, en, Phys. Rev. B **37**, 785 (1988).
- [44] A. D. Becke, “Density-functional thermochemistry. III. The role of exact exchange”, en, J. Chem. Phys. **98**, 5648 (1993).
- [45] S. Obara and A. Saika, “Efficient recursive computation of molecular integrals over Cartesian Gaussian functions”, en, J. Chem. Phys. **84**, 3963 (1986).
- [46] F. Weigend and R. Ahlrichs, “Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy”, en, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
- [47] R. M. Martin, *Electronic structure: basic theory and practical methods* (Cambridge University Press, 2004).

- [48] Z.-L. Cai, K. Sendt, and J. R. Reimers, “Failure of density-functional theory and time-dependent density-functional theory for large extended  $\pi$  systems”, *J. Chem. Phys.* **117**, 5543 (2002).
- [49] L. Hedin, “New method for calculating the one-particle green’s function with application to the electron-gas problem”, *Phys. Rev.* **139**, A796 (1965).
- [50] L. Hedin and S. Lundqvist, *Effects of electron-electron and electron-phonon interactions on the one-electron states of solids*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich, Vol. 23, Solid State Physics (Academic Press, 1970), pp. 1–181.
- [51] L. J. Sham and T. M. Rice, “Many-particle derivation of the effective-mass equation for the wannier exciton”, *Phys. Rev.* **144**, 708 (1966).
- [52] D. Golze, M. Dvorak, and P. Rinke, “The GW Compendium: A Practical Guide to Theoretical Photoemission Spectroscopy”, *Frontiers in Chem.* **7**, 377 (2019).
- [53] M. S. Hybertsen and S. G. Louie, “Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies”, *Phys. Rev. B* **34**, 5390 (1986).
- [54] M. S. Hybertsen and S. G. Louie, “First-principles theory of quasiparticles: calculation of band gaps in semiconductors and insulators”, *Phys. Rev. Lett.* **55**, 1418 (1985).
- [55] M. van Schilfgaarde, T. Kotani, and S. Falaleev, “Quasiparticle self-consistent *GW* theory”, *Phys. Rev. Lett.* **96**, 226402 (2006).
- [56] A. Stan, N. E. Dahlen, and R. Van Leeuwen, “Levels of self-consistency in the *GW* approximation”, *J. Chem. Phys.* **130**, 114105 (2009).
- [57] R. W. Godby, M. Schlüter, and L. J. Sham, “Self-energy operators and exchange-correlation potentials in semiconductors”, *Phys. Rev. B* **37**, 10159 (1988).

- [58] P. E. Bloechl, S. Lebègue, M. Alouani, P. E. Bloechl, S. Lebègue, M. Alouani, B. Arnaud, B. Arnaud, and B. Arnaud, “Implementation of an all-electron GW approximation based on the projector augmented wave method without plasmon pole approximation: Application to Si, SiC, AlAs, InAs, NaH, and KH”, Phys. Rev. B - Condens. Matter Mater. Phys. **67**, 155208 (2003).
- [59] M. M. Rieger, L. Steinbeck, I. D. White, H. N. Rojas, and R. W. Godby, “GW space-time method for the self-energy of large systems”, Comp. Phys. Comm. **117**, 211 (1999).
- [60] C. Friedrich, S. Blügel, and A. Schindlmayr, “Efficient implementation of the GW approximation within the all-electron FLAPW method”, Phys. Rev. B - Condens. Matter and Mater. Phys. **81**, 125102 (2010).
- [61] 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”, J. Chem. Theory Comput. **11**, 5665 (2015).
- [62] P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, “Cubic scaling GW: Towards fast quasiparticle calculations”, Phys. Rev. B **94**, 165109 (2016).
- [63] L. Hedin, “On correlation effects in electron spectroscopies and the GW approximation”, J. Phys. Condens. Matter **11**, R489 (1999).
- [64] M. J. Van Setten, F. Weigend, and F. Evers, “The GW-method for quantum chemistry applications: Theory and implementation”, J. Chem. Theory Comput. **9**, 232 (2013).
- [65] R. W. Godby and R. J. Needs, “Metal-insulator transition in Kohn-Sham theory and quasiparticle theory”, Phys. Rev. Lett. **62**, 1169 (1989).
- [66] D. R. Lide, *Handbook of chemistry and physics* (Chemical Rubber Company Press, Boca Raton, FL, 1998).

