

# Subsystem-Based Methods for Global and Local Optical Properties

Von der Naturwissenschaftlichen Fakultät der  
Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades  
Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

Marina Jansen, M.Sc.

2023

Referentin: Prof. Dr. rer. nat. Carolin König  
Koreferent: Prof. Dr. rer. nat. Jörg August Becker  
Tag der Promotion: 20.10.2023

# Kurzzusammenfassung

Die genaue Berechnung optischer Eigenschaften in komplexen Systemen stellt aufgrund der quantenmechanischen Natur optischer Prozesse eine herausfordernde Aufgabe dar. Subsystembasierte Methoden bieten einen vielversprechenden Ansatz, um diese Herausforderungen zu bewältigen, indem das gesamte System in kleinere, rechnerisch effizientere Teilsysteme aufgeteilt wird. Diese Dissertation beschäftigt sich mit zwei unterschiedlichen subsystembasierten Ansätzen zur Berechnung optischer Eigenschaften: lokale Einbettungsmethoden und globale Fragmentierungsschemata.

Für die lokalen Einbettungsmethoden wurden verschiedene polarisierbare Einbettungsschemata analysiert, um optische Anregungen von in Wasser gelöstem *para*-Nitroanilin und pentamerer Formylthiophenessigsäure zu berechnen. Durch die Aufschlüsselung der einzelnen Wechselwirkungseffekte in einem gemeinsamen theoretischen Rahmen und die umfassende Implementierung wurde ein direkter Vergleich zwischen *polarizable embedding* und *frozen-density embedding* ermöglicht. Wertvolle Erkenntnisse über die Bedeutung einzelner Wechselwirkungseffekte und die Stärken und Grenzen der Ansätze wurden verdeutlicht. Die Ergebnisse zeigen die wesentliche Bedeutung der gegenseitigen Grundzustandspolarisierung in den Einbettungsschemata sowie die teilweise bedeutsamen dynamischen Umgebungseffekte. Insbesondere eine starke Abhängigkeit von den zugrunde liegenden strukturellen Geometrien konnte beobachtet werden.

Darüber hinaus ermöglichten Fragmentierungsschemata die Berechnung globaler optischer Eigenschaften für verschiedene *zeolitic imidazolate frameworks* (ZIFs). In einer umfassenden Studie wurde ein rechnerisches Protokoll etabliert, um Geometrien zu erhalten und die Brechungsindices für ZIF-8 mit reduziertem Rechenaufwand zu berechnen. Die implementierten Fragmentierungsschemata wurden nachfolgend angewendet, um optische Eigenschaften für ZIFs mit derselben Topologie, aber substituierten organischen Linkern vorherzusagen. Die Ergebnisse verdeutlichen das hohe Potenzial dieses Ansatzes für die Charakterisierung und Vorhersage optischer Eigenschaften in leeren ZIFs. Darüber hinaus wurde das Fragmentierungsschema angewendet, um Gastmoleküle in die poröse Struktur der ZIFs integrieren zu können und erste Schätzungen für die bevorzugten Positionen und Anzahl der in die Poren eingelagerten Moleküle zu liefern.

Insgesamt stellt diese Arbeit neuartige theoretische und rechnerische Rahmenwerke für die genaue Berechnung optischer Eigenschaften in komplexen Systemen vor. Die vorgestellten Ansätze bieten eine vielversprechende Möglichkeit für die zukünftige Forschung im Bereich der optischen Eigenschaften. Sie ermöglichen eine tiefere Erkenntnis der Wechselwirkungen und ihrer Auswirkungen auf die optische Antwort verschiedener Moleküle. Sie ebnet den Weg für eine verbesserte Charakterisierung und Gestaltung fortschrittlicher Materialien mit genauen optischen Funktionen.

**Schlagwörter:** Optische Eigenschaften, Subsystem-basierte Methoden, Polarizable Embedding, Dichtebasierte Einbettung, Molekulare Fragmentierung, Metallorganische Gerüste, Lumineszente Konjugierte Oligothiophene

# Abstract

The accurate calculation of optical properties in complex systems is a challenging task due to the quantum-mechanical nature of optical processes and the computational demands involved. Subsystem-based methods offer a promising approach to tackle these challenges by partitioning the full system into smaller, computationally more efficient subsystems while maintaining the necessary accuracy. In this thesis, two distinct subsystem-based approaches for calculating optical properties are investigated: local embedding methods and global fragmentation schemes.

In the local embedding methods, different polarizable embedding schemes were employed to calculate optical excitations on solvated *para*-nitroaniline and pentameric formyl thiophene acetic acid. By dissecting the individual interaction effects in a common theoretical framework and developing an extensive computational setup, a one-to-one comparison between *polarizable embedding* and *frozen-density embedding* was performed. This comparison provided valuable insights into the importance of separate interaction effects and highlighted the strengths and limitations of each approach. The results revealed the major significance of mutual ground-state polarization in the embedding schemes and the partial importance of dynamical environment effects. Specifically, a strong dependence on the underlying structural geometries could be observed.

The global fragmentation schemes allow the calculation of optical properties for various zeolitic imidazolate frameworks (ZIFs). In an extensive study, a general computational protocol was established in order to obtain geometries and calculate refractive indices for ZIF-8, yielding accurate results with reduced computational demands. The implemented fragmentation schemes were then extended to predict optical properties for ZIFs with the same topology but substituted organic linkers. The results indicated the high potential of this approach for the rationalization and prediction of optical properties in empty ZIFs. Moreover, the framework was adapted to incorporate guest molecules into the porous structure of the ZIFs, providing rough estimates of the number of molecules incorporated per pore.

Overall, this thesis introduces novel theoretical and computational frameworks for the accurate calculation of optical properties in complex systems. The presented approaches offer a promising direction for future research in the field of optical properties, enabling a deeper understanding of the interactions and their effects on the optical response of various molecules. This paves the way for an improved characterization and design of advanced materials with precise optical functionalities.

**Keywords:** Optical Properties, Subsystem-Based Methods, Polarizable Embedding, Density-Based Embedding, Molecular Fragmentation, Metal-Organic Frameworks, Luminescent Conjugated Oligothiophenes

# Table of Contents

<b>Kurzzusammenfassung</b>	<b>3</b>
<b>Abstract</b>	<b>5</b>
<b>List of Abbreviations</b>	<b>11</b>
<b>1. Introduction</b>	<b>13</b>
1.1. Optical Properties <i>via</i> Subsystem-Based Approaches . . . . .	13
1.2. Local Subsystem-Based Methods . . . . .	15
1.3. Global Subsystem-Based Methods . . . . .	16
1.4. Outline of this Work . . . . .	18
<b>I. Background</b>	<b>21</b>
<b>2. Theoretical Background for Optical Properties</b>	<b>23</b>
2.1. Introduction . . . . .	23
2.2. Density-Functional Theory . . . . .	25
2.2.1. Kohn-Sham Density-Functional Theory . . . . .	27
2.3. Time-Dependent Density-Functional Theory . . . . .	30
2.4. Subsystem-Based Approaches for Optical Properties . . . . .	35
2.4.1. Local Optical Properties . . . . .	36
2.4.1.1. Density-Based Embedding . . . . .	36
2.4.1.2. Frozen-Density Embedding . . . . .	39
2.4.1.3. Polarizable Embedding . . . . .	40
2.4.2. Global Optical Properties . . . . .	43
2.4.2.1. Optical Properties from Energy-Based Fragmentation Schemes	45

<b>II. Research</b>	<b>47</b>
<b>3. Computational Details</b>	<b>49</b>
3.1. Computational Details for Calculations of Local Optical Properties . . . .	49
3.2. Computational Details for Calculations of Global Optical Properties . . .	51
<b>4. Dissection of Interaction Contributions for Local Optical Properties</b>	<b>53</b>
4.1. Theoretical Framework for Quantum-Derived Embedding Schemes . . . .	53
4.1.1. Density-Based Embedding . . . . .	57
4.1.2. Polarizable Embedding . . . . .	60
4.1.3. Comparison of Embedding Schemes in the Common Framework . .	64
4.2. Numerical Comparison of Quantum-Derived Embedding Schemes . . . .	66
4.2.1. Computational Setup and Implementation . . . . .	66
4.2.1.1. Polarizable Embedding . . . . .	68
4.2.1.2. Frozen-Density Embedding . . . . .	69
4.2.2. Numerical Comparison . . . . .	71
4.2.2.1. <i>para</i> -Nitroaniline . . . . .	72
4.2.2.2. Pentameric Formyl Thiophene Acetic Acid . . . . .	77
4.3. Summary . . . . .	82
<b>5. Refractive Indices of Zeolitic Imidazolate Frameworks via Fragmentation</b>	<b>85</b>
5.1. Refractive Indices for Zeolitic Imidazolate Frameworks . . . . .	86
5.1.1. Fragmentation Schemes . . . . .	86
5.1.2. Method Evaluation . . . . .	89
5.1.2.1. Exchange–Correlation Functional and Basis Set Evaluation	89
5.1.3. Refractive Indices for ZIF-8 . . . . .	94
5.1.3.1. Electrostatic Embedding . . . . .	98
5.1.4. Refractive Indices for Sodalite-Type Zeolitic Imidazolate Frameworks	100
5.2. Loading of Zeolitic Imidazolate Frameworks in Fragmentation Schemes .	105
5.2.1. Fragmentation Schemes . . . . .	105
5.2.2. Refractive Indices for Loaded Zeolitic Imidazolate Frameworks . .	107
5.2.2.1. ZIF-8 . . . . .	108
5.2.2.2. ZIF-71 . . . . .	112
5.3. Summary . . . . .	114
<b>6. Summary, Conclusions and Outlook</b>	<b>117</b>



<b>III. Appendix</b>	<b>123</b>
<b>A. Local Optical Properties</b>	<b>125</b>
A.1. Computational Setup . . . . .	125
A.1.1. Polarizable Embedding . . . . .	125
A.1.2. Frozen-Density Embedding . . . . .	126
A.2. Numerical Comparison . . . . .	141
A.2.1. Functional Evaluation . . . . .	141
A.2.2. <i>para</i> -Nitroaniline . . . . .	142
A.2.2.1. Excitation Energies and $S$ -Shifts for <i>para</i> -Nitroaniline . . .	142
A.2.2.2. Oscillator Strengths and $\mathcal{F}$ -shifts for <i>para</i> -nitroaniline . . .	150
A.2.3. Pentameric Formyl Thiophene Acetic Acid . . . . .	157
A.2.3.1. Electron-Spill-Out Effect . . . . .	160
A.2.3.2. Introduction of Atomic Pseudopotentials to PE . . . . .	163
A.2.3.3. Shifts of Pentameric Formyl Thiophene Acetic Acid . . . . .	165
A.2.4. Publications . . . . .	167
A.2.4.1. Jansen <i>et al.</i> 2022 . . . . .	167
A.2.4.2. Jansen <i>et al.</i> 2023 . . . . .	204
<b>B. Global Properties</b>	<b>219</b>
B.1. MOF Guest Placer Script . . . . .	219
B.2. Results . . . . .	231
B.2.1. DFT Functional Evaluation . . . . .	231
B.2.2. Program Comparison . . . . .	234
B.2.3. Refractive Indices for Empty Zeolitic Imidazolate Frameworks . . .	236
B.2.4. Loading of Zeolitic Imidazolate Frameworks . . . . .	239
<b>Bibliography</b>	<b>241</b>
<b>Acknowledgements</b>	<b>271</b>
<b>Lebenslauf</b>	<b>273</b>
<b>Publikationsliste</b>	<b>274</b>



# List of Abbreviations

<b>CF3IM</b>	2-trifluoromethyl-imidazolate
<b>DCIM</b>	4,5-dichloroimidazolate
<b>ICA</b>	imidazolate-2-carboxyaldehyde
<b>MIM</b>	2-methyl-imidazolate
<b>pNA</b>	<i>para</i> -nitroaniline
<b>4MR</b>	four-membered ring
<b>6MR</b>	six-membered ring
<b>ALDA</b>	adiabatic local density approximation
<b>BU</b>	bottom-up
<b>COSMO</b>	conductor-like screening model
<b>CP-SCF</b>	coupled-perturbed self-consistent field
<b>DFT</b>	density-functional theory
<b>EEF</b>	effective external field
<b>ESO</b>	electron spill-out
<b>FDE</b>	frozen-density embedding
<b>GGA</b>	generalized gradient approximation
<b>HF</b>	Hartree-Fock
<b>IBU</b>	inorganic building unit
<b>KS</b>	Kohn-Sham
<b>KSCED</b>	Kohn-Sham equations with constrained electron densities
<b>LCO</b>	luminescent conjugated oligothiophene
<b>LDA</b>	local density approximation

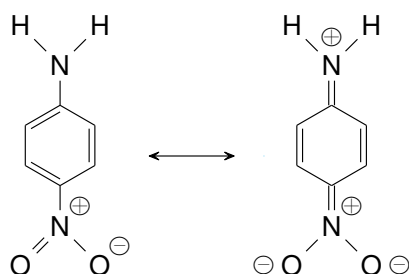
<b>MBE</b>	many-body expansion
<b>MD</b>	molecular dynamics
<b>MFCC</b>	molecular fractionation with conjugated caps
<b>MM</b>	molecular-mechanical
<b>MOF</b>	metal-organic framework
<b>pDFT</b>	periodic density-functional theory
<b>PE</b>	polarizable embedding
<b>PEAS</b>	polarizable embedding assistance script
<b>pFTAA</b>	pentameric formyl thiophene acetic acid
<b>QM</b>	quantum-mechanical
<b>QM/MM</b>	quantum-mechanical/molecular-mechanical
<b>QM/QM</b>	quantum-mechanical/quantum-mechanical
<b>SC XRD</b>	single crystal X-ray diffraction
<b>SCF</b>	self-consistent field
<b>SD</b>	Slater determinant
<b>sDFT</b>	subsystem density-functional theory
<b>TD</b>	top-down
<b>TD-DFT</b>	time-dependent density-functional theory
<b>XC</b>	exchange-correlation
<b>ZIF</b>	zeolitic imidazolate framework

# 1 Introduction

## 1.1. Optical Properties via Subsystem-Based Approaches

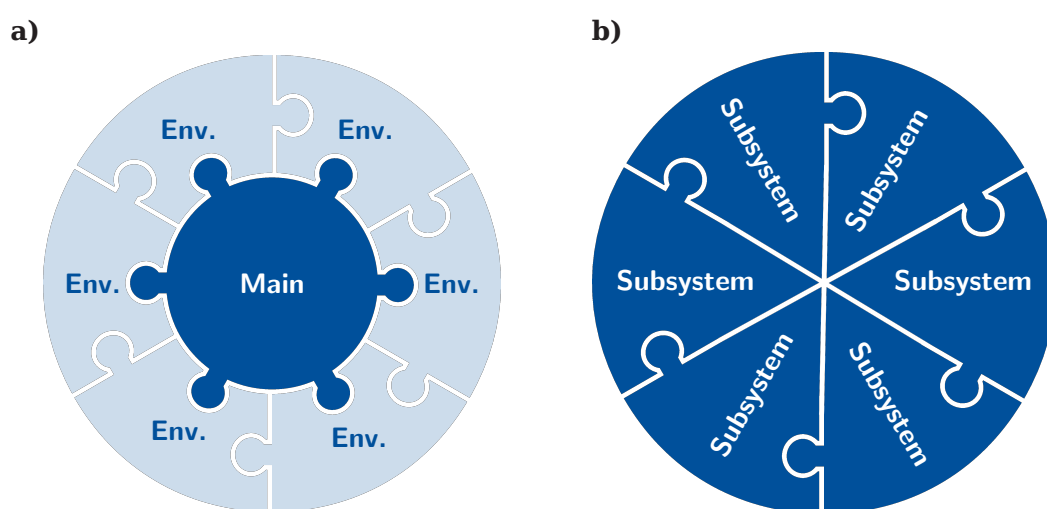
Optical properties refer to the fundamental characteristics and behaviour of materials in response to light-matter interactions. A wide array of phenomena, including absorption, emission, reflection, refraction, and scattering of light emerges from these interactions. To comprehend the intricate responses of materials to light-matter interactions and accurately predict their behaviour in this variety of applications, computational methods play a crucial role by enabling the calculation of optical properties.<sup>[1, 2]</sup> However, the calculation comes with its challenges, as light-matter interactions are inherently quantum-mechanical phenomena.<sup>[3, 4]</sup> The calculation of optical spectra is in many cases based on a density-functional theory (DFT) framework, that quickly reaches its limit concerning computational applicability when introducing extended system sizes.

For instance, smaller chromophores such as *para*-nitroaniline (*p*NA) (Fig. 1.1.1), can change their electronic structure significantly upon solvation and exhibit charge-transfer phenomena during excitation.<sup>[5, 6]</sup> The solvation process favours and stabilizes the zwitterionic structure (Fig. 1.1.1) in the excitation highlighting the importance of the solvent environment for the optical properties of complex systems.<sup>[5]</sup> Furthermore, chromophores that exhibit substantial changes in their optical properties upon solvation and specifically when docking to proteins have piqued interest in scientific research.<sup>[7-12]</sup> Luminescent conjugated oligothiophenes (LCOs) serve as fluorescent biomarkers for amyloid proteins in the detection of Alzheimer's and Parkinson's disease.<sup>[13, 14]</sup> Therefore, the underlying mechanisms governing the changes upon solvation are of utmost importance for the further improvement and design of these biomarkers, however, are not fully understood yet. They are known to originate from molecular conformation as well as the interaction with the solvent.<sup>[15-17]</sup> Previous studies have shown that environments of chromophores of up to 10-15 Å can contribute to solvation effects.<sup>[18, 19]</sup>



**Figure 1.1.1.:** Resonant Lewis structures of *para*-nitroaniline.

Moreover, extended periodic systems, such as porous metal-organic frameworks (MOFs), that are built from an inorganic building unit (IBU) linked by organic ligands, are difficult to describe due to their system size.<sup>[20, 21]</sup> The possibility to tune these properties by substitution of the IBUs or organic linkers is closely intertwined with the emergence of a plethora of experimentally accessible MOFs.<sup>[22-25]</sup> The interest in MOFs for optical properties<sup>[26-29]</sup> and particularly interest in the refractive index has grown recently.<sup>[22, 30-35]</sup> This is specifically due to the tunability of the refractive index when incorporating guest molecules into the porous structure.<sup>[35-37]</sup> The plain supermolecular calculation of these extended systems, specifically when introducing guest molecules, is extremely demanding for DFT-based methods, such as periodic density-functional theory (pDFT). Moreover, considering only isolated small extracted clusters, which are computationally feasible, neglects important interactions when determining global optical properties.



**Figure 1.1.2.:** Schematic representation of **a)** local and **b)** global subsystem-based methods. **a)** Main and Env. refers to the main and environmental subsystems, respectively.

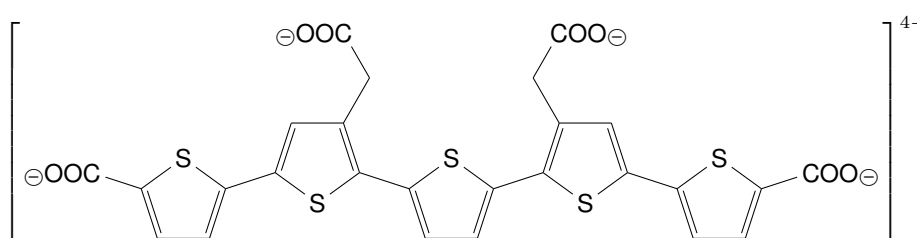
Subsystem-based methods offer a solution for the aforementioned limitations by reformulating the theoretical basis to account for the partitioning of the supermolecular system into smaller, manageable parts.<sup>[38]</sup> The division allows the individual calculation of these subsystems exploiting their computationally efficient size. In these schemes, two major approaches are commonly distinguished. The *local* approach makes a distinction between a main subsystem and its environment (Fig. 1.1.2 **a)**). Here, optical processes can be attributed to that main subsystem, while the surrounding environment is assumed to have a less pronounced influence on the optical behaviour.<sup>[39-42]</sup> The *global* approach (Fig. 1.1.2 **b)**) involves the calculation of subsystem properties that can be recombined to describe the property of the entire system.<sup>[43, 44]</sup>

## 1.2. Local Subsystem-Based Methods

Local subsystem-based methods (Fig. 1.1.2 a)), often referred to as embedding methods, rely on the quantum-mechanical (QM) description of the main system, while the environment is described more approximately. In a solvated system, it is possible to consider the environment implicitly by introducing a solvent continuum.<sup>[45, 46]</sup> However, directed solute–solvent interactions requiring explicit solvent representation or more heterogeneous environments are not covered by this approach. Therefore, two more extensive main methodologies are distinguished in the calculation of optical properties for embedded systems. QM/classical approaches describe the environment in a classical manner. Conversely, quantum-mechanical/quantum-mechanical (QM/QM) approaches introduce a QM treatment for the environmental subsystems.

The most well-known representative of the former is quantum-mechanical/molecular-mechanical (QM/MM) modelling, a hybrid approach originating from the ground-breaking work by Warshel and Levitt.<sup>[39, 42, 47, 48]</sup> Often this is done by employing point charges on environmental moieties in order to partially recover electrostatic interactions but can present in a variety of forms and complexities.<sup>[12, 49–53]</sup> Regardless of their success, it could be shown that the more accurate description of the environmental systems (*e.g.* polarization effects) strongly improves the accuracy of QM/MM models necessary for optical properties.<sup>[39, 54–56]</sup> Consequently, more advanced approaches have been developed to address the aforementioned limitations.

These include subsystem-based approaches with arbitrary environmental subsystems (*e.g.* solvent or protein). QM/classical polarizable embedding (PE) approaches describe the interactions between the subsystems *via* distributed multipoles and classical polarization contributions in the environment. Density-based QM/QM methods go beyond the classical description of the environment.<sup>[57–61]</sup> These approaches and combinations thereof<sup>[62–64]</sup> differ in the calculation of the main subsystem's and environmental energy and the definition of their interaction energy. In the calculation of optical properties, the different methodologies are combined either with state-specific methods, where the individual excited states are optimized<sup>[65, 66]</sup>, or with linear-response methods, where the influence of a time-dependent external field is treated perturbatively,<sup>[67–71]</sup>



**Figure 1.2.1.:** Lewis structures of pentameric formyl thiophene acetic acid anion.

Due to their different descriptions, methods within the extended QM/MM and approximate QM/QM methods have been extensively compared, however, comparisons have rarely been done between them.<sup>[72]</sup> Exploring the distinct interaction contributions and behaviours of these two different classes of embedding methods can offer valuable

insights into the strengths and limitations of each approach and provide guidance for selecting the most appropriate method. Moreover, the interaction contributions most important for the calculations of a specific system can be dissected.