- [67] E. E. Salpeter and H. A. Bethe, “A relativistic equation for bound-state problems”, *Phys. Rev.* **84**, 1232 (1951).
- [68] N. Nakanishi, “A General Survey of the Theory of the Bethe-Salpeter Equation”, en, *Prog. Theor. Phys. Supp.* **43**, 1 (1969).
- [69] G. Onida, L. Reining, and A. Rubio, “Electronic excitations: density-functional versus many-body green’s-function approaches”, *Rev. Mod. Phys.* **74**, 601 (2002).
- [70] N. Marom, F. Caruso, X. Ren, O. T. Hofmann, T. Körzdörfer, J. R. Chelikowsky, A. Rubio, M. Scheffler, and P. Rinke, “Benchmark of *GW* methods for azabenzenes”, *Phys. Rev. B* **86**, 245127 (2012).
- [71] Y. Ma, M. Rohlfing, and C. Molteni, “Excited states of biological chromophores studied using many-body perturbation theory: effects of resonant-antiresonant coupling and dynamical screening”, *Phys. Rev. B* **80**, 241405 (2009).
- [72] X. Blase, C. Attaccalite, and V. Olevano, “First-principles *GW* calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications”, *Phys. Rev. B* **83**, 115103 (2011).
- [73] B. Baumeier, D. Andrienko, and M. Rohlfing, “Frenkel and charge-transfer excitations in donor–acceptor complexes from many-body green’s functions theory”, *J. Chem. Theory Comput.* **8**, 2790 (2012).
- [74] B. T. Sutcliffe and R. G. Woolley, “On the quantum theory of molecules”, *J. Chem. Phys.* **137**, 22A544 (2012).
- [75] S. A. Adcock and J. A. McCammon, “Molecular dynamics: survey of methods for simulating the activity of proteins”, *Chem. Rev.* **106**, 1589 (2006).
- [76] S. E. Feller, D. Yin, R. W. Pastor, and A. D. MacKerell, “Molecular dynamics simulation of unsaturated lipid bilayers at low hydration: parameterization and comparison with diffraction studies”, *Biophys. J.* **73**, 2269 (1997).

- [77] S. E. Feller and A. D. MacKerell, “An improved empirical potential energy function for molecular simulations of phospholipids”, *J. Phys. Chem. B* **104**, 7510 (2000).
- [78] W. D. Cornell et al., “A second generation force field for the simulation of proteins, nucleic acids, and organic molecules”, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [79] C. J. Dickson et al., “Lipid14: the amber lipid force field”, *J. Chem. Theory Comput.* **10**, 865 (2014).
- [80] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. V. Gunsteren, “A biomolecular force field based on the free enthalpy of hydration and solvation: the gromos force-field parameter sets 53a5 and 53a6”, *J. Comput. Chem.* **25**, 1656 (2004).
- [81] W. L. Jorgensen and J. Tirado-Rives, “The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin”, *J. Am. Chem. Soc.* **110**, 1657 (1988).
- [82] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, “Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids”, *J. Am. Chem. Soc.* **118**, 11225 (1996).
- [83] T. M. Reed and K. E. Gubbins, *Applied statistical mechanics: thermodynamic and transport properties of fluids* (McGraw-Hill, 1973).
- [84] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, Computational Science Series, Vol 1 (Academic Press, 2001).
- [85] F. Castro-Román, R. W. Benz, S. H. White, and D. J. Tobias, “Investigation of finite system-size effects in molecular dynamics simulations of lipid bilayers”, *J. Phys. Chem. B* **110**, 24157 (2006).
- [86] J. Horbach, W. Kob, K. Binder, and C. A. Angell, “Finite size effects in simulations of glass dynamics”, *Phys. Rev. E* **54**, R5897 (1996).

- [87] J. J. Salacuse, A. R. Denton, and P. A. Egelstaff, “Finite-size effects in molecular dynamics simulations: static structure factor and compressibility. i. theoretical method”, Phys. Rev. E **53**, 2382 (1996).
- [88] T. Darden, D. York, and L. Pedersen, “Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems”, J. Chem. Phys. **98**, 2384 (1993).
- [89] P. P. Ewald, “Die berechnung optischer und elektrostatischer gitterpotentiale”, Annalen der Physik (Berl.) **369**, 253 (1921).
- [90] W. M. Brown, A. Kohlmeyer, S. J. Plimpton, and A. N. Tharrington, “Implementing molecular dynamics on hybrid high-performance computers—particle-particle-particle-mesh”, Comp. Phys. Comm. **183**, 449 (2012).
- [91] S. De Leeuw, J. Perram, E. Smith, and R. J. Shipley, “Simulation of electrostatic systems in periodic boundary conditions. i. lattice sums and dielectric constants”, Proceedings of the Royal Society of London A **373**, 27 (1980).
- [92] S. De Leeuw, J. Perram, and E. Smith, “Simulation of electrostatic systems in periodic boundary conditions. ii. equivalence of boundary conditions”, Proceedings of the Royal Society of London A **373**, 57 (1980).
- [93] S. De Leeuw, J. Perram, and E. Smith, “Simulation of electrostatic systems in periodic boundary conditions. iii. further theory and applications”, Proceedings of the Royal Society of London A **388**, 177 (1983).
- [94] K. Feng and M. Qin, *Symplectic geometric algorithms for hamiltonian systems* (Springer, 2010).
- [95] L. Verlet, “Computer “experiments” on classical fluids. i. thermodynamical properties of lennard-jones molecules”, Phys. Rev. **159**, 98 (1967).

- [96] E. Hairer, C. Lubich, and G. Wanner, “Geometric numerical integration illustrated by the störmer/verlet method”, *Acta Numerica* **12**, 399 (2003).
- [97] M. E. Tuckerman, *Statistical mechanics: theory and molecular simulation*, 1st ed. (Oxford University Press, Oxford ; New York, 2010), p. 720.
- [98] H. C. Andersen, “Molecular dynamics simulations at constant pressure and/or temperature”, *J. Chem. Phys.* **72**, 2384 (1980).
- [99] W. Hoover, “Canonical dynamics: equilibrium phase-space distributions”, *Phys. Rev. A* **31**, 1695 (1985).
- [100] S. Nosé, “A molecular dynamics method for simulations in the canonical ensemble”, *Molecular Phys.* **52**, 255 (1984).
- [101] S. Nosé, “A unified formulation of the constant temperature molecular dynamics methods”, *J. Chem. Phys.* **81**, 511 (1984).
- [102] F. S. Michael P. Allen, “A thermostat for molecular dynamics of complex fluids”, *Molecular Simulation* **33**, 21 (2007).
- [103] J. Tomasi, B. Mennucci, and R. Cammi, *Quantum mechanical continuum solvation models*, 2005.
- [104] B. Mennucci, S. Caprasecca, and C. A. Guido, “Computational Studies of Environmental Effects and Their Interplay With Experiment”, *Adv. Phys. Org. Chem.* **50**, 203 (2016).
- [105] *Continuum Solvation Models in Chemical Physics: From Theory to Applications / Wiley*.
- [106] I. Duchemin, D. Jacquemin, and X. Blase, “Combining the GW formalism with the polarizable continuum model: A state-specific non-equilibrium approach”, *J. Chem. Phys.* **144**, 164106 (2016).
- [107] R. E. Bulo, B. Ensing, J. Sikkema, and L. Visscher, “Toward a practical method for adaptive qm/mm simulations”, *J. Chem. Theory and Comput.* **5**, 2212 (2009).
- [108] J. M. H. Olsen, C. Steinmann, K. Ruud, and J. Kongsted, “Polarizable density embedding: A new QM/QM/MM-based computational strategy”, *J. Phys. Chem. A* **119**, 5344 (2015).