The above-mentioned *p*NA has been widely studied with multiple embedding schemes, but not directly compared for the advanced QM/MM-based and approximate QM/QM methods.<sup>[65, 73-77]</sup> Pentameric formyl thiophene acetic acid (pFTAA) (Fig. 1.2.1), a representative for the LCO has been studied with QM/MM methods<sup>[78, 79]</sup>, but not with QM/QM. Due to its highly anionic charge (4-), the classical description of the environment might be more error-prone. Thus, the dissection of the polarization and quantum effects on these systems could support the efforts to comprehend the underlying polarization contributions in the excitation and in the case of pFTAA support the further design of fluorescent amyloid biomarkers.

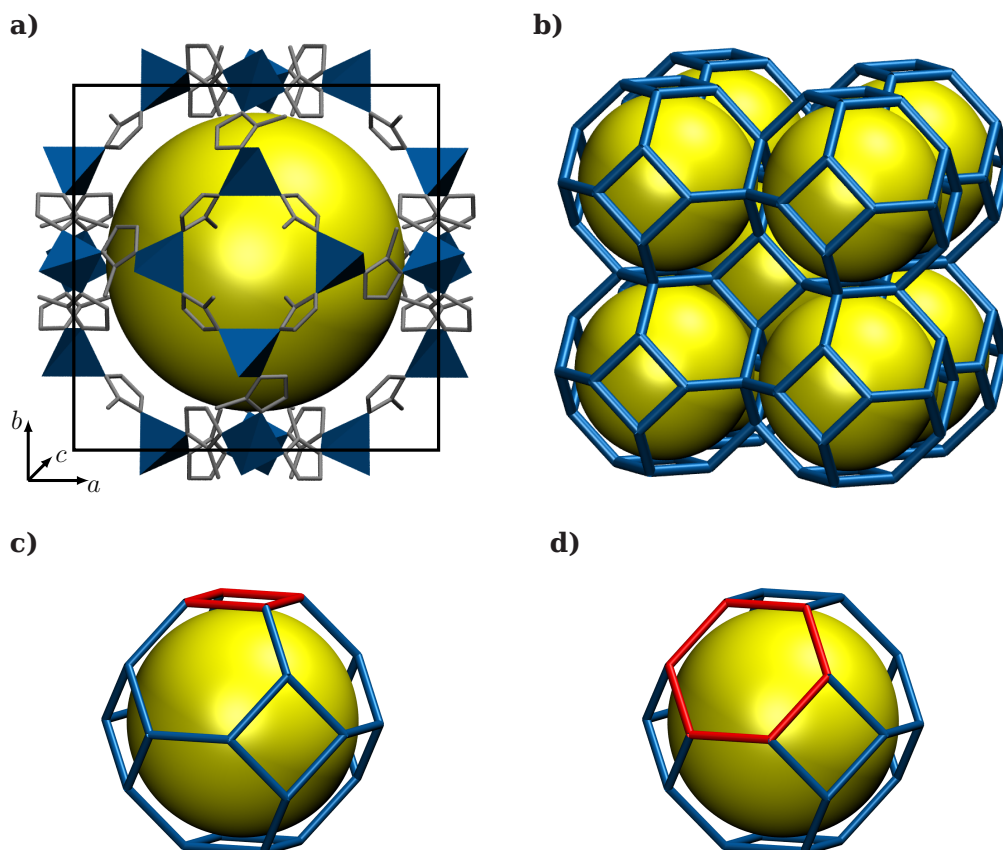
### 1.3. Global Subsystem-Based Methods

Subsystem-based methods for global optical properties determine the full supermolecular property by recombining the property of the additive individual subsystems (Fig. 1.1.2), which is in contrast to local embedding methods. This procedure is often referred to as molecular fragmentation<sup>[80-82]</sup> and involves the calculation of individual subsystem properties, that are subsequently combined to the (ideally exact) supermolecular system property.<sup>[43, 44, 80]</sup> Generally, two main types of fragmentation schemes can be distinguished. Density-based and energy-based fragmentation schemes.<sup>[83]</sup> The former involves the construction of the supermolecular electron density from the subsystem densities from where full system properties can be derived. In contrast, the latter concentrates on the calculation of subsystem energies that are combined to the supermolecular system energy, from which further properties can be derived.<sup>[83-85]</sup> Many energy-based fragmentation approaches were introduced, that mainly differ by the way the total energy expression is obtained (disjunct or overlapping initial fragments), the employment of capping substituents for the saturation of cut covalent bonds and the treatment of long-range interactions.<sup>[80, 83, 86-93]</sup>

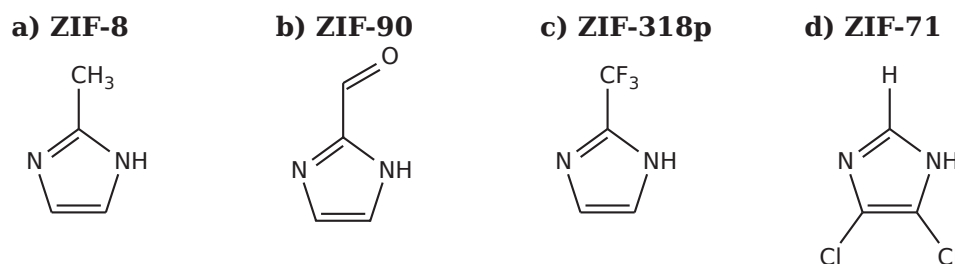
Specifically for systems that show periodic molecular patterns, (energy-based) fragmentation approaches alleviate a majority of the computational demand.<sup>[33, 89, 94, 95]</sup> For instance, the aforementioned porous MOFs are built from a repetitive pattern of an IBU linked by organic ligands<sup>[20, 21]</sup>, which makes them particularly attractive to study with these approaches. Specifically, zeolitic imidazolate frameworks (ZIFs), a class of MOFs built from transition metal ions coordinated with imidazolate linkers and resembling zeolite topology, have been found to show promising optical properties due to their wide range of refractive indices and the feasible synthesis of thin films of optical quality.<sup>[35, 96-101]</sup>

The composition of the most popular representative, ZIF-8, is displayed in Fig. 1.3.1 **a**).  $\text{Zn}^{2+}$  cations are tetrahedrally coordinated by 2-methyl-imidazolate (MIM) linkers and occur in sodalite topology. By the employment of different linkers, ZIFs in the same topology but with adapted optical properties can be formed (Fig. 1.3.2).<sup>[96, 102-106]</sup> The





**Figure 1.3.1.:** Representation of a **a)** ZIF-8 cubic unit cell at 110 plane and **b)** general structure with sodalite topology. **a)** Grey molecules represent MIM linkers tetrahedrally coordinating blue  $Zn^{2+}$  ions. **c)** and **d):** Schematic representation of the sodalite  $\beta$ -cage highlighting the **c)** four-membered ring (4MR) and **d)** a six-membered ring (6MR), that form the different pore windows. The yellow sphere highlights the empty pore inside the  $\beta$ -cage.



**Figure 1.3.2.:** Lewis structures of different organic linkers for constructing ZIFs. **a)** 2-methyl-imidazolate (MIM) for ZIF-8, **b)** imidazolate-2-carboxyaldehyde (ICA) for ZIF-90, **c)** 2-trifluoromethyl-imidazolate (CF3IM) for ZIF-318p and **d)** 4,5-dichloroimidazolate (DCIM) for ZIF-71.

topology contains  $\beta$ -cages that exhibit pores with a volume of approximately  $2500 \text{ \AA}^3$  (Fig. 1.3.1 **b**).<sup>[107]</sup> These large pores are connected by the so-called pore windows, six four-membered rings (4MRs) and eight six-membered rings (6MRs) marked in the sodalite cell in Fig. 1.3.1 **c**) and **d**), respectively. In gas-separation experiments it has been shown, that molecules can enter the pores *via* the pore windows.<sup>[108-111]</sup> It was found, that not only very small molecules but also larger molecules, such as methane could be incorporated.<sup>[35, 112]</sup> This implies some flexibility of the structure dependent on the membrane thickness<sup>[112]</sup> although showing high chemical stability.<sup>[96]</sup> Reversible gas adsorption enables the tailoring of optical properties.<sup>[35, 113]</sup> Keppler *et al.* studied optical-quality thin films of ZIF-8, tuning the refractive index by introducing *para*-methyl anisate, dimethylsulfoxide, tetrahydrofuran, toluene, iodobenzene or dimethylformamide as guest molecules. A refractive index range of  $\Delta n = 0.2$  could be obtained in the ellipsometric measurements.<sup>[35]</sup>

Previous computational characterization of properties for MOFs has been performed with full or partial molecular-mechanical (MM) methods (*cf.* Section 1.2)<sup>[114-119]</sup>, smaller clusters and DFT<sup>[24, 120-123]</sup> or DFT with periodic boundary conditions (pDFT)<sup>[124-129]</sup>. These lack applicability for optical properties either in terms of accuracy due to the neglect of QM effects, due to a not fully representative cluster size or in terms of computational demand.<sup>[119, 126-128]</sup> Treger *et al.* introduced an approximate fragmentation scheme to different MOFs including ZIFs.<sup>[129]</sup> The scheme enabled the calculation of the polarizability and the determination of the static refractive index by the Lorenz-Lorentz equation<sup>[130, 131]</sup>, however, interactions between subsystems were not regarded.

## 1.4. Outline of this Work

This thesis progresses with an outline of the theoretical background for the calculation of optical properties as well as the theoretical basis of the examined subsystem-based methods (Chapter 2). In the framework of this thesis, a dissection and analysis of polarization contributions as well as quantum effects for the calculation of optical properties with both an advanced polarizable QM/MM and a QM/QM embedding scheme are investigated on the example of *p*NA and *p*FTAA. In order to achieve this, a theoretical framework for the presented methodologies of local embedding schemes is derived. This framework facilitates the direct comparison of QM-classical and density-based embedding schemes. Additionally, a computational setup is implemented to numerically compare the embedding methods and enable a stepwise introduction of interaction effects. The numerical comparison involves the calculation of local optical properties for *p*NA and *p*FTAA systems solvated in varying sizes of a water environment. This study is presented in Chapter 4.

Subsequently, in Chapter 5, a set of different-order fragmentation schemes for refractive indices of MOFs is developed and evaluated in a comprehensive study. The developed computational protocol is then applied to other ZIFs in a sodalite topology (ZIF-8, ZIF-71, ZIF-90, and ZIF-318, Fig. 1.3.2). The impact of guest molecules, incorporated into the pores of ZIFs, on the refractive index is investigated.

This thesis targets the clarification of the importance of specific interaction contributions in local subsystem-based methods for the calculation of absorption properties. Embedding schemes with a classical and density-based environmental description are directly compared to highlight their strengths and limitations. Furthermore, the calculation of refractive indices with fragmentation schemes is developed and evaluated in order to allow for novel insights on the refractive indices of isomorphic and loaded ZIFs. The findings from these studies and the further research questions arising from these are summarized in Chapter 6.



# **Part I.**

## **Background**



## 2 Theoretical Background for Optical Properties

### 2.1. Introduction

The fundamental basis of quantum chemistry is the time-dependent Schrödinger equation. It can be written as,

$$\hat{H}_{\text{tot}}\Psi = E\Psi_{\text{tot}} = i\hbar\frac{\partial}{\partial t}\Psi_{\text{tot}}(\mathbf{R}, \mathbf{r}, t), \quad (2.1.1)$$

where the Hamilton operator  $\hat{H}_{\text{tot}}$  acts on the total wave function  $\Psi_{\text{tot}}(\mathbf{R}, \mathbf{r}, t)$ , that completely defines the system, yielding the total energy  $E$  as the eigenvalue. It depends on a set of coordinates for all nuclei ( $\mathbf{R}$ ) and electrons ( $\mathbf{r}$ ) and the time  $t$ . The imaginary unit and the reduced Planck constant are given as  $i$  and  $\hbar$ , respectively. For a molecule, the total Hamiltonian is comprised of the kinetic-energy operator ( $\hat{T}$ ) for electrons (el) and nuclei (nuc) and the potential-energy operators ( $\hat{V}$ ) referring to the electrostatic interaction of electrons and nuclei with and among each other

$$\hat{H}_{\text{tot}} = \hat{T}_{\text{nuc}}(\mathbf{R}) + \hat{T}_{\text{el}}(\mathbf{r}) + \hat{V}_{\text{nuc,el}}(\mathbf{r}, \mathbf{R}) + \hat{V}_{\text{el,el}}(\mathbf{r}) + \hat{V}_{\text{nuc,nuc}}(\mathbf{R}). \quad (2.1.2)$$

Expressing the kinetic energy through momenta and the potential energy through the Coulomb potentials of the involved particles, the molecular Hamiltonian reads

$$\hat{H}_{\text{tot}} = -\sum_I^N \frac{1}{2}\nabla_I^2 - \sum_i^n \frac{1}{2}\nabla_i^2 - \sum_i^n \sum_I^N \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{i<j}^n \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I<J}^N \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (2.1.3)$$

where the total number of electrons  $i$  and nuclei  $I$  are denoted as  $n$  and  $N$  so that  $Z_I$  is the nuclear charge and  $\mathbf{r}_i$  and  $\mathbf{R}_I$  denotes the electronic and nuclear coordinates, respectively.

In order to find a solution to the Schrödinger equation, if no time-dependent perturbation is applied to the system, a separation of the time-dependence from the total wave function *via* a product ansatz can be performed due to the assumption of (time-independent) stationary states. Following from this, the time-independent Schrödinger equation is obtained as,

$$\hat{H}_{\text{tot}}\tilde{\Psi}(\mathbf{R}, \mathbf{r}) = E_{\text{tot}}\tilde{\Psi}(\mathbf{R}, \mathbf{r}), \quad (2.1.4)$$

with the time-independent wave function  $\tilde{\Psi}(\mathbf{R}, \mathbf{r})$  and the total energy of the stationary state  $E_{\text{tot}}$ . Following from the immense computational complexity of solving this equation, commonly a further separation is aspired. The nuclei and electrons show a large disparity in their mass, *i.e.* the electronic movement can be assumed to instantaneously adapt to a change in the set of nuclear coordinates. In result, starting from a product ansatz, in the “clamped-nuclei” approximation the nuclear position and momentum are fixed. The nuclear kinetic energy term is omitted in the Hamiltonian (see Eq. (2.1.2)), thus, yielding a purely electronic Schrödinger equation

$$\hat{H}_{\text{el}}\psi_i(\mathbf{r}; \mathbf{R}) = E_{\text{el},i}(\mathbf{R})\psi_i(\mathbf{r}; \mathbf{R}) \quad (2.1.5)$$

with the electronic Hamiltonian  $\hat{H}_{\text{el}}$ . The electronic wave functions  $\psi_i$  and electronic energy  $E_{\text{el},i}$  are eigenfunctions and eigenvalues of the electronic Hamiltonian, respectively. Since the nuclear potential energy  $\hat{V}_{\text{nuc,nuc}}$  only depends on the set of nuclear coordinates, it can be regarded as an additive constant, to the electronic energy eigenvalue of the system. Although the nuclear motion is neglected in this approach, the electronic wave functions show a parametric dependence on the set of nuclear coordinates  $\mathbf{R}$ . This is due to the existence of an electronic Hamiltonian for every possible nuclear configuration, thus, the corresponding eigenfunctions and eigenvalues so that a set of coupled potential-energy surfaces arises. Performing the Born–Oppenheimer approximation<sup>[132]</sup>, the “non-adiabatic coupling” is neglected, yielding separate potential-energy surfaces on which the nuclei move. These allow for the description of molecular structures corresponding to a set of nuclear coordinates that show the minima of the electronic energy on the potential-energy surface.

In order to solve the electronic Schrödinger equation for a given set of nuclear coordinates (a point on the potential-energy surface), approximations have to be introduced, since it is only exactly solvable for a single electron. When introducing more than one electron the antisymmetry of the wave function with respect to an interchange of electrons has to be ensured. This is often taken into account by expressing the many-electron wave function in terms of a Slater determinant (SD) of single-electron wave functions.<sup>[133]</sup> Assuming that it can be approximated by a single SD, methods like the Hartree–Fock (HF) method, where the electron moves in a mean-field potential of the remaining electrons, were developed.<sup>[134, 135]</sup> Although showing a favourable scaling in computational demand, the major drawback is the lack of electron correlation (part of the electron–electron potential energy).<sup>[136]</sup> In an attempt to improve for this inaccuracy, post-HF methods have been proposed, such as configuration interaction<sup>[133, 137]</sup>, coupled-cluster<sup>[138–140]</sup>, or Møller Plesset perturbation theory<sup>[141]</sup>. In a dilemma between accuracy and computational complexity (*e.g.* CI is only solvable for extremely small systems), alternative ansatzes, that do not rely on expressing the wave function, have been formulated. One very popular representative is the so-called density-functional theory (DFT), which has been a workhorse in quantum chemistry due to its combination of accuracy and comparably low computational demands.<sup>[142]</sup>

In the following chapter, this approach will be laid out in detail (Section 2.2). Subsequently, an extension to the time-dependent case, the time-dependent density-functional



theory (TD-DFT) (Section 2.3), will be illustrated, which allows for the calculation of (optical) properties in an external potential. Furthermore, in Section 2.4, subsystem methods will be introduced and the application of DFT and time-dependent density-functional theory (TD-DFT) in order to calculate optical properties in these frameworks will be elaborated.

## 2.2. Density-Functional Theory

The general hypothesis in DFT is that the complexity of the many-electron problem can be reduced from the wave function object  $\Psi(\mathbf{r}, \sigma)$  depending on the spatial and spin coordinates of all electrons to the electron density  $\rho$  depending on three spatial coordinates. As early as 1928, the Thomas-Fermi model for the uniform electron gas emerged from statistical considerations proposing a purely electron-density-dependent energy expression<sup>[143-145]</sup>

$$E_{\text{TF}}[\rho] = T_{\text{TF}}[\rho] + V_{\text{nuc,el}}[\rho] + J[\rho]. \quad (2.2.1)$$

Here,  $J[\rho]$  denotes the classical Coulomb part of the electronic distribution, thus, neglecting the non-classical part of the  $V_{\text{el,el}}$  term (Eq. (2.1.2)). The kinetic Thomas-Fermi energy functional  $T_{\text{TF}}[\rho]$  is derived from the uniform electron gas and was further improved by considering the inhomogeneity of the electron density by integrating over small volume elements with an approximately constant electron density

$$T_{\text{TF}}[\rho] = \int \frac{3}{10} (3\pi^2)^{\frac{2}{3}} \rho(\mathbf{r})^{\frac{5}{3}}. \quad (2.2.2)$$

Although conceptually important, the model does not hold for the formation of stable bonds, since the energy is steadily decreasing for increasing interatomic distances.<sup>[146, 147]</sup> Corrections to the Thomas-Fermi theory by Dirac<sup>[148]</sup> and von Weizsäcker<sup>[149]</sup> did improve the model but likewise did not allow a sufficiently accurate description of stable bonds.<sup>[150-153]</sup>

More than 30 years later Hohenberg and Kohn laid the formal basis for DFT proving that for an interacting electron gas in an external potential  $v(\mathbf{r})$ , there exists a functional of the electron density  $F_{\text{HK}}[\rho]$ , that - independent of  $v(\mathbf{r})$  - minimizes to the ground-state energy. For this, they followed *via reductio ad absurdum* that the external potential uniquely corresponds to the ground-state electron density (first Hohenberg-Kohn theorem).<sup>[154]</sup> The number of electrons is readily obtained *via*

$$n = \int \rho(\mathbf{r}) d\mathbf{r}. \quad (2.2.3)$$

Consequently, the Hamiltonian of the system and, therefore, the electronic ground state of the system and its properties can be fully and uniquely expressed by the electron density  $\rho$  within one additive constant  $V_{\text{nuc,nuc}}$ . For the electronic energy, it follows

$$E_{\text{el}} = T_{\text{el}}[\rho] + V_{\text{el,el}}[\rho] + V_{\text{nuc,el}}[\rho] \quad (2.2.4)$$

$$= \underbrace{\langle \Psi[\rho] | \hat{T}_{\text{el}}[\rho] + \hat{V}_{\text{el,el}}[\rho] | \Psi[\rho] \rangle}_{F_{\text{HK}}[\rho]} + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \quad (2.2.5)$$

Replacing the external potential  $v(\mathbf{r})$  with an external potential  $V_0$  of a system with ground-state density  $\rho_0$  and ground-state energy  $E_0$ , it follows from the variational principle,

$$\begin{aligned} E_{v_0}[\rho] = E_{v_0}[\Psi[\rho]] &> E_0 && \text{with } \rho \neq \rho_0 \\ E_0 = E_{v_0}[\Psi[\rho_0]] &= E_{v_0}[\rho_0] \end{aligned} \quad (2.2.6)$$

and therefore conclude that the ground-state energy can be obtained by minimization of  $E_{v_0}[\rho]$ . This leaves us with the Hohenberg–Kohn variational principle, which is considered the second theorem, however, without knowledge of the exact form of  $F_{\text{HK}}[\rho]$ .