- [109] 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", Phys. Rev. B **97**, 35108 (2018).
- [110] X. Wen, D. S. Graham, D. V. Chulhai, and J. D. Goodpaster, "Absolutely Localized Projection-Based Embedding for Excited States", J. Chem. Theory Comput. **16**, 385 (2020).
- [111] 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", J. Phys. Chem. Lett. **7**, 2814 (2016).
- [112] 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**, 4430 (2018).
- [113] E. D. Hedegård and M. Reiher, "Polarizable Embedding Density Matrix Renormalization Group", J. Chem. Theory Comput. **12**, 4242 (2016).
- [114] J. Tölle, T. Deilmann, M. Rohlfsing, and J. Neugebauer, "Subsystem-based gw/bethe-salpeter equation", J.Chem. Theory Comput. **17**, 2186 (2021).
- [115] A. Stone, *The Theory of Intermolecular Forces*, 2nd editio (Oxford University Press, Oxford, 2013), p. 352.
- [116] B. T. Thole, "Molecular polarizabilities calculated with a modified dipole interaction", Chemical Phys. **59**, 341 (1981).
- [117] P. T. Van Duijnen and M. Swart, "Molecular and atomic polarizabilities: Thole's model revisited", J. Phys. Chem. A **102**, 2399 (1998).
- [118] A. Guerrero and G. Garcia-Belmonte, "Recent advances to understand morphology stability of organic photovoltaics", Nano-Micro Lett. **9**, 1 (2017).

- [119] V. Savikhin, M. Babics, M. Neophytou, S. Liu, S. D. Oosterhout, H. Yan, X. Gu, P. M. Beaujuge, and M. F. Toney, “Impact of Polymer Side Chain Modification on OPV Morphology and Performance”, *Chem. Mater.* **30**, 7872 (2018).
- [120] F. C. Grozema, P. T. Van Duijnen, Y. A. Berlin, M. A. Ratner, and L. D. Siebbeles, “Intramolecular charge transport along isolated chains of conjugated polymers: Effect of torsional disorder and polymerization defects”, *J. Phys. Chem. B* **106**, 7791 (2002).
- [121] J. Wildman, P. Repiščák, M. J. Paterson, and I. Galbraith, “General Force-Field Parametrization Scheme for Molecular Dynamics Simulations of Conjugated Materials in Solution”, *J. Chem. Theory Comput.* **12**, 3813 (2016).
- [122] V. Zoete, M. A. Cuendet, A. Grosdidier, and O. Michielin, “SwissParam: A fast force field generation tool for small organic molecules”, *J. Comput. Chem.* **32**, 2359 (2011).
- [123] K. Vanommeslaeghe, E. P. Raman, and A. D. MacKerell, “Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges”, *J. Chem. Inf. Modeling* **52**, 3155 (2012).
- [124] I. Ramos Sasselli, R. V. Ulijn, and T. Tuttle, “CHARMM force field parameterization protocol for self-assembling peptide amphiphiles: The Fmoc moiety”, *Phys. Chem. Chem. Phys.* **18**, 4659 (2016).
- [125] L. S. Dodd, I. C. De Vaca, J. Tirado-Rives, and W. L. Jorgensen, “LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands”, *Nucleic Acids Research* **45**, W331 (2017).
- [126] N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark, and W. F. Van Gunsteren, “Definition and testing of the GROMOS force-field versions 54A7 and 54B7”, *Eur. Biophys. J.* **40**, 843 (2011).

- [127] A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, and A. E. Mark, “An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0”, *J. Chem. Theory Comput.* **7**, 4026 (2011).
- [128] U. C. Singh and P. A. Kollman, “An approach to computing electrostatic charges for molecules”, *J. Comput. Chem.* **5**, 129 (1984).
- [129] W. R. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. Van Gunsteren, “The GROMOS biomolecular simulation program package”, *J. Phys. Chem. A* **103**, 3596 (1999).
- [130] 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”, *J. Comput. Chem.* **11**, 361 (1990).
- [131] F. Neese, “The ORCA program system”, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2**, 73 (2012).
- [132] T. Yanai, D. P. Tew, and N. C. Handy, “A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP)”, *Chem. Phys. Lett.* **393**, 51 (2004).
- [133] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, “RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency”, *Chem. Phys. Lett.* **294**, 143 (1998).
- [134] J. Wehner, L. Brombacher, J. Brown, C. Junghans, O. Çaylak, Y. Khalak, P. Madhikar, G. Tirimbò, and B. Baumeier, “Electronic Excitations in Complex Molecular Environments: Many-Body Green’s Functions Theory in VOTCA-XTP”, *J. Chem. Theory Comput.* **14**, 6253 (2018).
- [135] G. Tirimbò, V. Sundaram, O. Ca&cedil;ylak, W. Scharpach, J. Sijen, C. Junghans, J. Brown, F. Z. Ruiz, N. Renaud, J. Wehner, and B. Baumeier, “Excited-state electronic structure of molecules using many-body Green’s functions: Quasiparticles and electron-

- hole excitations with VOTCA-XTP”, J. Chem. Phys. **152**, 114103 (2020).
- [136] H. Bekker, H. Berendsen, E. Dijkstara, S. Achterop, R. Vondrumen, D. Vanderspoel, A. Sijbers, H. Keegstra, and M. Renardus, *GROMACS - A parallel computer for molecular-dynamics simulations* (World Scientific Publishing, 1993), pp. 252–256.
  - [137] S. S. Zade and M. Bendikov, “Twisting of conjugated oligomers and polymers: Case study of oligo- And polythiophene”, Chem. - Eur. J. **13**, 3688 (2007).
  - [138] W. L. Jorgensen and J. Tirado-Rives, “The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin”, J. Am. Chem. Soc. **110**, 1657 (1988).
  - [139] G. Bussi, D. Donadio, and M. Parrinello, “Canonical sampling through velocity rescaling”, J. Chem. Phys. **126**, 014101 (2007).
  - [140] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, “The Missing Term in Effective Pair Potentials1”, J. Phys. Chem **91**, 6269 (1987).
  - [141] H. J. Berendsen, J. P. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak, “Molecular dynamics with coupling to an external bath”, J. Chem. Phys. **81**, 3684 (1984).
  - [142] M. Grzybowski, E. Glodkowska-Mrowka, G. Clermont, M. Blanchard-Desce, and D. T. Gryko, “Synthesis and optical properties of water-soluble diketopyrrolopyrroles”, Chem. of Heterocyclic Compounds **53**, 72 (2017).
  - [143] S. Tang, E. H. Ghazvini Zadeh, B. Kim, N. T. Toomey, M. V. Bondar, and K. D. Belfield, “Protein-induced fluorescence enhancement of two-photon excitable water-soluble diketopyrrolopyrroles”, Org. Biomol. Chem. **15**, 6511 (2017).
  - [144] R. Xie, Y. Lee, M. P. Aplan, N. J. Caggiano, C. Müller, R. H. Colby, and E. D. Gomez, “Glass Transition Temperature of Conjugated Polymers by Oscillatory Shear Rheometry”, Macromolecules **50**, 5146 (2017).

- [145] D. Leman, M. A. Kelly, S. Ness, S. Engmann, A. Herzing, C. Snyder, H. W. Ro, R. J. Kline, D. M. DeLongchamp, and L. J. Richter, “In situ characterization of polymer-fullerene bilayer stability”, *Macromolecules* **48**, 383 (2015).
- [146] S. Luo, T. Wang, M. U. Ocheje, S. Zhang, J. Xu, Z. Qian, X. Gu, G. Xue, S. Rondeau-Gagné, J. Jiang, W. Hu, E. Zhuravlev, and D. Zhou, “Multiamorphous Phases in Diketopyrrolopyrrole-Based Conjugated Polymers: From Bulk to Ultrathin Films”, *Macromolecules*, acs.macromol.9b02738 (2020).
- [147] S. Sakka and J. D. Mackenzie, “Relation between apparent glass transition temperature and liquids temperature for inorganic glasses”, *J. Non-Crystalline Solids* **6**, 145 (1971).
- [148] F. Weigend, A. Köhn, and C. Hättig, “Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations”, *J. Chem. Phys.* **116**, 3175 (2002).
- [149] C. Adamo and V. Barone, “Toward reliable density functional methods without adjustable parameters: The PBE0 model”, *J. Chem. Phys.* **110**, 6158 (1999).
- [150] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Guru-moorthi, J. Chase, J. Li, and T. L. Windus, “Basis set exchange: A community database for computational sciences”, *J. Chemical Information and Modeling* **47**, 1045 (2007).
- [151] G. Tirimbò, X. de Vries, C. H. L. Weijtens, P. A. Bobbert, T. Neumann, R. Coehoorn, and B. Baumeier, “Quantitative Predictions of Photoelectron Spectra in Amorphous Molecular Solids from Multiscale Quasiparticle Embedding”, *101*, 035402 (2020).
- [152] B. Baumeier, M. Rohlfing, and D. Andrienko, “Electronic excitations in push-pull oligomers and their complexes with fullerene from many-body Green’s functions theory with polarizable embedding”, *J. Chem. Theory Comput.* **10**, 3104 (2014).