Furthermore, in the Hohenberg–Kohn theorems, it is assumed that the electron density is  $v$ -representable: It can be obtained from a single antisymmetric wave function corresponding to a ground-state Hamiltonian that contains terms for the electron–electron interaction and a local external potential. However, this is not the case for all non-negative trial densities and the general conditions for  $v$ -representability are unknown. To extend  $F_{\text{HK}}$  to arbitrary non-negative densities, Levy<sup>[155]</sup> and Lieb<sup>[151, 152]</sup> developed the so-called Levy–Lieb functional

$$F_{\text{LL}}[\rho] = \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{\text{el,el}} | \Psi \rangle, \quad (2.2.7)$$

where the density  $\rho(\mathbf{r})$  is obtained for the minimum of all antisymmetrized  $N$ -particle wave functions, thus, is  $N$ -representable. For any  $v$ -representable electron density  $\rho(\mathbf{r})$ , the Levy–Lieb functional is equal to the Hohenberg–Kohn functional. Performing the constrained search in Eq. (2.2.7) and subsequently minimizing

$$E_0 = \min_{\rho} E_v[\rho] \quad (2.2.8)$$

$$= \min_{\rho} \left( F_{\text{LL}}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \right) \quad (2.2.9)$$

is called the Levy–Lieb constrained search formalism. The potential term is taken out the constraint minimization in Eq. (2.2.7) since only those wave functions are considered that yield the same electron density. By this two-step procedure, the search

in Eq. (2.2.5) over all  $v$ -representable electron densities is circumvented and solely requires  $N$ -representability of the electron density. This condition is satisfied if<sup>[156]</sup>

$$\rho(\mathbf{r}) \geq 0, \quad \int \rho(\mathbf{r}) d\mathbf{r} = n \quad \text{and} \quad \int |\nabla \rho(\mathbf{r})|^2 d\mathbf{r} < \infty. \quad (2.2.10)$$

### 2.2.1. Kohn-Sham Density-Functional Theory

The energy functional (Eq. (2.2.9)) can be minimized *via* the Euler-Lagrange equation for stationary energies by introducing a Lagrangian multiplier  $\mu$

$$\mu = \frac{\delta E_v[\rho]}{\delta \rho(\mathbf{r})} \quad (2.2.11)$$

$$= \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} + v(\mathbf{r}). \quad (2.2.12)$$

Here, for  $F[\rho]$  the functional derivative can employ the Hohenberg-Kohn form  $F_{\text{HK}}[\rho]$  or an extended form, for instance, the Levy-Lieb form  $F_{\text{LL}}[\rho]$ . Aforementioned approaches such as the Thomas-Fermi model and extended forms early on attempted the minimization of energy functionals in the form of the Hohenberg-Kohn functional by expressing all terms solely as explicit functionals of the electron density. Their lack of referencing the  $N$ -particle wave function, however, leads to insufficient accuracy in expressing the kinetic and exchange energy.<sup>[146, 148]</sup>

In 1965, Kohn and Sham proposed an approach that combines the advantages of density-dependent and wave function approaches.<sup>[157]</sup> In Kohn-Sham (KS)-DFT besides the real, interacting system with the ground-state electron density  $\rho_0$ , there exists a non-interacting reference system  $S$  with that same ground-state electron density  $\rho_0^s = \rho_0$ . For the ground state of a non-interacting system with a local external potential, the wave function can be expressed in terms of a Slater determinant (SD) of one-electron functions  $\psi_i$

$$\Psi^s = \Psi^{\text{SD}} = \frac{1}{\sqrt{n!}} |\psi_1^s \dots \psi_n^s|. \quad (2.2.13)$$

with the number of electrons  $n$ . By reintroducing this wave function, additionally, the expression for the kinetic energy is known exactly for the non-interacting system,

$$T^s[\rho] = \sum_i^N \left\langle \psi_i^s \left| -\frac{1}{2} \nabla^2 \right| \psi_i^s \right\rangle. \quad (2.2.14)$$

If the proposed non-interacting system with the same ground-state energy exists, this means, that derived from the non-interactive system, accordingly the interactive system

can be expressed in terms of one-electron functions  $\psi_i$  (the KS orbitals). Subsequently, the electronic energy (Eq. (2.2.4)) can then be rewritten as

$$E^{\text{KS}}[\rho] = T^{\text{s}}[\rho] + J[\rho] + V_{\text{nuc,el}}[\rho] + \underbrace{(V_{\text{el,el}}[\rho] - J[\rho] + T[\rho] - T^{\text{s}}[\rho])}_{E_{\text{xc}}}, \quad (2.2.15)$$

where the so-called exchange–correlation (XC) energy functional was introduced, which includes a correction term to the electron–electron interaction effects, that are not considered in the Coulomb term ( $V_{\text{el,el}}[\rho] - J[\rho]$ ) as well as the correction term for the kinetic energy due to being described in a non-interacting reference system in contrast to the real system ( $T[\rho] - T^{\text{s}}[\rho]$ ). Returning to the Euler–Lagrange expression, for the constraint of orthonormal orbitals, the Lagrange multipliers are obtained as

$$\mu = \frac{\delta E^{\text{KS}}[\rho]}{\delta \rho(\mathbf{r})} \quad (2.2.16)$$

$$= \frac{\delta T^{\text{s}}[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta V_{\text{nuc,el}}[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \quad (2.2.17)$$

$$= \frac{\delta T^{\text{s}}[\rho]}{\delta \rho(\mathbf{r})} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} + \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \quad (2.2.18)$$

$$= -\frac{1}{2}\nabla^2 + \underbrace{v_{\text{nuc}}(\mathbf{r}) + v_{\text{coul}}(\mathbf{r}) + v_{\text{xc}}(\mathbf{r})}_{v_{\text{eff}}} \quad (2.2.19)$$

where all terms except  $v_{\text{xc}}(\mathbf{r})$  can be determined analytically and  $v_{\text{eff}}$  denotes the effective potential. In contrast to the approximated terms in “orbital-free” approaches (*e.g.* the Thomas–Fermi model), however, this term is commonly rather small.

Since the ground-state density is also known in terms of KS orbitals

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2, \quad (2.2.20)$$

the KS equations can be derived, as the following effective one-electron eigenvalue equations<sup>[156]</sup>

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right) \psi_i = \varepsilon_i \psi_i. \quad (2.2.21)$$

$$(2.2.22)$$

This minimization reflects the projection of the real system onto a non-interacting single-determinant system. Since all approximations were corrected with the XC functional  $E_{\text{xc}}$  and therefore in the XC potential  $v_{\text{xc}}$ , KS-DFT is in principle exact. The central question remains the analytical form of the XC functional.

Over the years many different approximations have been developed that can be divided into a few main classes. Firstly, the XC functional based on the local density approximation (LDA) was introduced. With the following expression for the XC functional

$$E_{\text{xc}}^{\text{LDA}} = \int \varepsilon_{\text{xc}}^{\text{LDA}}[\rho]\rho(\mathbf{r})d\mathbf{r} = \int \varepsilon_{\text{x}}^{\text{LDA}}[\rho]\rho(\mathbf{r})d\mathbf{r} + \int \varepsilon_{\text{c}}^{\text{LDA}}[\rho]\rho(\mathbf{r})d\mathbf{r} \quad (2.2.23)$$

the energy density  $\varepsilon_{\text{xc}}$  is further divided into the exchange  $\varepsilon_{\text{x}}$  and correlation part  $\varepsilon_{\text{c}}$ . The former, which is commonly contributing more can be expressed by the exchange terms of the uniform electron gas.<sup>[148]</sup> In this Thomas–Fermi–Dirac model small volume elements of locally homogenous density form the full electron density (cf. Eq. (2.2.2)), so that  $\varepsilon_{\text{x}} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\rho^{\frac{1}{3}}$ .<sup>[148]</sup> The correlation energy density can be obtained by fitting a functional expression to Monte-Carlo simulations of the uniform electron gas.<sup>[158-160]</sup>

Extending this class to the so-called generalized gradient approximation (GGA), the XC functional becomes<sup>[161]</sup>

$$E_{\text{xc}}^{\text{GGA}} = \int \varepsilon_{\text{xc}}^{\text{GGA}}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]\rho(\mathbf{r})d\mathbf{r}. \quad (2.2.24)$$

Here, the energy density shows a dependence on both the electron density and its gradient, due to a Taylor expansion of the energy to second order. Due to the better description of the XC hole, GGA functionals yield improved results in comparison to LDA functionals.<sup>[161]</sup> The GGA is often referred to as semi-local in contrast to the local LDA and commonly reduces the error from LDA by approximately one order of magnitude.<sup>[162]</sup> Popular representatives are Becke’s exchange functional B88<sup>[163]</sup>, the correlation functional LYP by Lee, Yang and Parr<sup>[164]</sup> or the XC functional PBE by Perdew, Burke and Ernzerhof<sup>[165]</sup>. Including the terms of the second-order Taylor expanded energy systematically improves the description in the so-called meta-GGA functionals but to a minor degree.<sup>[162]</sup>

Being derived from the other classes, the hybrid functionals are combining the density-dependent KS exchange and correlation part with a non-local exchange term from HF theory. Due to the lack of accuracy when approximating the exchange, DFT suffers from a self-interaction error of the electron that results from the Coulomb term, while in the HF theory this is explicitly cancelled out by the exchange terms. Therefore, the latter offers an exact expression for the exchange energy, although, being based on a different type of orbitals. Thus, non-local “exact exchange” is introduced as  $E_{\text{x}}^{\text{HF}}$  using the adiabatic connection to relate the XC functional to the non-interacting electron density so that the XC functional becomes

$$E_{\text{xc}}^{\text{H}} = (1 - c_{\text{HF}})E_{\text{x}}^{\text{KS}}[\rho, \nabla\rho] + c_{\text{HF}}E_{\text{x}}^{\text{HF}}[\psi_i] + E_{\text{c}}^{\text{KS}}[\rho, \nabla\rho], \quad (2.2.25)$$

where  $\psi_i$  are the occupied KS orbitals.<sup>[162]</sup> The amount of necessary exact exchange can differ greatly between different properties and systems so that a scaling coefficient  $c_{\text{HF}}$  is introduced. There exists a plethora of representatives, however, the most popular and widely used hybrid XC functional is B3LYP that is combined of the GGA-type B88

exchange functional and LYP correlation functional.<sup>[159, 166, 167]</sup> In this special case, Eq. (2.2.25) is adapted to

$$E_{xc}^{B3LYP} = (1 - c_{HF})E_x^{LDA} + c_{HF}E_x^{HF} + bE_x^{B88} + (1 - c)E_c^{LDA} + cE_c^{LYP} \quad (2.2.26)$$

with  $c_{HF} = 0.2$ ,  $b = 0.72$  and  $c = 0.81$ .<sup>[168]</sup>

Since for long-range electron interactions the HF exchange displays better asymptotic behaviour (also due to the self-interaction error), in so-called range-separated/Coulomb-attenuated functionals for this kind of interactions HF exchange is employed.<sup>[169-173]</sup> To obtain this range-dependent blending in of exact exchange, the Coulomb operator is split into a KS and a HF part, as shown here on the example of the popular CAM-B3LYP functional

$$\frac{1}{r_1 - r_2} = \frac{1 - [\alpha + \beta \cdot \text{erf}(\mu(r_1 - r_2))]}{r_1 - r_2} + \frac{\alpha + \beta \cdot \text{erf}(\mu(r_2 - r_1))}{r_1 - r_2} \quad (2.2.27)$$

with  $\alpha = 0.19$ ,  $\beta = 0.46$ ,  $\mu = 0.33$  and the error function  $\text{erf}$ .<sup>[174]</sup> Here,  $\alpha$  and  $\beta$  determine the amount of short-range and long-range exact exchange, respectively. It builds on the popular B3LYP functional and was one of the first functionals to introduce the interaction switching parameter  $\mu$  that determines when to switch between short- and long-range parts.<sup>[161]</sup>

In principle, the introduction of more accurate terms, as done for the exchange part in hybrid functionals, is also possible for the correlation part. This is commonly done by combining the density-dependent KS term and a scaled term obtained with the second-order Møller-Plesset perturbation theory equation

$$E^{DH} = (1 - c_{HF})E_x^{KS} + c_{HF}E_x^{HF} + (1 - c_{MP2})E_c^{KS}[\rho, \nabla\rho] + c_{MP2}E_c^{MP2}[\psi_i, \psi_a] \quad (2.2.28)$$

that depends on the occupied ( $\psi_i$ ) and virtual orbitals ( $\psi_a$ ).<sup>[161, 175]</sup> An early representative is the B2PLYP functional by Grimme.<sup>[175]</sup>

Additionally, there exist, semiempirical “3c” methods, that are derived from other usually GGA functionals, including three correction terms originating from the HF-3c and PBEh-3c methods.<sup>[176, 177]</sup> B97-3c<sup>[178]</sup> is a popular representative, that stems from the regular B97 XC functional<sup>[179]</sup> but involves correction terms for the basis set superposition error by a geometrical counterpoise correction, a Grimme D3 dispersion correction as well as a correction for short-range correction for estimated bond lengths. The method is applied with a special modified Ahlrichs basis set has been found to be robust while showing high efficiency.

### 2.3. Time-Dependent Density-Functional Theory

In order to describe optical properties such as excitation energies, it is necessary to be able to describe the system in response to external radiation. Following the semi-classical approach, the interaction between this radiation and the system is assumed to origi-

nate from a classical electromagnetic field interacting with the quantum-mechanically treated system. In the case of a molecule, this concerns the interaction of the field with the molecule's dipole moment  $\boldsymbol{\mu}(\mathbf{r})$  and consequently a change in the molecular electron density.<sup>[180]</sup> The frequency-dependence of this change in the electron density is particularly interesting and is connected to the time-dependent frame *via* Fourier transformation. For instance, the frequency-dependent polarizability is an essential quantity that describes the change of the dipole moment in interaction with an electromagnetic field and therefore enables to, *inter alia*, determine excitation energies. The calculation of these properties in response to an external electromagnetic potential can be realized with so-called TD-DFT. In this framework that is described below (closely the following the derivation in ref. [181]), KS-DFT is extended in order to describe the time-dependent change of the electron density.

In the previous chapter, the Hohenberg–Kohn theorems were introduced (Section 2.2). Although a one-to-one correspondence between the ground-state electron density and the wave function of the system was established, and thus, also for excited states, the introduction of the external potential does not necessarily follow a unique correspondence to excited-state electron densities. This formal basis was established in the Runge–Gross theorem<sup>[182]</sup>, which extends the Hohenberg–Kohn theorems to a one-to-one correspondence between the time-dependent electron density and external potential. Subsequently, a time-dependent version of the Kohn–Sham equations is obtained as,

$$\underbrace{\left(-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r}, t)\right)}_{\hat{f}^{\text{KS}}} \psi_i(\mathbf{r}, t) = i\frac{\partial}{\partial t}\psi_i(\mathbf{r}, t). \quad (2.3.1)$$

Linearly expanding the one-electron wave functions  $\psi_i(\mathbf{r}, t) = \sum_{\mu}^n c_{i\mu}(t)\chi_{\mu}(\mathbf{r})$  with time-independent single-particle wave functions  $\chi_{\mu}(\mathbf{r})$  the electron density can be defined in terms of expansion coefficients

$$\rho(\mathbf{r}, t) = \sum_i^n |\psi_i(\mathbf{r}, t)|^2 = \sum_i \left( \sum_{\mu} c_{i\mu}(t)\chi_{\mu}(\mathbf{r}) \sum_{\nu} c_{i\nu}^*(t)\chi_{\nu}(\mathbf{r}) \right) \quad (2.3.2)$$

$$= \sum_{\mu\nu} P_{\mu\nu}\chi_{\mu}(\mathbf{r})\chi_{\nu}^*(\mathbf{r}). \quad (2.3.3)$$

with the density matrix element  $P_{\mu\nu} = \sum_i^n c_{i\mu}c_{i\nu}^*$ . With this expression and the Liouville–von Neumann equation equation

$$i\frac{\partial}{\partial t}\mathbf{P} = [\mathbf{F}, \mathbf{P}] \quad (2.3.4)$$

the time-dependent Schrödinger equation (Eq. (2.1.1)) can be rewritten as

$$i\frac{\partial}{\partial t}P_{pr} = \sum_q (F_{pq}P_{qr} - P_{pq}F_{qr}), \quad (2.3.5)$$

where  $p, q, \dots$  are defined as general orbitals and  $F_{pq}$  as Fock matrix elements.

Occupied and virtual orbitals will in the following be denoted as  $i, j, \dots$  and  $a, b, \dots$ , respectively. From here, linear-response TD-DFT is applied, where the time-dependent external electric field is applied to a time-independent unperturbed ground-state electron density.

When applying this field, the density matrix and Fock matrix experience a linear (up to first-order) perturbation

$$P_{pq} = P_{pq}^{(0)} + P_{pq}^{(1)} \quad (2.3.6)$$

$$F_{pq} = F_{pq}^{(0)} + F_{pq}^{(1)}. \quad (2.3.7)$$

with the unperturbed and first-order perturbed Fock matrix elements  $F^{(0)}$  and  $F^{(1)}$  and the unperturbed and first-order perturbed density matrix elements  $P^{(0)}$  and  $P^{(1)}$ , respectively. When these expressions are inserted into Eq. (2.3.5) with zeroth-order terms vanishing due to them being unperturbed in the ground state and collecting the linear terms, the first-order change of the time-dependent Kohn–Sham equations becomes

$$\sum_q \left( F_{pq}^{(0)} P_{qr}^{(1)} - P_{pq}^{(1)} F_{qr}^{(0)} + F_{pq}^{(1)} P_{qr}^{(0)} - P_{pq}^{(0)} F_{qr}^{(1)} \right) = i \frac{\partial}{\partial t} P_{pr}^{(1)}. \quad (2.3.8)$$

The time-dependent change of the density matrix to first order is defined as

$$P_{pq}^{(1)} = \frac{1}{2} \left( d_{pq} e^{-i\omega t} + d_{qp}^* e^{i\omega t} \right), \quad (2.3.9)$$

with the oscillation amplitudes  $d_{pq}$ . The first-order change of the Fock matrix is comprised of the applied perturbation by the time-dependent electric field and the changes of the two-electron part due to the changes of the density matrix,

$$F_{pq}^{(1)} = V_{pq}(t) + \sum_{st} \frac{\partial F_{pq}^{(0)}}{\partial P_{st}} P_{st}^{(1)}. \quad (2.3.10)$$

Using Eq. (2.3.6) and Eq. (2.3.7) in the time-dependent Schrödinger equation in Dirac form Eq. (2.3.8) and assuming the external time-dependent potential to be of the form  $\hat{V}(t) = \frac{1}{2} (V_0 e^{-i\omega t} + V_0^* e^{i\omega t})$ , the following expression is obtained for the terms that are multiplied with  $e^{-i\omega t}$

$$\sum_q \left[ F_{pq}^{(0)} d_{qr} - d_{pq} F_{qr}^{(0)} + V_{0,pq} P_{qr}^{(0)} + \sum_{st} \frac{\partial F_{pq}^{(0)}}{\partial P_{st}} d_{st} P_{qr}^{(0)} - P_{pq}^{(0)} V_{0,qr} - P_{pq}^{(0)} \sum_{st} \frac{\partial F_{qr}^{(0)}}{\partial P_{st}} d_{st} \right] = \omega d_{qr}. \quad (2.3.11)$$



The terms multiplied by  $e^{i\omega t}$  give the complex conjugated of Eq. (2.3.11). Further simplifications can be made due to the fact that the unperturbed Fock matrix and density matrix are diagonal and solely oscillation amplitudes between occupied and virtual orbitals contribute ( $d_{ii} = d_{aa} = 0$ ). With the property of canonical orbitals being orthonormal, the unperturbed Fock matrix elements and density matrix elements can be reduced to  $F_{pq}^{(0)} = \delta_{pq}\varepsilon_p$  and  $P_{ij}^{(0)} = \delta_{ij}$ , respectively, with the orbital energies  $\varepsilon_p$ . Density matrix elements involving virtual orbitals do not contribute ( $P_{ia}^{(0)} = P_{ai}^{(0)} = P_{ab}^{(0)} = 0$ ). Introducing these expressions into Eq. (2.3.11) and its complex conjugate the following non-Hermitian eigenvalue equation is obtained

$$\left[ \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = - \begin{pmatrix} \mathbf{K} \\ \mathbf{L} \end{pmatrix}. \quad (2.3.12)$$

with

$$\begin{aligned} A_{ai,bj} &= \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \frac{\partial F_{ai}^{(0)}}{\partial P_{bj}} & A_{ai,bj}^* &= \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \frac{\partial F_{ia}^{(0)}}{\partial P_{bj}} \\ B_{ai,bj} &= \frac{\partial F_{ai}^{(0)}}{\partial P_{jb}} & B_{ai,bj}^* &= \frac{\partial F_{ia}^{(0)}}{\partial P_{jb}} \\ X_{ai} &= d_{ai} & Y_{ai} &= d_{ia} \\ K_{ai} &= V_{0,ai} & L_{ai} &= V_{0,ia}, \end{aligned}$$

the so-called the TD-DFT equations. Assuming resonance conditions and thus, an infinitesimal perturbation  $V_{0,pq} = 0$ , Eq. (2.3.12) can be further reduced to the non-Hermitian eigenvalue equation,

$$\left[ \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = 0. \quad (2.3.13)$$

In these definitions, the only unknown ingredient is the derivatives of the unperturbed Fock matrix elements with respect to the density matrix elements  $\frac{\partial F_{pq}^{(0)}}{\partial P_{rs}}$ . It is assumed, that the electron density very slowly changes with time so that in this adiabatic local density approximation (ALDA) the time-dependent non-local XC kernel is replaced with a time-independent local counterpart and determined with XC functionals developed for time-independent calculations. In this case, the derivative reads,

$$\frac{\partial F_{pq}^{(0)}}{\partial P_{rs}} = \frac{\partial \langle \psi_p | \hat{f}^{\text{KS}} | \psi_q \rangle}{\partial P_{rs}} = \frac{\partial \langle \psi_p | \hat{T} + V_{\text{eff}}[\rho] | \psi_q \rangle}{\partial P_{rs}} = \frac{\partial \langle \psi_p | v_{\text{coul}}[\rho] + v_{\text{xc}}[\rho] | \psi_q \rangle}{\partial P_{rs}}, \quad (2.3.14)$$

where only the Coulomb and XC potential show a dependence on the electron density, thus, the kinetic energy and nuclear potential terms vanish in the derivative. Using functional derivative laws Eq. (2.3.14) is further reduced to

$$\frac{\partial F_{pq}^{(0)}}{\partial P_{rs}} = \langle \psi_p | \int \frac{\delta(v_{\text{coul}}[\rho] + v_{\text{xc}}[\rho])}{\delta\rho(\mathbf{r}_2)} \cdot \underbrace{\frac{\delta\rho(\mathbf{r}_2)}{\partial P_{rs}}}_{\rho_{rs}^t(\mathbf{r}_2)} d\mathbf{r}_2 | \psi_q \rangle \quad (2.3.15)$$

$$= \iint \rho_{pq}^t(\mathbf{r}_1) \left( \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{\delta^2 E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)} \right) \rho_{rs}^t(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2, \quad (2.3.16)$$

with the transition density  $\rho_{pq}^t = \psi_p(\mathbf{r})\psi_q(\mathbf{r})$ .

To determine the frequency-dependent polarizability, the TD-DFT equations out of resonance have to be applied (Eq. (2.3.12)). Generally, the polarizability is defined as the derivative of the induced dipole moment with respect to the frequency-dependent electric field

$$\alpha(\omega) = \frac{\partial(\delta\boldsymbol{\mu}^{\text{ind}})}{\partial\mathbf{E}(\omega)}. \quad (2.3.17)$$

It can be shown, that the polarizability tensor is defined by the Kramer-Heisenberg relation as

$$(\alpha_{uv})_{km} = \frac{1}{\hbar} \sum_n \left[ \frac{\langle \psi_k | \mu_u \psi_n \rangle \langle \psi_m | \mu_v \psi_n \rangle}{\omega + \omega_{nm}} - \frac{\langle \psi_n | \mu_u \psi_m \rangle \langle \psi_k | \mu_v \psi_n \rangle}{\omega - \omega_{nk}} \right] \quad (2.3.18)$$

with the dipole operator  $\boldsymbol{\mu}$  for the directions in space  $u, v \in \{x, y, z\}$  and the orbitals for a transition from  $\psi_k$  to  $\psi_n$ .<sup>[180, 183]</sup>  $\omega$  and  $\omega_{nk}$  denote the frequency of the perturbing field and the resonance frequency, respectively. It can be alternatively expressed as<sup>[184]</sup>

$$\alpha_{uv}(\omega) = 4 \sum_n \frac{\mathbf{D}_u^T \mathbf{p}_n \mathbf{p}_n^T \mathbf{D}_v}{\omega_n^2 - \omega^2} \quad (2.3.19)$$

with  $\mathbf{p}_n = \frac{1}{\sqrt{2}}(\mathbf{X} + \mathbf{Y})$ , the excitation frequencies  $\omega_n$  and the dipole matrix elements  $(D_u)_{ai} = \int \psi_a r_u \psi_i d\mathbf{r}$ . This expression holds for both, the resonant and the non-resonant case, however, involving a different set of eigenvectors in  $\mathbf{p}$  (cf. Eqs. (2.3.12) and (2.3.13)). Comparing to the Kramer-Heisenberg equation (Eq. (2.3.18)) allows to express the transition dipole moments as

$$\langle 0 | \mu_u | n \rangle = \frac{1}{\sqrt{\omega_{0n}}} \mathbf{D}_u^T \mathbf{p}_n, \quad (2.3.20)$$

that allow the solution of the TD-DFT equations.