- [153] A. S. Larsen, M. T. Ruggiero, K. E. Johansson, J. A. Zeitler, and J. Rantanen, “Tracking dehydration mechanisms in crystalline hydrates with molecular dynamics simulations”, *Crys. Growth Des.* **17**, 5017 (2017).
- [154] V. Negi, A. Lyulin, and P. Bobbert, “Solvent-dependent structure formation in drying p3ht:pcbm films studied by molecular dynamics simulations”, *Macromol. Theory Simul.* **25**, 550 (2016).
- [155] M. Tsige and G. S. Grest, “Molecular dynamics study of the evaporation process in polymer films”, *Macromolecules* **37**, 4333 (2004).
- [156] H. Tabe, K. Kobayashi, H. Fujii, and M. Watanabe, “Molecular dynamics simulation of evaporation coefficient of vapor molecules during steady net evaporation in binary mixture system”, *Int. J. Heat Mass Transf.* **188**, 122663 (2022).
- [157] M. M. Reif, M. Winger, and C. Oostenbrink, “Testing of the gromos force-field parameter set 54a8: structural properties of electrolyte solutions, lipid bilayers, and proteins”, *J. Chem. Theory Comput.* **9**, 1247 (2013).
- [158] V. Sundaram, A. V. Lyulin, and B. Baumeier, “Development and testing of an all-atom force field for diketopyrrolopyrrole polymers with conjugated substituents”, *J. Phys. Chem. B* **124**, 11030 (2020).
- [159] D. Moghe, G. K. Dutta, S. Patil, and S. Guha, “Photocurrent spectroscopic studies of diketopyrrolopyrrole-based statistical copolymers †”, *Phys. Chem. Chem. Phys* **16**, 4291 (2014).
- [160] F. Liu, Y. Gu, C. Wang, W. Zhao, D. Chen, A. L. Briseno, and T. P. Russell, “Efficient polymer solar cells based on a low bandgap semi-crystalline dpp polymer-pcbm blends”, *Adv. Mater.* **24**, 3947 (2012).
- [161] M. Brehm and B. Kirchner, “Travis - a free analyzer and visualizer for monte carlo and molecular dynamics trajectories”, *J.Chem. Inf. Modeling* **51**, 2007 (2011).

- [162] M. Brehm, M. Thomas, S. Gehrke, and B. Kirchner, “Travis—a free analyzer for trajectories from molecular simulation”, *J. Chem. Phys.* **152**, 164105 (2020).
- [163] K. Y. Chiu, S. H. Chang, and W.-C. Huang, “To cite this article: trinh tung ngo et al”, *Adv. Nat. Sci: Nanosci. Nanotechnol.* **3**, 45001 (2012).
- [164] A. A. Guillet, M. Zbiri, A. D. Dunbar, and J. Nelson, “Quantitative analysis of the molecular dynamics of p3ht:pcbm bulk heterojunction”, *J. Phys. Chem. B* **121**, 9073 (2017).
- [165] D. V. D. Spoel and H. J. Berendsen, “Molecular dynamics simulations of leu-enkephalin in water and dmso”, *Biophys. J.* **72**, 2032 (1997).
- [166] J. Wehner and B. Baumeier, “Multiscale simulations of singlet and triplet exciton dynamics in energetically disordered molecular systems based on many-body Green’s functions theory Multiscale simulations of singlet and triplet exciton dynamics in energetically disordered molecular s”, *New J. Phys* **22**, 33033 (2020).
- [167] X. De Vries, P. Friederich, W. Wenzel, R. Coehoorn, and P. A. Bobbert, “Full quantum treatment of charge dynamics in amorphous molecular semiconductors”, *Phys. Rev. B* **97**, 075203 (2018).
- [168] V. Rühle, A. Lukyanov, F. May, M. Schrader, T. Vehoff, J. Kirkpatrick, B. Baumeier, and D. Andrienko, “Microscopic simulations of charge transport in disordered organic semiconductors”, *J. Chem. Theory Comput.* **7**, 3335 (2011).
- [169] V. Sundaram, A. V. Lyulin, and B. Baumeier, “Effect of solvent removal rate and annealing on the interface properties in a blend of a diketopyrrolopyrrole-based polymer with fullerene”, *J. Phys. Chem. B* **126**, 7445 (2022).
- [170] F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller, “A simple, exact density-functional-theory embedding scheme”, *J. Chem. Theory Comput.* **8**, 2564 (2012).

- [171] S. J. R. Lee, F. Ding, F. R. Manby, and T. F. Miller, “Analytical gradients for projection-based wavefunction-in-dft embedding”, *J. Chem. Phys.* **151**, 064112 (2019).
- [172] D. Claudino and N. J. Mayhall, “Automatic partition of orbital spaces based on singular value decomposition in the context of embedding theories”, *J. Chem. Theory Comput.* **15**, 1053 (2019).
- [173] B. Hégely, P. R. Nagy, G. G. Ferenczy, and M. Kállay, “Exact density functional and wave function embedding schemes based on orbital localization”, *J. Chem. Phys.* **145**, 064107 (2016).
- [174] T. A. Barnes, J. D. Goodpaster, F. R. Manby, and T. F. Miller, “Accurate basis set truncation for wavefunction embedding”, *J. Chem. Phys.* **139**, 024103 (2013).
- [175] S. J. Bennie, M. Stella, T. F. Miller, and F. R. Manby, “Accelerating wavefunction in density-functional-theory embedding by truncating the active basis set”, *J. Chem. Phys.* **143**, 024105 (2015).
- [176] S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher, and M. Reiher, “Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds”, *J. Chem. Phys.* **132**, 164101 (2010).
- [177] A. S. P. Gomes, C. R. Jacob, and L. Visscher, “Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory”, *Phys. Chem. Chem. Phys.* **10**, 5353 (2008).
- [178] C. R. Jacob, J. Neugebauer, and L. Visscher, “A flexible implementation of frozen-density embedding for use in multilevel simulations”, *J. Comput. Chem.* **29**, 1011 (2008).
- [179] J. Pipek and P. G. Mezey, “A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions”, *J. Chem. Phys.* **90**, 4916 (1989).
- [180] S. Lehtola and H. Jónsson, “Unitary optimization of localized molecular orbitals”, *J. Chem. Theory Comput.* **9**, 5365 (2013).