The polarizability is also commonly expressed as the isotropic mean polarizability,

$$\bar{\alpha}(\omega) = \frac{1}{3} \sum_i \alpha_{ii}(\omega), \quad (2.3.21)$$

also defined as the averaged trace of the polarizability tensor. It can be rewritten as

$$\bar{\alpha}(\omega) = \sum_n \frac{2}{3} \frac{\omega_{nm} \mu_{nm}^2}{\omega_{nm}^2 - \omega^2} = \sum_n \frac{f_n}{\omega_n^2 - \omega^2}, \quad (2.3.22)$$

where  $f_n$  is the oscillator strengths, the residues of the mean polarizability, while the excitation energies can be identified as its poles.<sup>[185-187]</sup> With the help of Eq. (2.3.20) they can be determined from the eigenvectors of Eq. (2.3.12).

Following from this, at resonance (Eq. (2.3.13)), where  $\omega$  is the resonance frequency, the excitation energies are readily determined and the oscillator strengths can be calculated from the eigenvectors  $\mathbf{X}$  and  $\mathbf{Y}$ .

## 2.4. Subsystem-Based Approaches for Optical Properties

In the previous chapter, a background for obtaining optical properties was given. However, it referred to methods solely dealing with a single system that possibly is composed of a large number of molecules. The scaling for supermolecular systems increases prohibitively with the system size, specifically the number of electrons. Thus, dealing with large supermolecular systems such as chromophores embedded in proteins or metal-organic frameworks on a pure QM level becomes impossible. For the calculation of optical properties, however, QM approaches are indispensable.<sup>[12, 39, 42, 188-191]</sup> Here, subsystem-based approaches promise a circumvention of the scaling problem. The key idea is to partition the supermolecular system into smaller subsystems or fragments (used interchangeably) in physical space whose computation is less computationally-demanding.<sup>[38, 39]</sup> For this, full molecules but also parts of a molecule can be considered as subsystems.<sup>[43, 192]</sup> The challenge is to develop these approaches in a way, that in the best case, the system gives the relevant property to be calculated as accurately as for the unpartitioned system.

In this thesis, two main classes of such subsystem-based approaches will be distinguished. Firstly, in Section 2.4.1, methods that focus on the *local* optical properties of one QM-described main subsystem embedded in the remaining “environmental” subsystems that may also be described on a different footing. Secondly, approaches that target the calculation of *global*, *i.e.* supermolecular, optical properties *via* the combination of properties calculated for subsystems, that are introduced in Section 2.4.2.

### 2.4.1. Local Optical Properties

As mentioned above, obtaining optical properties becomes more challenging with the number of electrons in the system. Thus, the calculation of one supermolecular system can be divided into the more affordable calculations for several, individual subsystems. Describing absorption and emission phenomena, oftentimes the excitation is dominated by one main subsystem, that experiences changed properties due to the environmental influence. Compared to the isolated case, the main molecule can show strongly changed absorption/emission properties when embedded in the environment.<sup>[7-12]</sup> In contrast, the environmental subsystems will have similar (but not identical) properties in both cases. Two classes of local embedding schemes are presented in this chapter: advanced polarizable QM/MM and approximate QM/QM embedding schemes. Although starting from different starting points, the target of these schemes coincides and they attempt to describe the same effects. In this chapter these local embedding schemes will be presented, incorporating an accurate description of environmental molecules and polarization effects. At first, the density-based QM/QM embedding scheme is introduced, which relies on a subsystem DFT approach and includes QM effects in the environmental description (Section 2.4.1.1). Subsequently, the a QM/classical PE model is introduced, that treats the environment *via* localized multipoles and includes anisotropic dipole-dipole polarizabilities obtained from QM electron densities (Section 2.4.1.3).<sup>[193]</sup>

#### 2.4.1.1. Density-Based Embedding

In this thesis, density-based local embedding schemes are referring to models based on a subsystem density-functional theory (sDFT) formalism. Similarly, to the PE formulation (Section 2.4.1.3), in this approach, the main subsystem is described in a QM fashion. In contrast to the PE model, the environment is described *via* QM as well, however, not necessarily of the same type as for the main system. In the following, the sDFT formalism, which is formally exact within the limit of exact energy functional description, and a practical approximate approach, the so-called frozen-density embedding (FDE), are introduced.<sup>[194]</sup>

In sDFT the partitioning of the system is carried out as a partitioning of supermolecular electron density

$$\rho_{\text{tot}}(\mathbf{r}) = \sum_X \rho_X \quad (2.4.1)$$

where  $\rho_X$  is the corresponding subsystem's electron density in the supermolecule.<sup>[195-197]</sup> Assuming the subsystem to be a non-interacting particle system as well as *v*-representability of the subsystem electron densities, these can be expressed *via* subsystem orbitals  $\psi_i$ ,

$$\rho_X(\mathbf{r}) = \sum_{i \in X} |\psi_i(\mathbf{r})|^2, \quad (2.4.2)$$

where  $i$  is an electron in subsystem  $X$ . When deriving the KS equations (Eq. (2.2.21)) for sDFT, both the nuclear and Coulomb potential are linear in the electron density, thus, can be expressed additively as,

$$v_{\text{nuc}} = \sum_X v_{\text{nuc},X} = \sum_X \sum_{I \in X} -\frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \quad (2.4.3)$$

and

$$v_{\text{coul}}[\rho](\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' = \sum_X \int \frac{\rho_X(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' = \sum_X v_{\text{coul},X}[\rho_X](\mathbf{r}). \quad (2.4.4)$$

with the  $l$ th nucleus assigned to subsystem  $X$ . In comparison to isolated subsystems, where only intrafragmental Coulomb interactions occur, these expressions were extended to account for all electrostatic interactions of all possible subsystem combinations. The kinetic energy term  $T^s[\rho]$  (Eq. (2.2.14)), however, shows a dependence on the orbitals of the total system. Therefore, it is non-additive in the electron density. The additive part of the subsystem kinetic energy in terms of subsystem orbitals can be introduced as,

$$T_X^s[\rho_X] = \sum_{i \in X} \langle \psi_i^s | -\frac{1}{2} \nabla_i^2 | \psi_i^s \rangle. \quad (2.4.5)$$

If there is any interaction in between subsystems, non-orthogonality between different subsystems can arise, and in consequence, a QM correction term for the non-additivity is introduced,

$$T^{\text{s,nadd}}[\rho_{\text{tot}}, \rho_X] = T^s[\rho_{\text{tot}}] - \sum_X T_X^s[\rho_X], \quad (2.4.6)$$

that is added to the energy functional expression. Accordingly, for the XC functional a QM non-additive correction term arises

$$E_{\text{xc}}^{\text{nadd}}[\rho_{\text{tot}}, \rho_X] = E_{\text{xc}}[\rho_{\text{tot}}] - \sum_X E_{\text{xc},X}[\rho_X], \quad (2.4.7)$$

that incorporates higher-order intermolecular short-range XC interactions, for instance, dispersion effects. In total, the energy functional in sDFT reads<sup>[198]</sup>

$$\begin{aligned}
 E[\rho_{\text{tot}}] = \sum_X \left[ \underbrace{T_X^s[\rho_X] + \int v_{\text{nuc},X}(\mathbf{r})\rho_X(\mathbf{r})\mathbf{d}\mathbf{r} + \int v_{\text{coul},X}(\mathbf{r})\rho_X(\mathbf{r})\mathbf{d}\mathbf{r} + E_{\text{xc}}[\rho_X]}_{E_X^{\text{iso}}[\rho_X]} \right. \\
 + \underbrace{\sum_{Y \neq X} \left( \int v_{\text{nuc},Y}(\mathbf{r})\rho_X(\mathbf{r})\mathbf{d}\mathbf{r} + \frac{1}{2} \int v_{\text{coul},X}(\mathbf{r})\rho_Y(\mathbf{r})\mathbf{d}\mathbf{r} \right)}_{E_X^{\text{emb,elstat}}} \\
 \left. + \underbrace{T^{\text{s,nadd}}[\rho_{\text{tot}}, \rho_X] + E_{\text{xc}}^{\text{nadd}}[\rho_{\text{tot}}, \rho_X]}_{E_X^{\text{emb,nadd}}} \right], \quad (2.4.8)
 \end{aligned}$$

where the full expression was decomposed into the isolated components  $E_X^{\text{iso}}[\rho_X]$  of subsystem  $X$ , the electrostatic embedding components  $E_X^{\text{emb,elstat}}$ , that involve the electrostatic interactions between subsystem  $X$  and all other subsystems and finally, the non-additive embedding components  $E_X^{\text{emb,nadd}}$  as previously introduced in Eq. (2.4.6) and Eq. (2.4.7), that depend on the total electron density. Performing an Euler-Lagrange minimization of the energy functional with respect to the individual subsystem KS orbitals while keeping the other subsystems' electron densities fixed, yields the sDFT equivalent of the KS equations (Eq. (2.2.21)) as<sup>[194, 199]</sup>

$$\left( -\frac{1}{2}\nabla^2 + v_X^{\text{eff}}[\rho_X] + v_X^{\text{emb}}[\rho_X, \rho_{\text{tot}}] \right) \psi_{X,i}^{\text{KSCED}} = \epsilon_{X,i} \psi_{X,i}^{\text{KSCED}}. \quad (2.4.9)$$

the so-called Kohn-Sham equations with constrained electron densities (KSCED) that include the effective potential for the isolated subsystem  $v_X^{\text{eff}}$  and the embedding potential  $v_X^{\text{emb}}$ <sup>[58]</sup>

$$\begin{aligned}
 v_X^{\text{emb}}[\rho_X, \rho_{\text{tot}}] = & \left( \sum_{Y, Y \neq X} v_{\text{nuc},Y}(\mathbf{r}) \right) + v_{\text{coul}}[\rho_{\text{tot}} - \rho_X](\mathbf{r}) \\
 & + v^{\text{kin,nadd}}[\rho_X, \rho_{\text{tot}}](\mathbf{r}) + v^{\text{xc,nadd}}[\rho_X, \rho_{\text{tot}}](\mathbf{r}). \quad (2.4.10)
 \end{aligned}$$

which is comprised of all interactions by other subsystems with subsystem  $X$ . Up to this point the formulation is in principle exact and leads to the exact electron density in the case of exact functionals and simultaneous or self-consistent solution of the KSCED (Eq. (2.4.9)) (as a consequence of the dependence on the total electron density).<sup>[198, 199]</sup> Although the non-additive contributions are commonly rather small<sup>[58]</sup>, their exact form is unknown. In contrast to other embedding schemes, such as QM/MM<sup>[39, 42]</sup>, however, the major contributions emerging from the interactions with nuclei and electrons in the environment are treated exactly.

### 2.4.1.2. Frozen-Density Embedding

Next to other approaches<sup>[200, 201]</sup> Wesolowski and Warshel proposed the frozen-density embedding (FDE) scheme. It enables the solution of the KSCED and is also the approach focussed on in this work.<sup>[194]</sup> As suggested by the name, the main or *active* system ( $A$ ) is embedded in the frozen environment density, that might however be comprised of several subsystems ( $X$ )

$$\rho_{\text{tot}}(\mathbf{r}) = \rho_A + \rho_{\text{env}} = \rho_A + \sum_{X \neq A} \rho_X. \quad (2.4.11)$$

Initially, the environment density, which is interacting with the active system *via* the embedding potential, is set to a feasible guessed electron density. In reality, this often refers to a determination of the electron density from isolated subsystem KS DFT calculations. The density of the active subsystem is then calculated (“relaxed”) in the presence of this frozen environmental density by the obtained embedding potential. If the focus lies on the properties of the active system, it could be expected that this process already serves as a sufficient approximation. However, the inductive components of and with the environment have been neglected by the frozen treatment of the environmental densities which can introduce major changes to the interactions with the main system.<sup>[202]</sup> Therefore, it is possible to yield the environmental information from other models’ calculations.<sup>[203]</sup> Another approach to this problem is the circumvention by iteratively exchanging the active role in between all subsystems in order to relax all subsystems’ electron densities. When changing the role to a new subsystem the other subsystem’s electron densities, including the newly obtained density from the previously active subsystem, in turn induce a change on the next active system. When performing these so-called “Freeze-and-Thaw” cycles until self-consistency, the sDFT solutions can be recovered.<sup>[204]</sup>

Since the non-additive terms of the embedding potential are unknown, approximations for the XC part as well as for the kinetic energy have to be introduced. In the XC part, the usual approximations (Section 2.2), that are also employed for the isolated subsystem calculations, are extended to the non-additive part.<sup>[57, 58]</sup> Since for these approximations, that involve inter-subsystem interactions no supermolecular orbitals are available, the use of hybrid functionals, that calculate the exact change *via* orbitals, cannot be employed.

For the kinetic energy the same problems occur, however, this contribution vanishes if orthogonality between the subsystems is ensured.<sup>[201, 205]</sup> This can be obtained by different techniques, such as projection embedding<sup>[200]</sup> or the einsatz of additional Lagrangian multipliers.<sup>[201]</sup> In FDE this orthogonality is not necessarily given so that approximations have to be employed. Although the Thomas-Fermi approximation or other similar purely density-dependent approaches are usually too crude approximations for accurately determining the kinetic energy (Section 2.2), since the non-additive kinetic energy term is generally small, relatively good results can be obtained for weakly interacting systems with the PW91k functional.<sup>[204, 206-208]</sup> For more strongly interact-

ing systems, for instance, in covalently-bound systems, the description is qualitatively wrong.<sup>[209, 210]</sup>

In order to obtain optical properties, equal to the ground-state DFT calculations, the common Fock operator is used for the isolated subsystems (Eq. (2.3.1)), so that  $\hat{f}^{\text{KS}} = \hat{f}^{\text{iso}}$ , and is extended to the embedded Fock operator in Eq. (2.4.9) ( $\hat{f}^{\text{emb}}$ ) due to the last term, the embedding potential  $v^{\text{emb}}$ , so that

$$\hat{f}^{\text{emb}} = \hat{f}^{\text{iso}} + v^{\text{emb}}. \quad (2.4.12)$$

This operator can be employed in the derivation of the TD-DFT equations (Section 2.3) and is leading to changes due to including the embedding effects with the embedding potential in the unperturbed system. This affects the orbital energies in the **A** term that are determined for the unperturbed system as well as the derivatives of the unperturbed Fock matrix elements with respect to the density matrix in both the **A** and **B** terms (Eq. (2.3.12)). This type of TD-DFT treatment in FDE is called uncoupled FDE<sup>[211]</sup> since the excitonic coupling between the subsystems is neglected and solely the properties of the embedded active system are determined.<sup>[212]</sup> The lack of describing this coupling leads to a lack of the dynamical response of the environmental excitation effects that act on the embedded chromophore. There have been several works on coupled FDE to include the dynamical response into the FDE scheme, however, exceeding the local properties discussed here and therefore will not be addressed further in this context.<sup>[213-221]</sup>

#### 2.4.1.3. Polarizable Embedding

As a representative of QM/classical embedding schemes, in PE the main system is described via QM methods, while, contrasting to QM/QM embedding schemes (Section 2.4.1.1), the environment is described in a more approximate nature. Closely following refs. [193, 222, 223], the derivation in second quantization will be summarized in the following chapter, starting from the supermolecular Hamiltonian in terms of subsystems,

$$\hat{H} = \sum_X \hat{H}_X + \sum_{X<Y} \hat{V}_{XY}, \quad (2.4.13)$$

summing over the different subsystems  $X$  and  $Y$ .  $H_X$  denotes the Hamiltonian of the isolated subsystem  $X$ , while  $\hat{V}_{XY}$  determines the Coulomb interactions between different subsystems  $X$  and  $Y$ . The former is assumed to give the energy of the isolated systems as the eigenvalue for the individually orthonormal, non-overlapping subsystem wave functions. Consequently, the supermolecular energy is additively separable and the wave function can be expressed as a product of subsystem wave functions. The latter term in Eq. (2.4.13), the interaction Hamiltonian, is comprised of the interactions of nuclei in system  $X$  or  $Y$  with electrons in different subsystems, respectively, as well as the electron-electron and nuclear-nuclear interactions in between different subsystems.



Performing Rayleigh–Schrödinger perturbation theory on the supermolecular system comprised of isolated subsystems defining the inter-subsystem interactions as the perturbation of the system, the energy up to second order includes

$$E = E^{(0)} + E^{(1)} + E^{(2)} \quad (2.4.14)$$

$$= \sum_X \langle \Psi_{X,0} | \hat{H}_X | \Psi_{X,0} \rangle + \sum_{X<Y} \langle \Psi_{X,0} \Psi_{Y,0} | \hat{V}_{XY} | \Psi_{X,0} \Psi_{Y,0} \rangle \quad (2.4.15)$$

$$- \sum_{X<Y} \sum_{\substack{ij \\ i+j \neq 0}} \frac{\langle \Psi_{X,0} \Psi_{Y,0} | \hat{V}_{XY} | \Psi_{X,i} \Psi_{Y,j} \rangle \langle \Psi_{X,i} \Psi_{Y,i} | \hat{V}_{XY} | \Psi_{X,0} \Psi_{Y,0} \rangle}{(E_{X,i} + E_{Y,j}) - (E_{X,0} + E_{Y,0})},$$

where the superscript (0), (1) or (2) refers to the zeroth-, first- or second-order perturbation terms. The wave function of subsystem  $X$  is denoted as  $\psi_X$  with the subscript 0 or  $i$  referring to the ground state or excited state, respectively. The individual subsystem wave functions are assumed to be known and orthonormalized. Accordingly, the subsystem energy of the ground state or excited state  $i$  is determined as  $E_{X,0}$  or  $E_{X,i}$ , respectively. The first term in Eq. (2.4.14) is defined by the energy of the isolated subsystems. The second term describes the electrostatic interaction in between subsystems. In the last term it was used, that the wave functions of different subsystems are assumed to fulfil orthogonality so that the summation over excited states can be given as fragment contributions.<sup>[223]</sup> It represents the induced energy and the dispersion energy resulting from the contributions involving the simultaneous excitation of one and two fragments, respectively.

In the PE model as a local embedding model, the focus lies on the description of the main system's energy due to the interactions with the PE potential resulting from the surrounding fragments,

$$E_{\text{QM-PE}} = E_{\text{QM}} + E_{\text{PE}}, \quad (2.4.16)$$

where  $E_{\text{QM}}$  is the energy of the isolated main system and  $E_{\text{PE}}$  is the energy resulting from the interactions with the PE potential comprised of the electrostatic, induction and dispersion term,

$$E_{\text{PE}} = E_{\text{PE}}^{\text{es}} + E_{\text{PE}}^{\text{ind}} + E_{\text{PE}}^{\text{disp}}. \quad (2.4.17)$$

For this, the description of the environmental subsystems in the interaction Hamiltonian is adjusted to be classical *via* a multipole expansion leading to an alternative formu-

lation of the interaction Hamiltonian of the main subsystem A and the environmental subsystems as,

$$\begin{aligned}\hat{V}_{AY} &= \sum_{|k| \in 0}^{\infty} \frac{(-1)^{|k|}}{k!} (\hat{F}_{A,\text{nuc}}^{(k)} + \hat{F}_{A,\text{el}}^{(k)}) Q_Y^{(k)} \\ &= \sum_{|k| \in 0}^{\infty} \frac{(-1)^{|k|}}{k!} \left[ \sum_{I \in A} Z_I T_{AY}^{(k)}(\mathbf{R}_I) - \sum_{pq \in A} \left( \int \rho_{pq}(\mathbf{r}) T_{AY}^{(k)}(\mathbf{r}) d\mathbf{r} \right) \hat{E}_{pq} \right] Q_Y^{(k)},\end{aligned}\quad (2.4.18)$$

with the singlet excitation operators  $\hat{E}_{pq}$  describing the excitation from an orbital  $\psi_p$  to an orbital  $\psi_q$ .<sup>[224]</sup> The multi-index  $k = (k_x, k_y, k_z)$  has the norm  $|k| = k_x + k_y + k_z$ , that denotes the order of the expansion. Here,  $Q_Y^{(k)}$  is the multipole moment of order  $k$  for subsystem Y so that the multipole moment of order  $k = 0, 1, 2$  equals charges, dipoles and quadrupoles, respectively. The field operators  $\hat{F}_{A,\text{nuc}}$  and  $\hat{F}_{A,\text{el}}$  yield the electrostatic potential, electric field and field gradient for the order  $k = 0, 1, 2$ , respectively, for the electrons and nuclei in the main subsystem A. Both terms of the field operator contain the so-called interaction tensor elements

$$\mathbf{T}_{ij}^{(k)} = \nabla_j^k \frac{1}{|\mathbf{R}_j - \mathbf{R}_i|} \quad (2.4.19)$$

where  $k$  refers to the power of the partial derivative operator  $\nabla_j^k = \frac{\partial^{k_x+k_y+k_z}}{\partial x^{k_x} \partial y^{k_y} \partial z^{k_z}}$  as well as the rank of the tensor which is obtained by a Taylor expansion at the expansion point  $R_O$ , the center of mass of a subsystem.<sup>[193, 225]</sup> The convergence of the multipole expansion is dependent on the distance between the expansion point  $R_O$  and  $\mathbf{R}_j$ , where short distances have been shown to lead to convergence problems.<sup>[222]</sup> The distribution of the multipole from the center of mass of a subsystem to atomic sites (and possibly bond-midpoints) in the subsystem improves this behaviour.<sup>[222, 226]</sup> The resulting electrostatic energy  $E^{\text{es}}$  (Eq. (2.4.18)) referring to the electrostatic interactions of the main subsystem A and the environmental subsystems can therefore be expressed in terms of these sites  $s$  as follows,

$$E^{\text{es}} = \sum_{|k|} \sum_{s=1} \frac{(-1)^{|k|}}{k!} \left( \hat{F}_{s,\text{nuc}}^{(k)} + \langle 0 | \hat{F}_{s,\text{el}}^{(k)} | 0 \rangle \right) Q_s^{(k)}, \quad (2.4.20)$$

where  $|0\rangle$  is the ground state and the sum is over the order of the expansion  $|k|$ , that denotes the order of the expansion and allows for a truncation.<sup>[223]</sup>

Employing the multipole expansion to further express the induction energy of the environmental multipoles on the active system, it is obtained as,

$$E^{\text{ind}} = -\frac{1}{2} \sum_{s=1} (F_{\text{nuc},s} + F_{\text{el},s} + F_{\text{mul},s}) + \boldsymbol{\mu}_s^{\text{ind}}(\mathbf{F}_{\text{tot}}) \quad (2.4.21)$$

with the field originating from the active subsystems nuclei  $\mathbf{F}_{\text{nuc}}$  and electrons  $\mathbf{F}_{\text{el}}$  or other multipoles in the environment  $\mathbf{F}_{\text{mul}}$ . The induced dipole moment  $\boldsymbol{\mu}^{\text{ind}}$  is dependent on the total field that additionally to the aforementioned contributions includes the electric field from other induced dipoles<sup>[222]</sup>

$$\boldsymbol{\mu}^{\text{ind}} = \sum_s \boldsymbol{\mu}_s^{\text{ind}} = \sum_s \boldsymbol{\alpha}_s \mathbf{F}_{\text{tot}} = \boldsymbol{\alpha}_s (\mathbf{F}_{\text{nuc}} + \mathbf{F}_{\text{el}} + \mathbf{F}_{\text{mul}} + \mathbf{F}_{\text{ind}}(\boldsymbol{\mu}_{s'}^{\text{ind}})) = \mathbf{R} \mathbf{F} \quad (2.4.22)$$

with the polarizability  $\boldsymbol{\alpha}_s$  at site  $s$ . The response matrix  $\mathbf{R}$  is defined as

$$\mathbf{R} = \begin{pmatrix} \boldsymbol{\alpha}_1^{-1} & -\mathbf{T}_{12}^{(2)} & \dots & -\mathbf{T}_{1S}^{(2)} \\ -\mathbf{T}_{21}^{(2)} & \boldsymbol{\alpha}_2^{-1} & \dots & -\mathbf{T}_{2S}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{T}_{S1}^{(2)} & -\mathbf{T}_{S2}^{(2)} & \dots & \boldsymbol{\alpha}_S^{-1} \end{pmatrix}^{-1} \quad (2.4.23)$$

with the inverse polarizabilities as diagonal elements and interaction tensors between sites as off-diagonal elements. Since the induced electric field depends on the induced dipole moment from other sites  $s'$  the equation has to be solved iteratively, *e.g.* via the Jacobi method.<sup>[222]</sup>

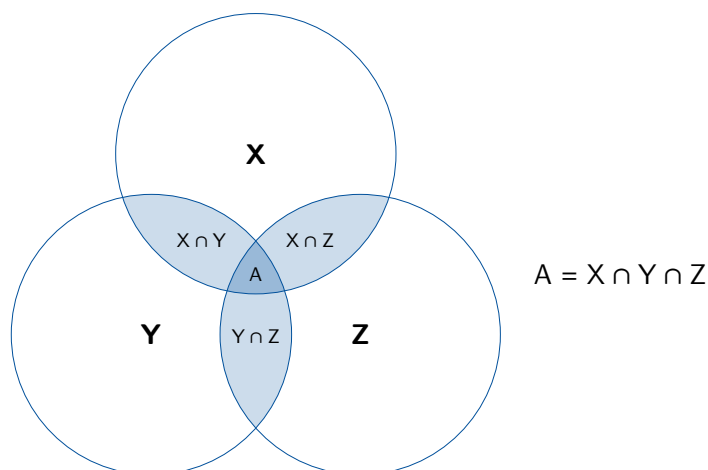
The minimization of the obtained energy with respect to the electron density gives an effective Fock operator

$$\hat{f}^{\text{emb}} = \hat{f}^{\text{iso}} + v^{\text{emb}} \quad (2.4.24)$$

with the KS Fock operator for the isolated active subsystem and an embedding potential, that reflects the changes due to embedding the active subsystem in the environment. In order to obtain local optical properties for the active system a linear-response TD-DFT ansatz (see Section 2.3) is performed. Inserting the effective Fock operator instead of the isolated KS Fock operator into the systems the density-dependent terms in the embedding potential lead to extra contributions changing the matrices of the unperturbed system. These terms include the linear response of the static environment, and additionally, the dynamical response of the environment to the perturbation.

### 2.4.2. Global Optical Properties

The above-mentioned reduction of computational effort in embedding schemes is rooted in the partitioning of the supermolecular system into smaller subsystems. In contrast to the previously presented local embedding schemes, where the focus lies on a main subsystem influenced by the environment, subsystem-based methods can also be employed to determine a global system property. Focussing on the energy-based fragmentation formalism, the supermolecular energy can be expressed in terms of individual subsystems or “monomers” and overlapping subsystems. It is composed of isolated monomer energies  $E_A$  and their interactions. The interaction in the energy expression can be further structured into different orders, that originate from the interactions obtained



**Figure 2.4.1.:** Schematic representation of subsystems  $X$ ,  $Y$ ,  $Z$  and their intersections as used in Eq. (2.4.26).

for two combined subsystems (dimers), for three combined subsystems (trimers) and so forth. These terms are denoted as two-body, three-body terms and higher-order terms so that in an infinite “many-body expansion (MBE)” the exact supermolecular energy can be recovered due to the full consideration of interactions. In this ansatz, the so-called bottom-up (BU) approach, the first term consists of the sum of non-interacting subsystem energies.<sup>[44]</sup> Interactions are added sequentially, starting with the two-body terms so that the full energy expression can be formulated as,<sup>[227-229]</sup>

$$\begin{aligned}
 E &= \sum_A E_A + \sum_{A<B} \underbrace{[E_{AB} - (E_A + E_B)]}_{\Delta_{AB}} \\
 &+ \sum_{A<B<C} \underbrace{[E_{ABC} - (\Delta_{AB} + \Delta_{AC} + \Delta_{BC}) - (E_A + E_B + E_C)]}_{\Delta_{ABC}} + \dots \\
 &= \sum_A E_A + \sum_{A<B} \Delta_{AB} + \sum_{A<B<C} \Delta_{ABC} + \dots,
 \end{aligned} \tag{2.4.25}$$

where  $E_A$ ,  $E_B$  and  $E_C$  represents the energy of a specific monomeric subsystem  $A$ ,  $B$  or  $C$ .  $E_{AB}$  and  $E_{ABC}$  denote the energy of  $A$  and  $B$  and  $A$ ,  $B$  and  $C$  calculated from the combined subsystems as a dimer and trimer, respectively. The differences between those higher-body terms (*e.g.*  $E_{AB}$ ) and the lower-body terms (*e.g.*  $E_A$ ), that are denoted  $\Delta$ , therefore, annihilate the double-counting of previously introduced terms and serve as the incremental addition of interactions towards the supermolecular energy.

Additionally, so-called top-down (TD) approaches exist, where the initial terms are built from a combination of subsystems presenting an inter-subsystem overlap, which is subtracted in the energy expression.<sup>[44, 92]</sup> An exemplary system of three overlapping subsystems is schematically presented in Fig. 2.4.1. Formally, these terms can be

expressed in the general MBE expression of disjunct fragments in Eq. (2.4.25).<sup>[80]</sup> Alternatively, the energy expression can be formulated as<sup>[93, 230-233]</sup>

$$E = \sum_X E_X - \sum_{X<Y} E_{X\cap Y} + \sum_{X<Y<Z} E_{X\cap Y\cap Z} - \dots \quad (2.4.26)$$

with the energy of the overlap of two subsystems  $X$  and  $Y$ ,  $E_{X\cap Y}$  and energy of the overlap of three subsystems  $X$ ,  $Y$ ,  $Z$  denoted  $E_{X\cap Y\cap Z}$  (Fig. 2.4.1). This expression can coincide with the expression resulting from the BU approach depending on the selection of fragments.<sup>[82]</sup> Although the total energy expressions are ultimately the same for the full interactions considered, this does not transfer to the calculation behaviour for the truncated terms. The important difference for these “intersecting-fragment” (TD) methods is, that due to the employment of overlapping subsystems in the first terms, a faster convergence is obtained.<sup>[92, 93, 234]</sup>

These TD approaches, as well as the MBE-based BU approaches, are easier to implement for systems that do not show covalent bonding due to the occurrence of dangling bonds in the fragmentation process. Methods accounting for these ends have been suggested (for proteins and peptides), such as the molecular fractionation with conjugated caps (MFCC) approach<sup>[235, 236]</sup>, but are disregarded here due to the focus on fragmentation of coordinative systems. Heterolytic cutting of coordinative bonds is also critically evaluated<sup>[80, 236, 237]</sup>, specifically in metal-ligand bonds. This can be circumvented by keeping metal-ligand bonds intact<sup>[238]</sup>, although proving impractical for efficiently-sized fragments. Moreover, the performance of the fragmentation with cut coordinative bonds was found to perform well for optical properties for linker-dominated MOFs as well as for TD approaches that at least partially include fragments with intact bonds.<sup>[129, 239]</sup>

#### 2.4.2.1. Optical Properties from Energy-Based Fragmentation Schemes

So far fragmentation methods have been introduced allowing the recombination of the subsystem energies to the supermolecular energy. As previously mentioned (Chapter 1 and Section 2.3), the electric polarizability  $\alpha$ , a central optical property, allows the derivation of further properties, such as the refractive index. It can also be defined as the second-order derivative of the energy with respect to an electric field  $\mathbf{F}$ ,

$$\alpha_{uv} = -\frac{\partial^2 E}{\partial F_u \partial F_v}. \quad (2.4.27)$$

For weak electric fields, *i.e.* close to  $\mathbf{F}=0$ , if the energy of the supermolecular system can be obtained from subsystem energies *via* fragmentation due to the sum rule of derivatives, additivity is assumed for derivatives of the energy. For fragmentation schemes that do not include any molecular interactions, this additivity hypothesis can be made, that was applied to molecular<sup>[240]</sup> and non-molecular crystals<sup>[241, 242]</sup>. Furthermore, the examination of hydrogen-bonded urea clusters proved a rather small influence of intermolecular interactions on the linear polarizability in contrast to intramolecular

chemical bonds<sup>[243]</sup>. The additivity of the polarizability was also applied to derive related properties of minerals.<sup>[244, 245]</sup> Thus, Eqs. (2.4.25) and (2.4.26) can be adapted to extrapolate the isotropic polarizability (Eq. (2.3.22)) for both, the BU and TD fragmentation approach. For the most simple BU approach, composed of monomer terms, it follows,<sup>[129]</sup>

$$\alpha^{\text{BU}} = \sum_A \alpha_A. \quad (2.4.28)$$

This approach does, however, not include any many-body effects as presented in Eq. (2.4.25).

The individual subsystem polarizabilities are accessible *via* DFT, where two main approaches can be distinguished: the linear-response and coupled-perturbed approaches.<sup>[246-249]</sup> The former was presented in Section 2.3, where the frequency-dependent, *i.e.* dynamic polarizability is determined from the non-resonant case (Eq. (2.3.19)) since the resonant case refers to frequencies, where the polarizability shows singularities (*cf.* Section 2.3). The coupled-perturbed approach is based on the analytic calculation of the second-order energy derivatives based on the perturbation of the wave function. The term “coupled-perturbed” refers to the Fock matrix and the density matrix, since the derivative Fock matrix depends on the density matrix *via* the molecular orbital coefficients. After the construction of the derivative Fock matrix, the density matrix is updated so that a new derivative Fock matrix can be constructed, leading to the iterative coupled-perturbed self-consistent field (CP-SCF) approach.<sup>[246, 248, 249]</sup>

Having obtained the total polarizability from DFT calculations and fragmentation schemes, the frequency-dependent refractive index can be simply determined by the Lorenz-Lorentz equation<sup>[130, 131]</sup>,

$$\frac{n^2 - 1}{n^2 + 2} = \frac{N\alpha}{3\epsilon_0} \Leftrightarrow n = \sqrt{\frac{1 + 2\frac{N\alpha}{3\epsilon_0}}{1 - \frac{N\alpha}{3\epsilon_0}}}, \quad (2.4.29)$$

with the number density  $N$  and the dielectric constant  $\epsilon_0$ . This procedure has been widely used in order to determine the refractive index from the polarizability obtained with DFT calculations.<sup>[129, 250-258]</sup> Using that the polarizability volume is defined as  $\alpha' = \frac{\alpha}{4\pi\epsilon_0}$ , the equation can be rewritten to

$$n = \sqrt{\frac{1 + 2\frac{N\alpha'4\pi}{3}}{1 - \frac{N\alpha'4\pi}{3}}}. \quad (2.4.30)$$

The polarizability volume can directly be obtained from the fragmentation scheme and the above-mentioned DFT calculations.

# **Part II.**

## **Research**





## 3 Computational Details

In this chapter, the general computational procedures and technical details employed in this work are presented in two main sections. First, the computational details for the calculation of local optical properties are presented. Subsequently, the computational details for the calculation of global optical properties are shown.

### 3.1. Computational Details for Calculations of Local Optical Properties<sup>†</sup>

#### Model generation

The target systems (*p*NA and *p*FTAA) were solvated in water and molecular dynamics (MD) simulations were performed in order to obtain different structural configurations. From the available number of structures seven and eight snapshots for *p*NA and *p*FTAA, respectively, were extracted, so-called “snapshots”. The simulation of *p*NA was kindly performed by Peter Reinholdt<sup>[259]</sup> but is described here for the sake of completeness. For the simulation of *p*NA, the molecule was parametrized with the General AMBER force field<sup>[260]</sup> and RESP charges<sup>[261]</sup> calculated with B3LYP<sup>[163, 164, 166]</sup> 6-31+G\* basis set<sup>[262-264]</sup> (with PCM<sup>[45]</sup> using the dielectric constant of water) in the AMBER software.<sup>[265]</sup> In the tleap module of the AMBER software the molecule was solvated in 3160 OPC-model represented water molecules.<sup>[266]</sup> At first, a steepest-descent minimization with 10000 steps was performed, followed by a conjugate-gradient minimization with 10000 steps. Subsequently, the system was equilibrated for 1 ns in the NPT ensemble (at 1 atm. pressure), where in the first 20 ps the system was heated from 0 to 298 K. In the following 100 ns production run in the NPT ensemble (at 298 K), a Langevin thermostat and Monte-Carlo barostat were employed. The Particle Mesh Ewald method was employed for the electrostatic treatment.<sup>[267]</sup> The cut-off threshold for non-bonded interactions was set to 12 Å and hydrogen bonds were constrained *via* the SHAKE algorithm.<sup>[268, 269]</sup> This simulation yielded one hundred snapshots, of which the seven snapshots were chosen arbitrarily and for further calculation adapted to solely include environmental water molecules within a 3, 4, 5, and 12 Å environment of *p*NA.

The simulation of *p*FTAA in water was performed in the GROMACS 2019.3 software<sup>[270-277]</sup> with an adapted CHARMM force field.<sup>[15, 16, 278, 279]</sup> The solvation in 4028 water molecules was represented by the TIP3P model.<sup>[280]</sup> At first, a steepest-

<sup>†</sup> This section is adapted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689-5703. Copyright 2023 American Chemical Society.

descent minimization with 50000 steps was performed. Subsequently, the system was equilibrated for 10 ns in the NPT ensemble (at 1 atm. pressure), where in the first 0.2 ps the system was heated from 0 to 300 K. In the following 100 ns production run in the NVT ensemble (at 300 K), a velocity-rescaling thermostat<sup>[281]</sup> and a Berendsen barostat<sup>[282]</sup> were employed. The Particle Mesh Ewald method was employed for the electrostatic treatment.<sup>[267, 283]</sup> The cut-off threshold for non-bonded interactions was set to 10 Å and non-bonded interactions were constrained *via* the LINCS algorithm.<sup>[284]</sup> This simulation yielded 61 snapshots, of which the eight snapshots were chosen arbitrarily and for further calculation adapted to solely include environmental water molecules within a 3 Å environment of pFTAA. For half of these adapted snapshots, one or two sodium counterions were found at an average distance of 2.3 Å to pFTAA, while for the other half no sodium atoms entered the system.

The supermolecular reference TD-DFT calculations were performed with Dalton 2020.1<sup>[285]</sup> employing the range-separated CAM-B3LYP XC functional.<sup>[174]</sup> Due to computational expense the calculations for pFTAA could only be performed for a 3 Å environment and an aug-cc-pVDZ basis set.<sup>[286]</sup>

#### **Polarizable Embedding**

The workflow for the PE scheme is shown in Fig. A.1. The polarizable embedding assistance script (PEAS)<sup>[287]</sup> was employed to read in the adapted snapshots and perform the partitioning of the system. For the *pNA* or pFTAA subsystem a Dalton input file was produced, then for the environmental subsystems, the Openmolcas 2.1<sup>[288]</sup> implementation of LoProp<sup>[289]</sup> was called. Here, the calculations were performed with a B3LYP XC functional<sup>[163, 164, 166]</sup> and ANO-type reconstructions of the aug-cc-pVDZ and aug-cc-pVTZ basis set for water<sup>[286, 290-292]</sup> and an ANO-L basis set for sodium ions.<sup>[293]</sup> PEAS then constructed a PE potential containing the obtained multipoles and polarizabilities. This potential could then be read-in *via* the PELib module<sup>[294]</sup> in Dalton and was used in ground-state as well as linear response calculations. Atomic pseudopotentials for the sodium ions were added to the QM system to counteract electron-spill-out effects.<sup>[295]</sup>

#### **Frozen-Density Embedding**

The workflow for the FDE scheme all programs were called with the PyADF 0.98 scripting environment, that called all other modules and programs.<sup>[61, 296]</sup> A supermolecular quadrature grid with a “good” Becke grid quality<sup>[297]</sup> was calculated with the ADF code of the AMS suite<sup>[298]</sup> and was used to write all following results on this grid. After partitioning the supermolecule into subsystems, PyADF called Dalton in order to perform individual subsystem ground-state DFT calculations. Here, for *pNA* and pFTAA a CAM-B3LYP XC functional was employed. For *pNA*, the calculations were performed with an aug-cc-pVDZ and an aug-cc-pVTZ basis set, whereas for pFTAA solely an aug-cc-pVDZ basis set was performed. The results were given to the DensityEvaluator extracting the electron densities and electrostatic potentials from the results. The PyEmbed module was then called, using these objects, determining the non-additive

kinetic energy with a PW91k functional<sup>[160, 208]</sup> and the non-additive XC energy with a BP86 functional<sup>[163, 299]</sup>. The Coulomb and non-additive potential were then added and given out as an embedding potential, that was further used in ground-state and linear-response calculations with Dalton where the active molecule was calculated with the CAM-B3LYP XC functional in presence of the embedding potential. Three freeze-and-thaw cycles have been used throughout as this setting had been found to generally yield sufficient results.<sup>[300]</sup> After calculating the five lowest excitations for the reference as well as for the embedding calculations (and up to 25 solely for pFTAA and PE), they were sorted by the oscillator strength of the transition. The strongest  $\pi \rightarrow \pi^*$  transition (ensured via inspection of response vectors and orbitals) was compared with other results. The oscillator strengths are presented in length form. Orbital representations were obtained with Molden 6.9 and an isocontour value of 0.05 for pNA and of 0.02 for pFTAA.<sup>[301]</sup> Molecular representations were attained with VMD 1.9.3.<sup>[302]</sup>

## 3.2. Computational Details for Calculations of Global Optical Properties

The employed crystal structures were obtained from publications in the crystal information file format as periodic structures.<sup>[96, 102-104]</sup> The VMD 1.9.3<sup>[302]</sup> software was employed in order to extract substructures from these for further calculations and for all molecular presentations shown. For the calculation of polarizabilities a general protocol was followed either in ORCA 5.0.1<sup>[303, 304]</sup> or TURBOMOLE V7.5.1 and V7.6.0<sup>[305-307]</sup>, where V7.5.1 version was employed for the initial evaluation calculations and V7.6.0 for all other calculations, respectively.

- In ORCA, SHARK<sup>[308]</sup>, libint2<sup>[309]</sup> and libXC<sup>[310]</sup> are employed in the calculations. All calculations were performed in a DFT framework with the Grimme D3 dispersion correction with Becke-Johnson damping<sup>[311, 312]</sup>, and the resolution of identity approximation for Coulomb integrals (RI-J)<sup>[313, 314]</sup>, Fock exchange integrals (RI-COSX)<sup>[315]</sup> and MP2 correlation integrals (RI-MP2)<sup>[316, 317]</sup>, where applicable and automatic generation of auxiliary basis sets where necessary (AutoAux)<sup>[318]</sup>. The default auxiliary basis sets were applied for all calculations. For the calculation of the static polarizability tensor, the CP-SCF algorithm with the Pople solver was employed.<sup>[246]</sup>
- In TURBOMOLE, optimizations were performed with the jobex script<sup>[319, 320]</sup>, self-consistent field (SCF) calculations with DFT were performed with the ridft script<sup>[313, 321]</sup> and TD-DFT calculations for the attainment of the frequency-dependent polarizability tensor were performed with the escf script<sup>[322-325]</sup>. If not stated otherwise a default m3 grid<sup>[326]</sup> and the resolution of identity approximation was applied where available<sup>[317]</sup> employing the default auxiliary basis sets.

For the ORCA calculations a XC functional benchmark was performed with the following functionals used: BP86<sup>[160, 163]</sup>, PBE<sup>[165]</sup>, TPSS<sup>[327]</sup>, B3LYP<sup>[163, 164, 166]</sup>, CAM-

B3LYP<sup>[163, 164, 166, 174]</sup>, PBE0<sup>[165, 328]</sup>, B2-PLYP<sup>[163, 164, 175]</sup>, B2GP-PLYP<sup>[163, 164, 329]</sup> and DSD-PBEPBE-D3BJ<sup>[165, 330]</sup>. The Gaussian-type split-valence atomic-orbital basis sets in these calculations were employed in double- $\zeta$  quality to quintuple- $\zeta$  quality ( $X = D, T, Q, 5$ ) for Dunning-type basis sets without diffuse basis functions<sup>[286, 331, 332]</sup> (cc-pVXZ) and up to triple- $\zeta$  quality ( $X = D, T$ ) with diffuse basis functions<sup>[291, 333]</sup> (aug-cc-pVXZ) as well as for the Ahlrichs-type basis sets with (def2-XZVPD) and without (def2-XZVP) diffuse functions ( $X = S, T, Q$ ).<sup>[334]</sup>

In TURBOMOLE for the geometry optimizations two different protocols were used, where a m4 grid was employed.<sup>[326]</sup> In the first protocol, the DFT XC functional PBE<sup>[165]</sup> was employed with a def2-SVP and a def2-TZVP basis set<sup>[313, 314, 326, 335]</sup> for the pre-optimization and main optimization, respectively, both with a Grimme D3 dispersion correction with Becke-Johnson damping<sup>[311, 312]</sup>. In the second protocol, the B97-3c method was employed with the def2-mTZVP basis set.<sup>[178]</sup>

In the DFT and TD-DFT calculation in order to determine the polarizability tensor the PBE<sup>[165]</sup> or CAM-B3LYP XC functional<sup>[163, 164, 166, 174]</sup> were employed. The basis sets employed in these calculations are the Dunning-type basis sets with (aug-cc-pVXZ)<sup>[291, 333]</sup> and without diffuse basis functions (cc-pVXZ)<sup>[286, 331, 332]</sup> for double- $\zeta$  to quadruple- $\zeta$  quality and double- $\zeta$  to quintuple- $\zeta$  quality, respectively. Additionally, the Ahlrichs-type basis sets with (def2-XZVP) and without diffuse basis functions (def2-XZVPD) basis sets were employed in double-, triple- and quadruple- $\zeta$  quality.<sup>[334-336]</sup> In all cases, where the auxiliary basis set was not set for Zn by default (quintuple- $\zeta$  calculations), a universal basis set was employed.<sup>[314]</sup>

In some cases, the TD-DFT calculations in TURBOMOLE failed due to non-real unstable reference states so that the SCF convergence criteria in the ridft ground-state calculations were increased from the default energy difference of  $10^{-8} E_h$  between steps, to  $10^{-9} E_h$ . Subsequently, if the TD-DFT calculation still proved difficult, the grid was increased from the default m3 grid to an m4 grid. If the problems persisted further, the SCF convergence criteria were even further tightened to  $10^{-10} E_h$ ,  $10^{-11} E_h$  and  $10^{-12} E_h$ . In some cases, these calculations did not converge or still resulted in problems for the TD-DFT calculations. In order to converge the reference calculation better, the conductor-like screening model (COSMO)<sup>[337]</sup> was utilized in order to simulate the environment of the molecule as a dielectric continuum. A permittivity of  $\epsilon = 2.263 e^2/a_0 E_h$ ,  $\epsilon_r = 1.825 e^2/a_0 E_h$ ,  $\epsilon_r = 1.787 e^2/a_0 E_h$  and  $\epsilon_r = 2.119 e^2/a_0 E_h$  for ZIF-8, ZIF-90, ZIF-318 and ZIF-71, respectively, was employed in the ground-state ridft calculation, but not in the TD-DFT calculation.