- [181] R. McWeeny, “Some recent advances in density matrix theory”, Rev. Mod. Phys. **32**, 335 (1960).
- [182] J. Tölle, T. Deilmann, M. Rohlfing, and J. Neugebauer, “Correction to “subsystem-based gw/bethe–salpeter equation””, J.Chem. Theory Comput. **19**, 2699 (2023).
- [183] G. Senatore and K. Subbaswamy, “Density dependence of the dielectric constant of rare-gas crystals”, Phys. Rev. B: Condens. Matter Mater. Phys. **34**, 5754 (1986).
- [184] M. Johnson, K. Subbaswamy, and G. Senatore, “Hyperpolarizabilities of alkali halide crystals using the local-density approximation”, Phys. Rev. B: Condens. Matter Mater. Phys. **36**, 9202 (1987).
- [185] P. Cortona, “Self-consistently determined properties of solids without band-structure calculations”, Phys. Rev. B: Condens. Matter Mater. Phys. **44**, 8454 (1991).
- [186] T. A. Wesolowski and A. Warshel, “Frozen density functional approach for ab initio calculations of solvated molecules”, J. Phys. Chem. **97**, 8050 (1993).
- [187] T. A. Wesolowski and J. Weber, “Kohn-sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interacting molecules”, Chem. Phys. Lett. **248**, 71 (1996).
- [188] 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”, J.Phys. Chem. Lett. **7**, 2814 (2016).
- [189] 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”, Chem. Sci. **9**, 4430 (2018).
- [190] S. Baral, B. Baumeier, and E. Lyman, “Predicting spectral properties of polarity sensitive dyes with qm/mm simulation”, Biophys. J. **114**, 272a (2018).

- [191] L. C. Emery, J. M. Sheldon, W. Edwards, and J. L. McHale, “Intermolecular ground and excited state potential energy surfaces and charge-transfer spectrum of benzene: tcne electron donor–acceptor complex”, *Spectrochimica Acta Part A: Molecular Spectroscopy* **48**, 715 (1992).
- [192] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, “Packmol: a package for building initial configurations for molecular dynamics simulations”, *J. Comput. Chem.* **30**, 2157 (2009).
- [193] J. M. Martínez and L. Martínez, “Packing optimization for automated generation of complex system’s initial configurations for molecular dynamics and docking”, *J. Comput. Chem.* **24**, 819 (2003).
- [194] C. Degitz, M. Schmid, F. May, J. Pfister, A. Auch, W. Brütting, and W. Wenzel, “From molecule to device: prediction and validation of the optical orientation of iridium phosphors in organic light-emitting diodes”, *Chem. Mater.* **35**, 295 (2023).
- [195] R. A. Marcus, “Electron transfer reactions in chemistry: theory and experiment (nobel lecture)”, *Angewandte Chemie International Edition in English* **32**, 1111 (1993).
- [196] G. R. Hutchison, M. A. Ratner, and T. J. Marks, “Hopping transport in conductive heterocyclic oligomers: reorganization energies and substituent effects”, *J. Am. Chem. Soc.* **127**, 2339 (2005).
- [197] J. M. Martin and A. Sundermann, “Correlation consistent valence basis sets for use with the stuttgart–dresden–bonn relativistic effective core potentials: the atoms ga–kr and in–xe”, *J. Chem. Phys.* **114**, 3408 (2001).
- [198] S. O. Odoh and G. Schreckenbach, “Performance of relativistic effective core potentials in dft calculations on actinide compounds”, *J. Phys. Chem. A* **114**, 1957 (2010).
- [199] L. Maron, C. Teichteil, R. Poteau, and F. Alary, “Recent progress in atomic and chemical group effective potentials”, *J. Chem. Info. Comput.* **41**, 43 (2001).

- [200] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, “Self-consistent molecular orbital methods. xx. a basis set for correlated wave functions”, *J. Chem. Phys.* **72**, 650 (1980).
- [201] T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. V. R. Schleyer, “Efficient diffuse function-augmented basis sets for anion calculations. iii. the 3-21+ g basis set for first-row elements, li-f”, *J. Comput. Chem.* **4**, 294 (1983).
- [202] N. S. Hush, “Adiabatic theory of outer sphere electron-transfer reactions in solution”, *Trans. Faraday Soc.* **57**, 557 (1961).
- [203] B. Bagheri, M. Karttunen, and B. Baumeier, “Solvent effects on optical excitations of poly para phenylene ethynylene studied by qm/mm simulations based on many-body green’s functions theory”, *Eur Phys J Spec Top.* **225**, 1743 (2016).
- [204] B. T. Draine and J. Goodman, “Beyond clausius-mossotti-wave propagation on a polarizable point lattice and the discrete dipole approximation”, *Astrophysical Journal, Part 1* (ISSN 0004-637X), vol. 405, no. 2, p. 685-697. **405**, 685 (1993).