When problems occurred in the ORCA calculations some TD-DFT test calculations for Zn were performed with the ADF code of the AMS suite<sup>[298]</sup> and the DZP, TZP or QZ4P basis set<sup>[338-341]</sup> in order to determine the polarizability tensor.

Optimization of structures with PBEsol and reference calculation of the refractive indices with HSE06 in pDFT were kindly performed by Marvin Treger<sup>[129]</sup> with primitive cells for the MOFs. Geometry optimizations were performed with a PBEsol XC functional with ultra-soft pseudopotentials.<sup>[342]</sup> Single-point calculations of these geometries were then performed with the HSE06 XC functional.<sup>[343]</sup> For further details, the reader is referred to the computational protocol by Treger *et al.*<sup>[129]</sup>

## 4 | Dissection of Interaction Contributions for Local Optical Properties

The two presented classes of local embedding methods (Sections 2.4.1.1 and 2.4.1.3) target the same effects in order to describe environmental effects on a main system, however, follow different philosophies to obtain this goal. In contrast to QM/QM embeddings models, that are commonly introduced in a first quantization formalism such as the density-based embedding schemes (Section 2.4.1.1), the QM/MM PE embedding scheme is formulated in second quantization (Section 2.4.1.3). In order to dissect and directly compare the interaction contributions described by these methods, a common framework is indispensable. The derivation of this framework, that is held in first quantization, allows for this direct comparison of the effects accounted for and will be introduced in the following (Section 4.1).

For a further numerical comparison of the embedding methods, the previously established framework is implemented in a computational setup enabling a one-to-one comparison between the embedding schemes and the stepwise introduction of interaction effects. This setup as well as the following numerical comparison are described in Sections 4.2.1 and 4.2.2. The latter is performed by applying the calculation of local optical properties to the *p*NA and *p*F<sub>TAA</sub> systems solvated in different sizes of a water environment. The results for the dissected interactions are then discussed in Section 4.2.2.

### 4.1. Common framework for quantum-derived embedding schemes<sup>†</sup>

One mutuality between QM/MM embedding models and QM/QM embedding schemes is the attempt to recreate the Hamiltonian in Eq. (2.4.13). In the interaction Hamiltonian

---

<sup>†</sup> This section is largely adapted from Jansen, M.; Thi Minh, N. N.; Hedegård, E. D.; König, C. In *Chemical Modelling: Volume 17*, 2022. Partially adapted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society. This particularly holds true for the outline of the derivation of the framework and the introduced polarization models.

$\hat{V}_{XY}$ , only those terms from Eq. (2.1.2) are left, that refer to interactions for particles in different subsystems  $X$  and  $Y$ ,

$$\hat{V}_{XY} = \sum_{i \in X} \sum_{j \in Y} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i \in X} \sum_{J \in Y} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} - \sum_{I \in X} \sum_{j \in Y} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_j|} + \sum_{I \in X} \sum_{J \in Y} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (4.1.1)$$

with a set of electronic coordinates  $\mathbf{r}_i$ , nuclear coordinates  $\mathbf{R}_I$  and charges  $Z_I$  assigned to a subsystem  $X$  and accordingly for the electrons  $j$  and nuclei  $J$ , that are assigned to a subsystem  $Y$ . Since density-based embedding schemes such as the here presented QM/QM model are often formulated in terms of real-space electron densities rather than operators acting on a set of electronic coordinates, the density operator is introduced

$$\hat{\rho}_X(\mathbf{r}_x) = \sum_{i \in X} \delta(\mathbf{r}_i - \mathbf{r}_x), \quad (4.1.2)$$

where  $i$  and  $x$  refer to a single-electron and real-space coordinate, respectively. This relation can then be used to obtain the equivalent definition of the interaction Hamiltonian as follows,

$$\begin{aligned} \hat{V}_{XY} = & \int \int \frac{\hat{\rho}_X(\mathbf{r}_i) \hat{\rho}_Y(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j - \sum_{I \in X} \int \frac{\hat{\rho}_Y(\mathbf{r}_j) Z_I}{|\mathbf{R}_I - \mathbf{r}_j|} d\mathbf{r}_j \\ & - \sum_{J \in Y} \int \frac{\hat{\rho}_X(\mathbf{r}_i) Z_J}{|\mathbf{R}_J - \mathbf{r}_i|} d\mathbf{r}_i + \sum_{I \in X} \sum_{J \in Y} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \end{aligned} \quad (4.1.3)$$

Emerging from this operator, the interaction energy  $E_{\text{int}}$  can also be expressed in terms of energy functionals as follows,

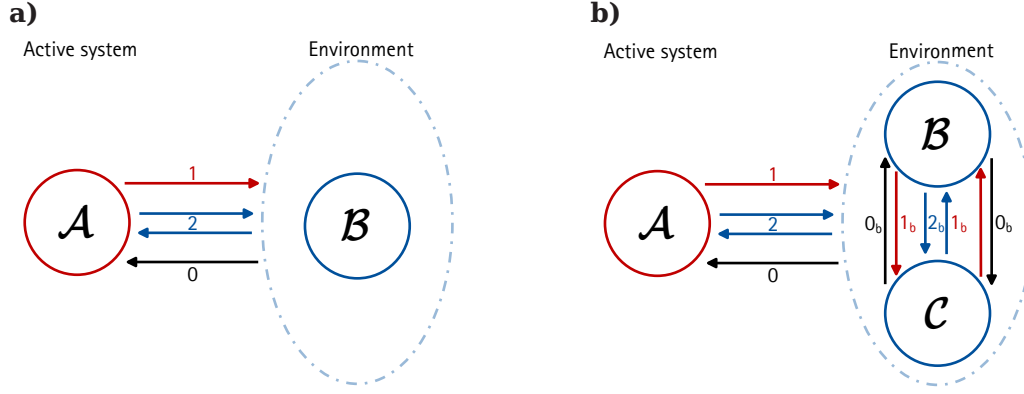
$$E_{\text{int}}[\rho_{\text{tot}}] = E_{\text{tot}}[\rho_{\text{tot}}] - \sum_X E_X[\rho_X], \quad (4.1.4)$$

where  $E_X$  is the energy of the isolated subsystem with the corresponding subsystem density  $\rho_X$  and  $E_{\text{tot}}$  is the total energy of the supermolecular system with the total electron density  $\rho_{\text{tot}}$ . The interaction energy can be further decomposed into a QM part  $E_{\text{int}}^{\text{QM}}$  and a Coulomb part  $E_{\text{int}}^{\text{C}}$ ,

$$E_{\text{int}}[\rho_{\text{tot}}] = E_{\text{int}}^{\text{C}}[\rho_{\text{tot}}] + E_{\text{int}}^{\text{QM}}[\rho_{\text{tot}}]. \quad (4.1.5)$$

The former, containing the classical and additive Coulomb interaction terms, is obtained in all discussed embedding schemes, whereas the latter vanishes for non-overlapping subsystems. Since this is assumed for the PE model (Section 2.4.1.3), these are only included in the density-based embedding model, where non-additive terms can occur (Section 2.4.1.1).

Introducing an exemplary supermolecule with a main subsystem  $\mathcal{A}$  and one environmental subsystem  $\mathcal{B}$ , displayed in Fig. 4.1.1 **a**), or two environmental subsystems  $\mathcal{B}$  and



**Figure 4.1.1.:** Schematic representation of the supermolecules discussed in this chapter. The active system  $\mathcal{A}$  and the interactions with the environment are displayed. The electrostatic interactions, single subsystem polarization and mutual polarization are denoted (0), (1) and (2), respectively. In **a)** there is one environmental subsystem  $\mathcal{B}$ , in **b)** there are two environmental subsystems  $\mathcal{B}$  and  $\mathcal{C}$ , that show further intra-environmental interactions with the subscript  $b$ .

$\mathcal{C}$ , displayed in Fig. 4.1.1 **b)**, the total electron density of this system is assumed to be the sum of the subsystem densities (*cf.* Eq. (2.4.1))

$$\rho_{\text{tot}} = \rho_{\mathcal{A}} + \rho_{\text{env}} \stackrel{\mathbf{a)}}{=} \rho_{\mathcal{A}} + \rho_{\mathcal{B}} \quad (4.1.6)$$

$$\stackrel{\mathbf{b)}}{=} \rho_{\mathcal{A}} + \rho_{\mathcal{B}} + \rho_{\mathcal{C}}. \quad (4.1.7)$$

For these systems the separate Coulomb and QM interaction energy (Eq. (4.1.4) and Eq. (4.1.5)) become

$$E_{\text{int}}^{\mathcal{C}}[\rho_{\mathcal{A}}, \rho_{\text{env}}] \stackrel{\mathbf{a)}}{=} E_{\text{tot}}^{\mathcal{C}}[\rho_{\text{tot}}] - E_{\mathcal{A}}^{\mathcal{C}}[\rho_{\mathcal{A}}] - E_{\mathcal{B}}^{\mathcal{C}}[\rho_{\mathcal{B}}] \quad (4.1.8)$$

$$\stackrel{\mathbf{b)}}{=} E_{\text{tot}}^{\mathcal{C}}[\rho_{\text{tot}}] - E_{\mathcal{A}}^{\mathcal{C}}[\rho_{\mathcal{A}}] - E_{\mathcal{B}}^{\mathcal{C}}[\rho_{\mathcal{B}}] - E_{\mathcal{C}}^{\mathcal{C}}[\rho_{\mathcal{C}}] \quad (4.1.9)$$

$$E_{\text{int}}^{\text{QM}}[\rho_{\mathcal{A}}, \rho_{\text{env}}] \stackrel{\mathbf{a)}}{=} E_{\text{tot}}^{\text{QM}}[\rho_{\text{tot}}] - E_{\mathcal{A}}^{\text{QM}}[\rho_{\mathcal{A}}] - E_{\mathcal{B}}^{\text{QM}}[\rho_{\mathcal{B}}] \quad (4.1.10)$$

$$\stackrel{\mathbf{b)}}{=} E_{\text{tot}}^{\text{QM}}[\rho_{\text{tot}}] - E_{\mathcal{A}}^{\text{QM}}[\rho_{\mathcal{A}}] - E_{\mathcal{B}}^{\text{QM}}[\rho_{\mathcal{B}}] - E_{\mathcal{C}}^{\text{QM}}[\rho_{\mathcal{C}}] \quad (4.1.11)$$

$$(4.1.12)$$

where  $E_{\mathcal{A}}[\rho_{\mathcal{A}}]$ ,  $E_{\mathcal{B}}[\rho_{\mathcal{B}}]$  and  $E_{\mathcal{C}}[\rho_{\mathcal{C}}]$  are the energies of the individual subsystems  $\mathcal{A}$ ,  $\mathcal{B}$  and  $\mathcal{C}$ , respectively.

Furthermore, the Coulomb interaction energy for the ground state (Eq. (4.1.9)) and its further decomposition is presented in Fig. 4.1.1. The electrostatic contributions between the main system and the environment are denoted (0) [Fig. 4.1.1 **a)** and **b)**], whereas the electrostatic contributions in between environmental subsystems are denoted ( $0_b$ ) [Fig. 4.1.1 **b)**]. Electron densities and subsystem wave functions, that are solely prone to these interactions, but are frozen in the process, *i.e.* are

unpolarized, will in the following be denoted  $\rho^{(0)}$  and  $\Psi^{(0)}$ , respectively. In case of a further polarization of the subsystem  $\mathcal{A}$  denoted (1)[Fig. 4.1.1 **a**] and **b**] or the subsystems in the environment  $\mathcal{B}$  and  $\mathcal{C}$  denoted ( $1_b$ ) [Fig. 4.1.1 **b**], the electron density is denoted  $\rho^{(1)}$ . A further mutual (self-consistent) polarization between the main system and the environment (2) or between the environmental subsystems ( $2_b$ ), respectively, leads to the mutually polarized electron density  $\rho^{(2)}$ . Accordingly, the subsystem wave function for the polarized and mutually polarized systems are denoted  $\Psi^{(1)}$  and  $\Psi^{(2)}$ , respectively. Additionally, in the process of an excitation, there naturally occurs a dynamical response of the system to the perturbation. This change is, however, not expressed in this ground state decomposition and will be discussed separately in the following chapters. Resulting from this decomposition, the Coulombic interaction energy of system **a**) for frozen densities (0) can according to Eq. (4.1.3) be expressed as,

$$E_{\text{int}}^{\text{C}}[\rho_{\text{tot}}^{(0)}] = \langle \Psi_{\text{tot}}^{(0)} | \sum_X \sum_{Y \neq X} \left[ \frac{1}{2} \int \int \frac{\hat{\rho}_X(\mathbf{r}_i) \hat{\rho}_Y(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j - \sum_{I \in X} \int \frac{Z_I \hat{\rho}_Y(\mathbf{r}_j)}{|\mathbf{R}_I - \mathbf{r}_j|} d\mathbf{r}_j \right. \\ \left. - \sum_{J \in Y} \int \frac{Z_J \hat{\rho}_X(\mathbf{r}_i)}{|\mathbf{R}_J - \mathbf{r}_i|} d\mathbf{r}_i + \sum_{I \in X} \sum_{J \in Y} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \right] | \Psi_{\text{tot}}^{(0)} \rangle, \quad (4.1.13)$$

where,  $\Psi_{\text{tot}}^{(0)}$  denotes the unpolarized wave function of the total system, that can also be expressed by a product of the non-overlapping subsystem wave functions  $\Psi_{\mathcal{A}}^{(0)}$  and  $\Psi_{\mathcal{B}}^{(0)}$ . Here, the Coulomb interactions contribute but do not polarize the involved electron densities  $\rho^{(0)}$ . Introducing the polarization of subsystem  $\mathcal{A}$ , the total electron density becomes  $\rho_{\text{tot}}^{(1)} = \rho_{\mathcal{A}}^{(1)} + \rho_{\text{env}}^{(0)}$ , and only the terms, that involve the electron density of subsystem  $\mathcal{A}$  change from the unpolarized density to the polarized electron density  $\rho_{\mathcal{A}}^{(1)}$ , so that the interaction energy is then a function of the polarized density of subsystem  $\mathcal{A}$ ,  $E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}^{(1)}, \rho_{\text{env}}^{(0)}]$ . The expression for the mutually polarized densities with superscript (2) is more complicated and will be discussed in more detail for the respective methods.

Along the lines of local embedding schemes, the obtained expressions can be used to formulate an effective Hamiltonian for the active subsystem, that yields the energy of subsystem  $\mathcal{A}$  polarized by the potential resulting from embedding it in the environmental subsystems.

$$\hat{H}_{\text{eff}} = \hat{H}_{\mathcal{A}} + \hat{v}_{\mathcal{A}}^{\text{emb}} \quad (4.1.14)$$

$$= \hat{H}_{\mathcal{A}} + \int \hat{\rho}_{\mathcal{A}} v_{\mathcal{A}}^{\text{emb}}(\mathbf{r}_i) d\mathbf{r}_i \quad (4.1.15)$$

with the embedding potential operator  $\hat{v}_{\mathcal{A}}^{\text{emb}}$  acting on subsystem  $\mathcal{A}$ . With the embedding potential resulting from the interactions of the environmental system on the main subsystem  $\mathcal{A}$ , the embedding energy is obtained as

$$E_{\text{emb}}[\rho_{\text{tot}}] = E_{\text{tot}}[\rho_{\text{tot}}] - E_{\mathcal{A}}[\rho_{\mathcal{A}}]. \quad (4.1.16)$$



with  $\rho_{\text{tot}}$  defined in Eq. (4.1.6). Decomposing into the Coulomb and QM part (Eq. (4.1.5)), the embedding potential is given as,

$$v_{\mathcal{A}}^{\text{emb}} = \frac{\delta E_{\text{tot}}^{\text{C}}[\rho_{\text{tot}}] - E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\text{tot}}^{\text{QM}}[\rho_{\text{tot}}] - E_{\mathcal{A}}^{\text{QM}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}}. \quad (4.1.17)$$

In the following chapters, the recreation of the interaction between the main system and the environment will be derived based on the presented common fundament for density-based embedding schemes (Section 4.1.1) and PE (Section 4.1.2).

#### 4.1.1. Density-Based Embedding

In the density-based embedding schemes all terms in the interaction Hamiltonian are treated by real-space electron densities connected *via* the delta function (Eq. (4.1.2)). Focussing on system **a**) and the case of unpolarized systems, that follow the (0) model in Fig. 4.1.1 **a**), the Coulombic interaction energy is given as,

$$\begin{aligned} E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}^{(0)}, \rho_{\mathcal{B}}^{(0)}] &= \int \int \frac{\rho_{\mathcal{A}}(\mathbf{r}_i)\rho_{\mathcal{B}}(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_i d\mathbf{r}_j - \sum_{I \in \mathcal{A}} \int \frac{Z_I \rho_{\mathcal{B}}(\mathbf{r}_j)}{|\mathbf{R}_I - \mathbf{r}_j|} d\mathbf{r}_j \\ &\quad - \sum_{J \in \mathcal{B}} \int \frac{Z_J \rho_{\mathcal{A}}(\mathbf{r}_i)}{|\mathbf{R}_J - \mathbf{r}_i|} d\mathbf{r}_i + \sum_{I \in \mathcal{A}} \sum_{J \in \mathcal{B}} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \end{aligned} \quad (4.1.18)$$

where the unpolarized subsystem densities  $\rho_{\mathcal{A}}^{(0)}$  and  $\rho_{\mathcal{B}}^{(0)}$  stay unchanged. Allowing the polarization of the active subsystem in the frozen environment, a polarization of the electron density occurs so that the electron density of the total system becomes  $\rho_{\text{tot}}^{(1)} = \rho_{\mathcal{A}}^{(1)} + \rho_{\mathcal{B}}^{(0)}$ . The Coulombic interaction energy then adapts to,

$$E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}^{(1)}, \rho_{\mathcal{B}}^{(0)}] = E_{\text{tot}}^{\text{C}}[\rho_{\text{tot}}^{(1)}] - E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}^{(1)}] - E_{\mathcal{B}}^{\text{C}}[\rho_{\mathcal{B}}^{(0)}], \quad (4.1.19)$$

with the polarized subsystem  $\mathcal{A}$ , yielding the polarized subsystem electron density  $\rho_{\mathcal{A}}^{(1)}$ . This polarization of the main subsystem in the environment potential with frozen electron densities is denoted FDE (Section 2.4.1.1). The polarization of subsystem  $\mathcal{A}$  will in turn induce a change on the environmental subsystem  $\mathcal{B}$ , that causes a further change of subsystem  $\mathcal{A}$  and so on. According to the aforementioned description in Section 2.4.1.1, this mutual polarization of the electron densities can be described by exchanging the active role in the FDE calculation with all other systems. In this procedure, the electron density is polarized for the respective active subsystem and by cycling the active role in freeze-and-thaw cycles, after a sufficient number of cycles self-consistency and therefore mutual polarization [Fig. 4.1.1, (2) and (2<sub>b</sub>)] is achieved.

Formulating the effective Hamiltonian for the active subsystem in the supermolecule with  $E_{\text{tot}}^{\text{FDE}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}] = E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}] + E_{\mathcal{A}}^{\text{QM}}[\rho_{\mathcal{A}}] + E_{\mathcal{B}}^{\text{C}}[\rho_{\mathcal{B}}] + E_{\mathcal{B}}^{\text{QM}}[\rho_{\mathcal{B}}] + E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}] + E_{\text{int}}^{\text{QM}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]$ , corresponding embedding potential (Eq. (4.1.17)) can be formulated as,

$$v_{\mathcal{A}}^{\text{FDE}}(\mathbf{r}_i) = \frac{\delta E_{\text{tot}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}] - E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\text{tot}}^{\text{QM}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}] - E_{\mathcal{A}}^{\text{QM}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}} \quad (4.1.20)$$

$$= \frac{\delta E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\mathcal{B}}^{\text{C}}[\rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\text{int}}^{\text{QM}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\mathcal{B}}^{\text{QM}}[\rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}}. \quad (4.1.21)$$

The functional derivative for the energy of subsystem  $\mathcal{B}$  vanishes for both, the Coulomb and QM term, since the electron density of the system  $\rho_{\mathcal{B}}$  is considered frozen, *i.e.* constant. The functional derivative for the Coulomb part of the interaction energy is obtained straightforwardly from Eq. (4.1.18). The derivative of the quantum mechanical interaction energy, however, is not additive (Section 2.4.1.1). It can be decomposed in a kinetic and XC term so that the FDE embedding potential (*cf.* Eq. (2.4.10)) reads,

$$v_{\mathcal{A}}^{\text{FDE}}(\mathbf{r}_i) = \frac{\delta E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\text{int}}^{\text{QM}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} \quad (4.1.22)$$

$$= - \sum_{J \in \mathcal{B}} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} + \int \frac{\rho_{\mathcal{B}}(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j + \frac{\delta E_{\text{int}}^{\text{kin}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\text{int}}^{\text{xc}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]}{\delta \rho_{\mathcal{A}}} \\ = v_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{B}}](\mathbf{r}_i) + v_{\mathcal{A}}^{\text{nadd,kin}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}](\mathbf{r}_i) + v_{\mathcal{A}}^{\text{nadd,xc}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}](\mathbf{r}_i), \quad (4.1.23)$$

where the Coulomb potential  $v_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{B}}]$  denotes the result from the functional derivative of the Coulombic interaction energy with subsystem  $\mathcal{A}$  and  $v_{\mathcal{A}}^{\text{nadd,kin}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]$  and  $v_{\mathcal{A}}^{\text{nadd,xc}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}]$  denote the non-additive kinetic and XC potential, respectively, resulting from the functional derivative of the QM interaction energy. In order to extend to a multiplicity of environmental subsystems, the potential has to be adapted to include interactions with more than one subsystem. For the QM terms, this is denoted by a change of the interaction terms from the density of subsystem  $\mathcal{B}$  to the environmental density  $\rho_{\text{env}}$ ,  $v_{\mathcal{A}}^{\text{nadd,kin}}[\rho_{\mathcal{A}}, \rho_{\text{env}}] + v_{\mathcal{A}}^{\text{nadd,xc}}[\rho_{\mathcal{A}}, \rho_{\text{env}}]$ . For the Coulomb potential, a sum over all environmental subsystems  $X$  is introduced,

$$v_{\mathcal{A}}^{\text{C}}[\rho_{\text{env}}](\mathbf{r}_i) = - \sum_{X \neq \mathcal{A}} \sum_{J \in X} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} + \sum_{X \neq \mathcal{A}} \int \frac{\rho_X(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (4.1.24)$$

$$= - \sum_{X \neq \mathcal{A}} \sum_{J \in X} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} + \int \frac{\rho_{\text{env}}(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j. \quad (4.1.25)$$

For the polarization of the main subsystem  $\mathcal{A}$ , this involves the frozen densities of the environment only. When performing freeze-and-thaw cycles the environmental densities stay frozen, however, for every step in the cycle, when exchanging the role of the active system to a different subsystem, a new potential has to be formed for every active

subsystem, incorporating the newly obtained densities from the last step. This way, when changing to an environmental subsystem as the active system, the interaction among environmental subsystems is incorporated (Fig. 4.1.1 **b**). The obtained formalism can be used in order to describe the local excitation of the active subsystem, the so-called uncoupled FDE approach.

Starting from the already introduced expression Eq. (2.3.14), but employing the effective Fock operator  $\hat{f}^{\text{eff}} = \hat{f}^{\text{iso}} + v^{\text{emb}}$ , the Fock matrix element derivative becomes,

$$\frac{\partial F_{pq}^{(0)}}{\partial P_{rs}} = \frac{\partial \langle \psi_p | \hat{f}^{\text{iso}} + v^{\text{emb,FDE}} | \psi_q \rangle}{\partial P_{rs}} = \frac{\partial \langle \psi_p | \hat{f}^{\text{iso}} | \psi_q \rangle}{\partial P_{rs}} + \frac{\partial \langle \psi_p | v^{\text{emb,FDE}} | \psi_q \rangle}{\partial P_{rs}}, \quad (4.1.26)$$

where the terms for the isolated Fock operator stay unchanged and can be separated from the embedding term<sup>[214, 345]</sup>,

$$\frac{\partial \langle \psi_p | v_{\mathcal{A}}^{\text{emb,FDE}} | \psi_q \rangle}{\partial P_{rs}} = \left\langle \psi_p(\mathbf{r}_i) \psi_s(\mathbf{r}'_i) \left| \frac{\delta v_{\mathcal{A}}^{\text{emb,FDE}}[\rho_{\mathcal{A}}(\mathbf{r}'_i)](\mathbf{r}_i)}{\delta \rho_{\mathcal{A}}(\mathbf{r}'_i)} \right| \psi_q(\mathbf{r}_i) \psi_r(\mathbf{r}'_i) \right\rangle. \quad (4.1.27)$$

$$= \int d\mathbf{r}_i \int d\mathbf{r}'_i \rho_{pq}^t(\mathbf{r}_i) \frac{\delta v_{\mathcal{A}}^{\text{emb,FDE}}(\mathbf{r}_i)}{\delta \rho_{\mathcal{A}}(\mathbf{r}'_i)} \rho_{rs}^t(\mathbf{r}'_i), \quad (4.1.28)$$

with the matrix elements for the transition density  $\rho_{pq}^t(\mathbf{r}) = \psi_p^*(\mathbf{r})\psi_q(\mathbf{r})$ . Inserting equation Eq. (4.1.23) in the functional derivative of the embedding potential, the Coulomb terms vanish due to the environmental electron densities being frozen. Therefore, the derivative is solely comprised of the contribution from the non-additive kinetic and XC terms<sup>[213, 345]</sup>

$$\frac{\delta v_{\mathcal{A}}^{\text{FDE}}(\mathbf{r}_i)}{\delta \rho_{\mathcal{A}}(\mathbf{r}'_i)} = \frac{\delta^2 E^{\text{xc}}[\rho_{\text{tot}}]}{\delta \rho_{\text{tot}}(\mathbf{r}_i) \delta \rho_{\text{tot}}(\mathbf{r}'_i)} - \frac{\delta^2 E^{\text{xc}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}(\mathbf{r}_i) \delta \rho_{\mathcal{A}}(\mathbf{r}'_i)} + \frac{\delta^2 T_s[\rho_{\text{tot}}]}{\delta \rho_{\text{tot}}(\mathbf{r}_i) \delta \rho_{\text{tot}}(\mathbf{r}'_i)} - \frac{\delta^2 T_s[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}(\mathbf{r}_i) \delta \rho_{\mathcal{A}}(\mathbf{r}'_i)}, \quad (4.1.29)$$

following Eqs. (2.4.6) and (2.4.7). These terms have, however, been shown to be rather small.<sup>[346, 347]</sup>

### 4.1.2. Polarizable Embedding

In the polarizable embedding scheme, that was presented in Section 2.4.1.3 in the second quantization formalism, the starting point for system **a**) (see Fig. 4.1.1) is the same as in the FDE scheme, the interaction energy of unpolarized subsystems, Eq. (4.1.18). As previously mentioned (see Section 2.4.1.3), a multipole expansion is introduced in order to express the environmental electron densities classically. For this, the inverse distances are expanded around an origin  $R_B$  at the centre of mass of subsystem  $B$ , exemplary derived for the electron-electron distance, as follows

$$\begin{aligned} |\mathbf{r}_i - \mathbf{r}_j|^{-1} &= \frac{1}{|\mathbf{r}_i - \mathbf{R}_B|} - \sum_{\alpha} \left( \frac{\partial}{\partial r_{i,\alpha}} \frac{1}{|\mathbf{r}_i - \mathbf{R}_B|} \right) (r_{j,\alpha} - R_{B,\alpha}) + \dots \\ &= T_{B_i}^{(0)} - \sum_{\alpha} T_{B_i,\alpha}^{(1)} (r_{j,\alpha} - R_{B,\alpha}) + \dots \end{aligned} \quad (4.1.30)$$

$$= \sum_{|k|=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{B_i}^{(k)} (\mathbf{r}_j - \mathbf{R}_B)^k \quad (4.1.31)$$

$$|\mathbf{r}_i - \mathbf{R}_J|^{-1} = \sum_{|k|=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{B_i}^{(k)} (\mathbf{R}_J - \mathbf{R}_B)^k \quad (4.1.32)$$

$$|\mathbf{R}_I - \mathbf{R}_J|^{-1} = \sum_{|k|=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{B_i}^{(k)} (\mathbf{R}_J - \mathbf{R}_B)^k \quad (4.1.33)$$

with the interaction tensors defined in Eq. (2.4.19) and the variable for the spatial coordinate  $\alpha = x, y, z$ . Here, the interaction tensor contains an implicit dependence on the particle coordinate  $\mathbf{r}_i$  or  $\mathbf{R}_I$ .<sup>[225]</sup> Choosing to describe the electron density of subsystem  $B$  via multipoles employing interaction tensors, the interaction energy can therefore be simplified to

$$E_{\text{int}}^{\text{C}}[\rho_{\mathcal{A}}^{(0)}, \rho_{\mathcal{B}}^{(0)}] \approx \quad (4.1.34)$$

$$\begin{aligned} E_{\text{mul}}^{\text{C}}[\rho_{\mathcal{A}}^{(0)}, \rho_{\mathcal{B}}^{(0)}] &= \left( -\langle \Psi_{\mathcal{A}}^{(0)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) T_{B_i}^{(0)} d\mathbf{r}_i | \Psi_{\mathcal{A}}^{(0)} \rangle + \sum_{I \in \mathcal{A}} Z_I T_{B_i}^{(0)} \right) \langle \Psi_{\mathcal{B}}^{(0)} | \hat{q}_{\mathcal{B}} | \Psi_{\mathcal{B}}^{(0)} \rangle \\ &\quad - \sum_{\alpha} \left( -\langle \Psi_{\mathcal{A}}^{(0)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) T_{B_i}^{(1)} d\mathbf{r}_i | \Psi_{\mathcal{A}}^{(0)} \rangle + \sum_{I \in \mathcal{A}} Z_I T_{B_i}^{(1)} \right) \langle \Psi_{\mathcal{B}}^{(0)} | \hat{\mu}_{\mathcal{B},\alpha} | \Psi_{\mathcal{B}}^{(0)} \rangle + \dots \end{aligned} \quad (4.1.35)$$

with the multipole moment operators

$$\hat{q}_{\mathcal{B}} = - \int \hat{\rho}_{\mathcal{B}}(\mathbf{r}_j) d\mathbf{r}_j + \sum_{J \in \mathcal{B}} Z_J \quad (4.1.36)$$

$$\hat{\mu}_{\mathcal{B},\alpha} = - \int \hat{\rho}_{\mathcal{B}}(\mathbf{r}_j) (r_{j,\alpha} - R_{B,\alpha}) d\mathbf{r}_j + \sum_{J \in \mathcal{B}} Z_J (R_{J,\alpha} - R_{B,\alpha}) \quad (4.1.37)$$

that were defined with respect to the charge density of subsystem  $\mathcal{B}$ . Introducing the multi-index  $k$  also for the multipole moments, one multi-order multipole moment operator emerges,

$$\hat{\mathbf{Q}}_{\mathcal{B}}^{(k)} = - \int \hat{\rho}_{\mathcal{B}}(\mathbf{r}_j)(r_{j,\alpha} - R_{\mathcal{B},\alpha})^k d\mathbf{r}_j + \sum_{J \in \mathcal{B}} Z_J (R_{J,\alpha} - R_{\mathcal{B},\alpha})^k. \quad (4.1.38)$$

Due to the convergence problems of the expansion for short distances (*cf.* Section 2.4.1.3), distributed multipoles are introduced, *i.e.*, they are located on so-called “sites” on the environmental subsystem’s atoms (occasionally also bond-midpoints are used as sites). Resulting from this, for a single environmental subsystem  $\mathcal{B}$ , the interaction energy can be expressed in terms of sites as,

$$E_{\text{mul}}^{\text{C}}[\rho_{\mathcal{A}}^{(0)}, \rho_{\mathcal{B}}^{(0)}] = \sum_s \sum_{|k|} \frac{(-1)^{|k|}}{k!} \left( - \langle \Psi_{\mathcal{A}}^{(0)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) \mathbf{T}_{\text{si}}^{(k)} d\mathbf{r}_i | \Psi_{\mathcal{A}}^{(0)} \rangle + \sum_{I \in \mathcal{A}} Z_I \mathbf{T}_{sI}^{(k)} \right) \mathbf{Q}_s^{(k)}.$$

For now, only the electrostatic contributions were examined, however, with this same expression, the interaction energy for one polarized main subsystem  $\mathcal{A}$  can be captured by introducing  $\rho_{\mathcal{A}}^{(1)}$  and  $\Psi^{(1)}_{\mathcal{A}}$ ,

$$E_{\text{mul}}^{\text{C}}[\rho_{\mathcal{A}}^{(1)}, \rho_{\mathcal{B}}^{(0)}] = \sum_s \sum_{|k|} \frac{(-1)^{|k|}}{k!} \left( - \langle \Psi_{\mathcal{A}}^{(1)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) \mathbf{T}_{\text{si}}^{(k)} d\mathbf{r}_i | \Psi_{\mathcal{A}}^{(1)} \rangle + \sum_{I \in \mathcal{A}} Z_I \mathbf{T}_{sI}^{(k)} \right) \mathbf{Q}_s^{(k)} \quad (4.1.39)$$

$$= \sum_s \sum_{|k|} \frac{(-1)^{|k|}}{k!} \left( - \int \rho_{\mathcal{A}}^{(1)}(\mathbf{r}_i) \mathbf{T}_{\text{si}}^{(k)} d\mathbf{r}_i + \sum_{I \in \mathcal{A}} Z_I \mathbf{T}_{sI}^{(k)} \right) \mathbf{Q}_s^{(k)}, \quad (4.1.40)$$

where the the multipole-moment vector is defined as  $\mathbf{Q}_s^{(k)}[\rho_{\mathcal{B}}^{(0)}] = \mathbf{Q}_s^{(k)} = \langle \Psi_{\mathcal{B}}^{(0)} | \hat{\mathbf{Q}}_s^{(k)} | \Psi_{\mathcal{B}}^{(0)} \rangle$ . Additionally, when determining the energy of the isolated QM subsystem, it is obtained from the polarized subsystem  $E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}^{(1)}]$ .

In order to take the mutual polarization into account, the second-order perturbed contribution  $E^{(2)}$  (the last term in Eq. (2.4.14)) has to be considered, since it contains the higher polarization contribution as well as the dispersion effects. For the two subsystem case discussed here, the second-order perturbation term becomes

$$E^{(2)} = E_{\text{PE}}^{\text{ind}} = E_{\mathcal{A}}^{\text{pol}} + E_{\mathcal{B}}^{\text{pol}} + E^{\text{disp}} \quad (4.1.41)$$

$$= \sum_{n \neq 0} \frac{\langle \Psi_{\mathcal{B},0}^{(0)} | \langle \Psi_{\mathcal{A},0}^{(0)} | \hat{V}_{\mathcal{A}\mathcal{B}} | \Psi_{\mathcal{A},n}^{(0)} \rangle | \Psi_{\mathcal{B},0}^{(0)} \rangle \langle \Psi_{\mathcal{B},0}^{(0)} | \langle \Psi_{\mathcal{A},n}^{(0)} | \hat{V}_{\mathcal{A}\mathcal{B}} | \Psi_{\mathcal{A},0}^{(0)} \rangle | \Psi_{\mathcal{B},0}^{(0)} \rangle}{E_{\mathcal{A},0}^{(0)} - E_{\mathcal{A},n}^{(0)}} \quad (4.1.42)$$

$$+ \sum_{n \neq 0} \frac{\langle \Psi_{\mathcal{B},0}^{(0)} | \langle \Psi_{\mathcal{A},0}^{(0)} | \hat{V}_{\mathcal{A}\mathcal{B}} | \Psi_{\mathcal{A},0}^{(0)} \rangle | \Psi_{\mathcal{B},n}^{(0)} \rangle \langle \Psi_{\mathcal{B},n}^{(0)} | \langle \Psi_{\mathcal{A},0}^{(0)} | \hat{V}_{\mathcal{A}\mathcal{B}} | \Psi_{\mathcal{A},0}^{(0)} \rangle | \Psi_{\mathcal{B},0}^{(0)} \rangle}{E_{\mathcal{B},0}^{(0)} - E_{\mathcal{B},n}^{(0)}} + E^{\text{disp}},$$

where the first term refers to the polarization of subsystem  $\mathcal{A}$  and the second term to the polarization of subsystem  $\mathcal{B}$ . The third term, that is comprised of terms with excitations in both subsystems, is identified as the dispersion energy<sup>[225]</sup>, that is not taken into account here.

Inserting the expressions for the interaction Hamiltonian on the example of the polarization of subsystem  $\mathcal{B}$ , including the terms for the previously introduced multipole expansion, leads to

$$E_{\mathcal{B}}^{\text{pol}} = \sum_{\alpha\beta} \left\{ \left( -\langle \Psi_{\mathcal{A},0}^{(0)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) T_{\mathcal{B}i}^{(1)} d\mathbf{r}_i | \Psi_{\mathcal{A},0}^{(0)} \rangle + \sum_{I \in \mathcal{A}} Z_I T_{\mathcal{B}I}^{(1)} \right) \right. \\ \cdot \sum_{n \neq 0} \frac{\langle \Psi_{\mathcal{B},0}^{(0)} | \hat{\mu}_{\mathcal{B},\alpha} | \Psi_{\mathcal{B},n}^{(0)} \rangle \langle \Psi_{\mathcal{B},n}^{(0)} | \hat{\mu}_{\mathcal{B},\beta} | \Psi_{\mathcal{B},0}^{(0)} \rangle}{E_{\mathcal{B},0}^{(0)} - E_{\mathcal{B},n}^{(0)}} \\ \left. \cdot \left( -\langle \Psi_{\mathcal{A},0}^{(0)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) T_{\mathcal{B}i}^{(1)} d\mathbf{r}_i | \Psi_{\mathcal{A},0}^{(0)} \rangle + \sum_{I \in \mathcal{A}} Z_I T_{\mathcal{B}I}^{(1)} \right) \right\}, \quad (4.1.43)$$

where only first-order terms survive since the terms of zeroth order in  $k$  vanish due to the orthogonality of the ground and excited state  $\langle \Psi_{\mathcal{B},0}^{(0)} | \Psi_{\mathcal{B},n}^{(0)} \rangle = 0$ . Since the expression is truncated after first order  $|k| = 1$  here, this solely leaves the interaction terms with the dipole moments of the environmental subsystems. The sum-over-states expression in the second term can be further identified as the static polarizability of subsystem  $\mathcal{B}$  following the definition

$$\alpha_{\mathcal{B},\alpha\beta}^0 = -2 \sum_{n \neq 0} \frac{\langle \Psi_{\mathcal{B},0}^{(0)} | \hat{\mu}_{\mathcal{B},\alpha} | \Psi_{\mathcal{B},n}^{(0)} \rangle \langle \Psi_{\mathcal{B},n}^{(0)} | \hat{\mu}_{\mathcal{B},\beta} | \Psi_{\mathcal{B},0}^{(0)} \rangle}{E_{\mathcal{B},0}^{(0)} - E_{\mathcal{B},n}^{(0)}}. \quad (4.1.44)$$

The field that is introduced by the electrons and nuclei in subsystem  $\mathcal{A}$  is denoted  $\mathcal{E}^e$  and  $\mathcal{E}^n$ , respectively. The field as the negative derivative of the potential at  $\mathbf{R}_{\mathcal{B}}$  arises as,

$$\mathcal{E}_{\alpha}^n = - \sum_I \frac{\partial V(\mathbf{R}_I)}{\partial R_{I,\alpha}} = - \sum_{I \in \mathcal{A}} Z_I T_{\mathcal{B}I,\alpha}^{(1)} \quad (4.1.45)$$

$$\mathcal{E}_{\alpha}^{\text{el}}[\rho_{\mathcal{A}}^{(0)}] = - \int \rho_{\mathcal{A}}^{(0)}(\mathbf{r}_i) \left( \frac{\partial V(\mathbf{r}_i)}{\partial r_{i,\alpha}} \right) d\mathbf{r}_i = \langle \Psi_{\mathcal{A}}^{(0)} | \int \hat{\rho}_{\mathcal{A}}(\mathbf{r}_i) T_{\mathcal{B}\alpha}^{(1)} d\mathbf{r}_i | \Psi_{\mathcal{A}}^{(0)} \rangle. \quad (4.1.46)$$

resulting from the potential due to the presence of electrons and nuclei in subsystem  $\mathcal{A}$ . Inserting Eqs. (4.1.44) to (4.1.46), Eq. (4.1.43) can be reduced to,

$$E_{\mathcal{B}}^{\text{pol}}[\rho_{\mathcal{B}}^{(0)}] = \frac{1}{2} \left( \mathcal{E}_{\mathcal{A}}^{\text{el}}[\rho_{\mathcal{A}}^{(0)}] + \mathcal{E}_{\mathcal{A}}^{\text{nuc}} \right)^{\text{T}} \boldsymbol{\mu}_{\mathcal{B}}^{\text{ind}}[\rho_{\mathcal{A}}] \quad (4.1.47)$$

by additionally recognizing the induced dipole moment as  $\boldsymbol{\mu}_{\mathcal{B}}^{\text{ind}}[\rho_{\mathcal{A}}] = \boldsymbol{\alpha}_{\mathcal{B}}^0 \mathcal{E}_{\mathcal{B}}^{\text{tot}}$ , that depends on the total field emerging from the nuclei and electrons in the main subsystem  $\mathcal{A}$  ( $\mathcal{E}^{\text{tot}} = \mathcal{E}_{\mathcal{A}}^{\text{el}}[\rho_{\mathcal{A}}^{(0)}] + \mathcal{E}_{\mathcal{A}}^{\text{nuc}}$ ). Furthermore, introducing multiple environmental subsystems

[Fig. 4.1.1 **b**)] along with atomic sites for the distributed multipoles, does not only lead to site-dependent fields and polarizabilities, but it also gives rise to additional fields. On the one hand, the sites of the multipoles of all other subsystems create a field  $\mathcal{E}^{\text{mul}}$ , that reflects the interactions with subscript  $b$  in Fig. 4.1.1 **b**). On the other hand, the existence of other sites  $s'$  causes other induced dipoles to arise and therefore an additional field  $\mathcal{E}^{\text{ind}}$ ,

$$\mathcal{E}_s^{\text{tot}} = \mathcal{E}_{\mathcal{A},s}^{\text{el}}[\rho_{\mathcal{A}}^{(0)}] + \mathcal{E}_{\mathcal{A},s}^{\text{nuc}} + \mathcal{E}_s^{\text{mul}} + \underbrace{\sum_{s' \neq s} \mathbf{T}_{ss'}^{(2)} \mu_{s'}^{\text{ind}}[\rho_{\text{tot}}]}_{\mathcal{E}_s^{\text{ind}}}. \quad (4.1.48)$$

Here, the induced dipole moment of one site depends on the induced dipole moment of all other sites, which requires a self-consistent treatment. For this, the matrix-vector equation proposed by Applequist *et al.* can be employed<sup>[348]</sup>

$$\begin{aligned} \boldsymbol{\mu}^{\text{ind}} &= \sum_s^S \boldsymbol{\mu}_s^{\text{ind}} = \sum_s^S (\boldsymbol{\alpha}_s \mathcal{E}_s^{\text{tot}}) = \sum_s^S \mathbf{R}_{ss'} \mathcal{E}_s^{\text{tot}} = \mathbf{R} \mathcal{E}^{\text{tot}} \\ &= \begin{pmatrix} \boldsymbol{\alpha}_1^{-1} & -\mathbf{T}_{12}^{(2)} & \dots & -\mathbf{T}_{1S}^{(2)} \\ -\mathbf{T}_{21}^{(2)} & \boldsymbol{\alpha}_2^{-1} & \dots & -\mathbf{T}_{2S}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{T}_{S1}^{(2)} & -\mathbf{T}_{S2}^{(2)} & \dots & \boldsymbol{\alpha}_S^{-1} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathcal{E}_1^{\text{tot}} \\ \mathcal{E}_2^{\text{tot}} \\ \vdots \\ \mathcal{E}_S^{\text{tot}} \end{pmatrix}. \end{aligned} \quad (4.1.50)$$

Shifting the focus from the interaction energies to the effective embedding potential, that allows for the calculation of the main subsystem energy from an effective Hamiltonian (Eq. (4.1.14)), the Coulomb part of the embedding potential (Eq. (4.1.17)) in the PE framework with a total energy of

$$E_{\text{tot}}^{\text{PE}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}] = E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}] + E_{\mathcal{B}}^{\text{C}}[\rho_{\mathcal{B}}] + E_{\text{mul}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}] + E_{\mathcal{A}}^{\text{pol}}[\rho_{\mathcal{B}}] + E_{\mathcal{B}}^{\text{pol}}[\rho_{\mathcal{A}}] \quad (4.1.51)$$

is given as,

$$v_{\mathcal{A}}^{\text{PE}}(\mathbf{r}_i) = \frac{\delta E_{\text{tot}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}^{(0)}] - E_{\mathcal{A}}^{\text{C}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}} = \frac{\delta E_{\mathcal{B}}^{\text{C}}[\rho_{\mathcal{B}}^{(0)}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\text{mul}}^{\text{C}}[\rho_{\mathcal{A}}, \rho_{\mathcal{B}}^{(0)}]}{\delta \rho_{\mathcal{A}}} + \frac{\delta E_{\mathcal{B}}^{\text{pol}}[\rho_{\mathcal{A}}]}{\delta \rho_{\mathcal{A}}}, \quad (4.1.52)$$

where the first term in the second equality vanishes when assuming the electron densities of subsystem  $\mathcal{B}$  to be constant. Simply performing the functional derivative of the second

term and using the functional derivative product rule for the third term<sup>[156]</sup>, the potential becomes

$$v_{\mathcal{A}}^{\text{PE}} = - \sum_s \left( \underbrace{\sum_{|k|=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{si}^{(k)}(\mathbf{r}_i) \mathbf{Q}_s^{(k)}[\rho_{\mathcal{B}}^{(0)}]}_{v_s^{\text{mul}}(\mathbf{r}_i)} - \underbrace{\left( \boldsymbol{\varepsilon}_s^{\text{tot}} \right)^T \boldsymbol{\alpha}_s^0 \frac{\delta \mathcal{E}_{\mathcal{A},s}^e[\rho_{\mathcal{A}}(\mathbf{r}_i)]}{\delta \rho_{\mathcal{A}}(\mathbf{r}_i)}}_{v_s^{\text{pol}}([\rho_{\mathcal{A}}](\mathbf{r}_i))} \right), \quad (4.1.53)$$

Having obtained the form of this effective operator (Eq. (4.1.14)), equally to the effective operator for density-based embedding schemes, it can be inserted in the TD-DFT equations (Eq. (2.3.13)), in the effective Fock operator and the just obtained PE embedding potential can be inserted in Eq. (4.1.26) to give,

$$\frac{\partial \langle \psi_p | v_{\mathcal{A}}^{\text{emb,PE}} | \psi_q \rangle}{\partial P_{rs}} = \left\langle \psi_p(\mathbf{r}_i) \psi_s(\mathbf{r}'_i) \left| \frac{\delta v_{\mathcal{A}}^{\text{emb,PE}}[\rho_{\mathcal{A}}(\mathbf{r}'_i)](\mathbf{r}_i)}{\delta \rho_{\mathcal{A}}(\mathbf{r}'_i)} \right| \psi_q(\mathbf{r}_i) \psi_r(\mathbf{r}'_i) \right\rangle \quad (4.1.54)$$

$$= \int d\mathbf{r}_i \int d\mathbf{r}'_i \rho_{pq}^t(\mathbf{r}_i) \frac{\delta v_{\mathcal{A}}^{\text{PE}}[\rho_{\mathcal{A}}(\mathbf{r}'_i)](\mathbf{r}_i)}{\delta \rho_{\mathcal{A}}(\mathbf{r}'_i)} \rho_{rs}^t(\mathbf{r}'_i), \quad (4.1.55)$$

$$= \int d\mathbf{r}_i \int d\mathbf{r}'_i \rho_{pq}^t(\mathbf{r}_i) \left( - \sum_s \sum_{s'} \mathbf{T}_{is'}^{(1)}(\mathbf{r}'_i) \mathbf{R}_{ss'} \mathbf{T}_{is}^{(1)}(\mathbf{r}_i) \right) \rho_{rs}^t(\mathbf{r}'_i), \quad (4.1.56)$$

which is the extra term introduced *via* the  $\mathbf{A}$  and  $\mathbf{B}$  terms in the TD-DFT equations (Eq. (2.3.13)). It accounts for the effects that occur due to the perturbation of the system, *i.e.* the polarization of the environmental subsystems, often named the dynamical response of the environment<sup>[193, 222]</sup>. Since this potential is acting on the active system this so-called differential polarization is included in this approach.

### 4.1.3. Comparison of Embedding Schemes in the Common Framework

Having successfully established a framework for both the density-based and PE embedding schemes, mutualities as well as differences in the description of interactions in the two classes become more clear. For instance, the decomposition of the interaction energy into Coulomb and QM parts has been performed. The former is reproduced in both classes, however, in different ways: while FDE enables to retrieve formally exact results with a density-based QM/QM treatment, in PE, the environmental subsystem densities are approximated *via* a multipole expansion and polarizabilities introduced in a perturbative approach of the interaction energy. The novel framework allows for the setup of different models that reflect the order of included polarizability in the interaction:

- (i) The NOPOL model, where solely the active subsystem is polarized in the interaction. This is reflected by a polarized subsystem  $\mathcal{A}$  [Fig. 4.1.1 (1)] in the frozen environment and refers to Eq. (4.1.23) and Eq. (4.1.53) with only the multipole potential, in the density-based embedding and PE, respectively. In FDE, the electron density



is relaxed due to the interactions with the environmental subsystems' densities, while in PE it is relaxed in a potential that is obtained from the multipole expansion of these environmental subsystem densities. These schemes are therefore comparable. It should, however, be noted that in the FDE scheme, QM contributions are accounted for in an approximate manner (for all models).

- (ii) The GSPOL model, that refers to a mutually polarized system in the ground state [Fig. 4.1.1 (2) and (2<sub>b</sub>)], which is reflected by the FDE approach including freeze-and-thaw cycles up to self-consistency account for the polarizability in the density-based embedding (Eq. (4.1.23)). In the PE scheme, the inclusion of the higher-order polarization is obtained by introducing anisotropic dipole polarizabilities from the perturbation of the Hamiltonian. Since the embedding potential is obtained by the multipole expansion and the polarization of the environmental subsystems due to the active system as well as the other multipoles and arising induced dipoles in this process (Eq. (4.1.53)), it reproduces the ground-state polarization and is comparable to the FDE approach with freeze-and-thaw cycles (Eq. (4.1.23)).
- (iii) The DPOL model, refers to the inclusion of differential polarization, *i.e.* the approximate dynamical response of the system arising due to an external field. This model is not reflected in Fig. 4.1.1, since it does not refer to ground state interactions. As previously mentioned, DPOL is covered in the PE scheme due to the polarization term that enters the TD-DFT equations and yields an additional term (Eq. (4.1.56)). This term accounts for the polarization of the environment due to the excitation in the active subsystem, that in turn then optimizes the active subsystem. In the uncoupled FDE scheme, this fraction of the polarization is not accounted for. There have, however, been approaches covering this part such as state-specific approaches that involve the optimization of individual states and therefore naturally include the dynamical response<sup>[65]</sup> or coupled FDE, where the response equations are reformulated to cover excitonically coupled subsystems.<sup>[213-218]</sup> In later approaches an effective contribution to Eq. (2.3.13) is employed, that includes non-resonant effects.<sup>[213]</sup> In the present work, these approaches will be disregarded.
- (iv) The DPOL model, refers to the inclusion of differential polarization, however, this does not take into account that the field of the dipole moments in the environment of the active molecule also modifies the external field itself, interacting with the active molecule. The resulting external effective field (EEF), that is therefore indirectly introduced by the polarization of the environment due to these local fields, can be further incorporated in the DPOL+EEF model. It is implemented additionally to the common PE formalism. For a derivation, the reader is referred to List *et al.*<sup>[349]</sup>

The (0) and (0<sub>b</sub>) interactions from Fig. 4.1.1 are not separately covered in these models, since they are incorporated in all other models and the electrostatic contributions are rarely employed separately apart from QM/MM point-charge models.<sup>[42]</sup>

Apart from the presented main models further details of the embedding schemes can be discussed. For PE schemes, the use of a multipole expansion can follow the occurrence of electron spill-out (ESO) effects, where the lack of short-range repulsion due to classical electrostatic interaction in the Coulomb terms with the active QM system causes the electron density to unphysically distribute in the environmental space.<sup>[350]</sup> The employment of adapted effective core potentials (pseudopotentials) in order to add repulsive terms for specific environmental regions has been found to alleviate these effects.<sup>[295]</sup> In FDE, ESO effects have also been observed, but are less common since the approximate non-additive functionals calculated from orbital-free DFT have shown to contain repulsion terms counteracting the electronic spill-out effect.<sup>[58, 209, 351, 352]</sup> Since these functionals perform well, they have also been introduced to PE, which is known as polarizable density embedding<sup>[353-356]</sup>, but exceeds the scope of this work. Furthermore, it is possible to counter occurring ESOs in FDE by the exertion of a correction to the kinetic energy functional for long distances.<sup>[352]</sup>

## 4.2. Numerical Comparison of Quantum-Derived Embedding Schemes<sup>‡</sup>

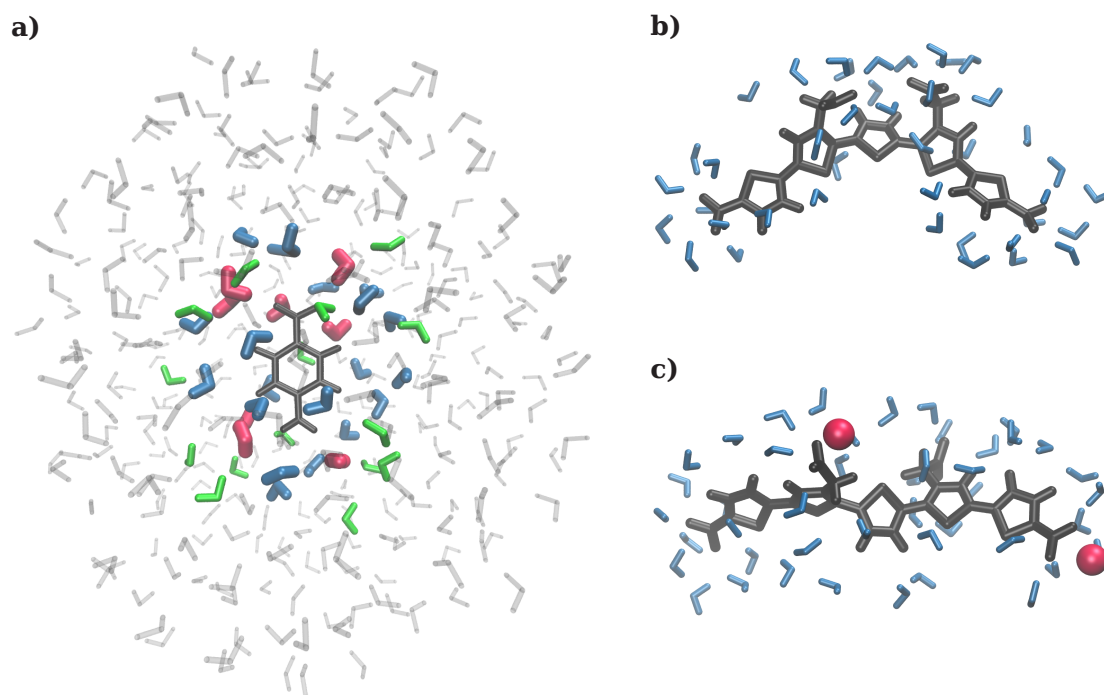
Subsequently to establishing a common theoretical framework for the PE and FDE embedding schemes as well as specific models with different orders of polarization effects included (NOPOL, GSPOL, DPOL and DPOL+EEF) in the previous chapter, the next step involves a transfer to a fair one-to-one comparison in actual calculations. In order to observe the numerical effects of these schemes and models, to fairly represent these and maximize comparability between them, in the following a computational setup is introduced in Section 4.2.1. Subsequently, the numerical comparison is performed by calculations in the newly implemented framework for the previously presented systems pNA and pFTAA (see Chapter 1) solvated in different sizes of a water environment [Fig. 4.2.1 **a**] and a 3 Å water environment [Fig. 4.2.1 **b**) and **c**)], respectively. The obtained results are discussed in Section 4.2.2.

### 4.2.1. Computational Setup and Implementation

The general procedure in the common theoretical framework can be broken down into a few main steps, that are shown in the middle of Fig. 4.2.2. The initial structure of the supermolecule is the input for both models and acquired from performing MD simulations (*cf.* Chapter 3), obtaining a variety of configurational structures and thus ruling out any bias towards a specific solvent configuration.

The supermolecular structure needs to be partitioned into subsystems in the initialization step including the identification and assignment of the main system. The initialization also involves any initial subsystem ground-state calculations, that are necessary in

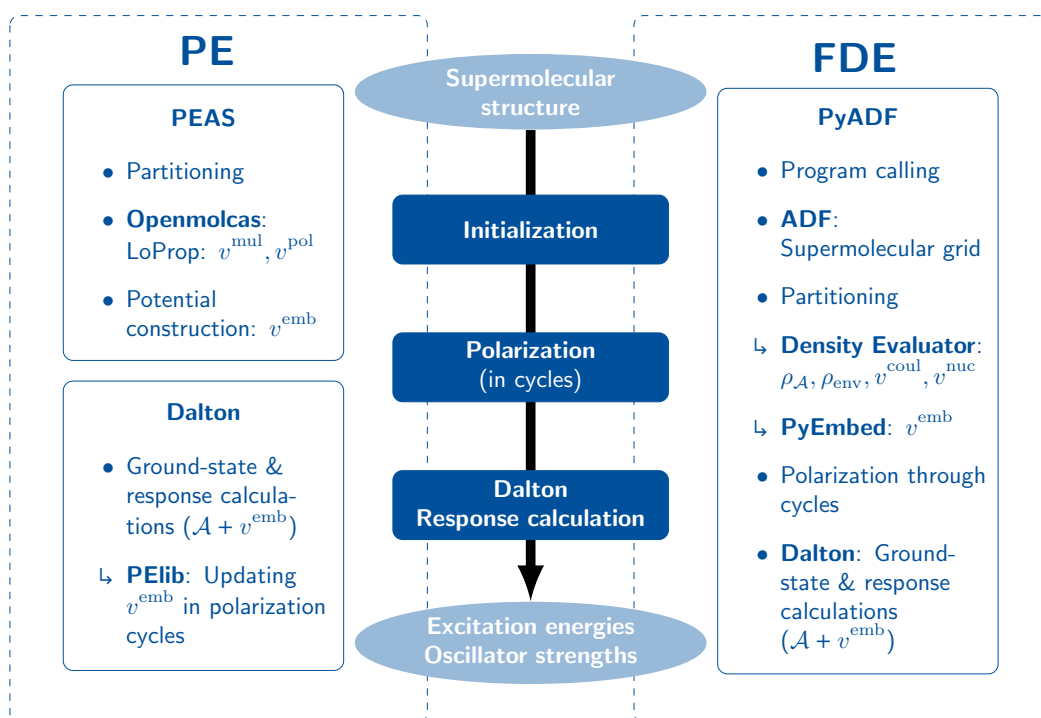
<sup>‡</sup> This section is adapted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689-5703. Copyright 2023 American Chemical Society. This particularly holds true for the outline of the computational setup and the presented results in the numerical comparison.



**Figure 4.2.1.:** **a)** Solvation of *pNA* (black) in a 3 Å (red), 4 Å (blue), 5 Å (green) and 12 Å (grey) water environment on the example of snapshot 3. **b)** and **c)** Solvation of *pFTAA* in a 3 Å water environment (blue). **b)** Snapshot 1, **c)** Snapshot 2, that exhibits sodium ions (red) in the close vicinity to *pFTAA*.

order to determine properties that yield the initial embedding potential. Subsequently, the polarization of the main subsystem can take place. In case of the NOPOL model, this involves only a single ground-state calculation of the main molecule in presence of the embedding potential. For the GSPOL model, further embedding-scheme-specific steps are involved in the further self-consistent polarization. Finally, the polarized main system is then employed in a response calculation in the presence of an embedding potential, yielding the desired optical properties, in this case, excitation energies and corresponding oscillator strengths. Here, possibly further polarization can be accounted for in the response calculations in the PE scheme.

The PE scheme is available in a Dalton framework<sup>[285]</sup> via PEAS<sup>[287]</sup> and the Dalton-implemented library PELib<sup>[294]</sup>. In the present work, the FDE scheme was implemented in a scheme focussed on Dalton, largely complying with the PE implementation. Dalton is in principle able to perform FDE calculations since the inclusion of a static embedding potential in DFT and TD-DFT calculations is possible, however, a FDE freeze-and-thaw formalism is not implemented in the present version, specifically, because the potential construction has to be performed externally. A framework allowing the determination of embedding potentials with the results from Dalton calculations therefore enables the performance of the freeze-and-thaw procedure using a Dalton fundament is introduced and implemented in this work. The embedding-scheme-specific workflows are summa-



**Figure 4.2.2.:** Workflow diagram of the general computational setup. The elliptical nodes represent the input and outputs of the calculation. The dark blue boxes represent the general work steps in both embedding schemes. The white boxes represent the workflow for the specific embedding scheme, with **programs** and **modules** called.

rized in Fig. 4.2.2, however, will be presented in more detail in the following sections (cf. Fig. A.1 for PE and Fig. 4.2.3 for PE).

#### 4.2.1.1. Polarizable Embedding

In the PE scheme (see Fig. A.1), the supermolecular structure is read in by PEAS, that is partitioning the structure into the multiple subsystems, and separating the active system so that a Dalton input file can be produced for this subsystem. For the environmental subsystems, ground-state DFT calculations in Openmolcas are performed in order to obtain the unpolarized subsystem electron densities.<sup>[288]</sup> From these the localized multipoles  $\mathbf{Q}_s[\rho_X^{(0)}]$  (here truncated to second order, *i.e.* quadrupoles) and static polarizabilities  $\alpha_s^0$  are retrieved in LoProp calculations in Openmolcas (Eq. (4.1.53)), which is automatically called.<sup>[289]</sup> PEAS subsequently creates a Dalton-readable potential from the multipoles and polarizabilities.

Dalton is initialized with the main system input file from PEAS. The subsequent calculations then closely follow the workflow presented by Olsen *et al.*<sup>[223]</sup>: Both the ground-state and response calculations are performed with Dalton including the PE embedding potential of the corresponding cycle. In this process in both calculation types, PElib handles the self-consistent calculations of the induced dipole moment (Eq. (4.1.50)) in both types of calculations due to its dependence on the total field changing it in every

SCF step.<sup>[294]</sup> For the NOPOL model, this means that it is represented by optimizing the main subsystem in the obtained multipoles (first term in Eq. (4.1.53)), but overall no polarization terms will be introduced in the calculations. In contrast, in the GSPOL model additionally the self-consistent ground-state polarization is included in the embedding potential *via* PELib (Eq. (4.1.53)) and the subsequent response calculation disregards the differential polarization terms when solving the TD-DFT equations (Eqs. (2.3.12) and (4.1.56)). When changing to the DPOL model, the inclusion of differential polarization by including the corresponding term in the response calculations (Eq. (4.1.56)), is the sole adaptation in comparison to the GSPOL model. Additionally, the effective external field (EEF) effects for the DPOL+EEF model can be added in the response calculation by a simple Dalton keyword.<sup>[285, 349]</sup>

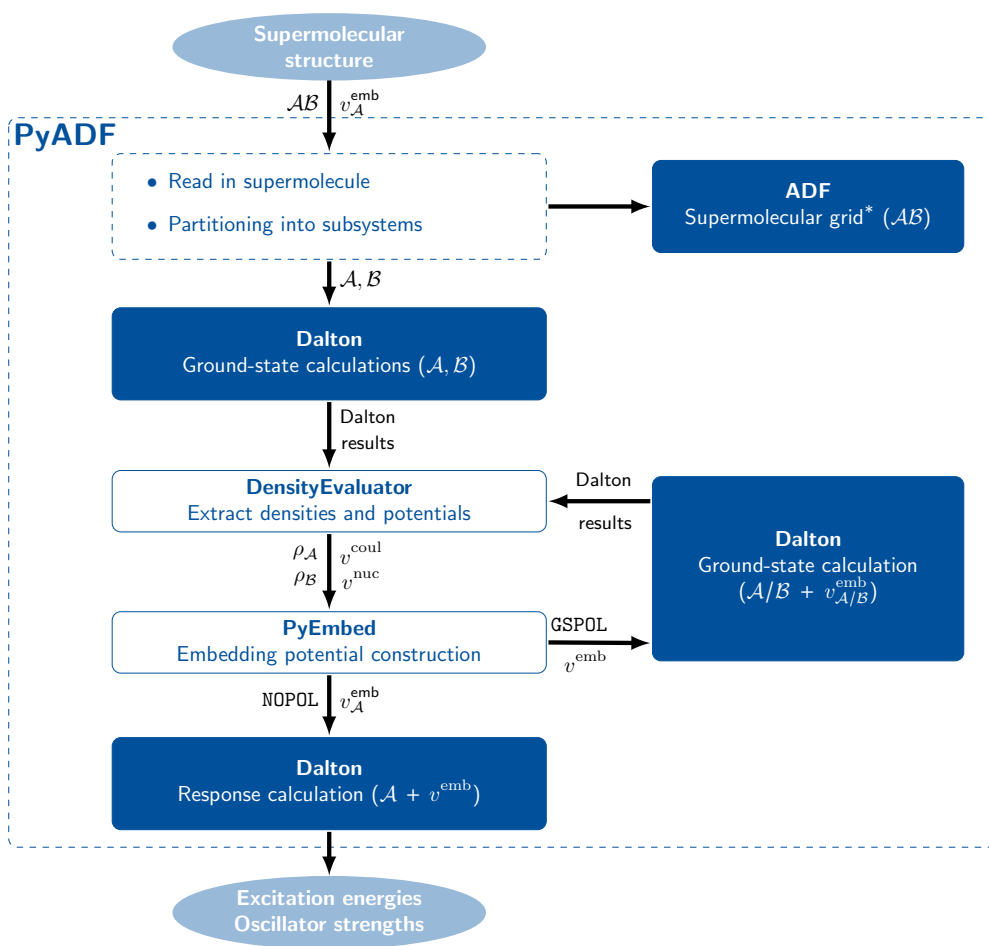
#### 4.2.1.2. Frozen-Density Embedding

In the FDE scheme (Fig. 4.2.3), the whole calculation is managed by the PyADF framework, that is able to call other programs, such as Dalton, and read and convert their results for further calculation steps.<sup>[61, 296]</sup> The implementation and an exemplary input code for both examined systems, as well as the test case for the implementation, are presented in listings A.1, A.2 and A.5 to A.16.

PyADF reads the supermolecular structure and performs a single step of an ADF ground-state DFT calculation in order to obtain a supermolecular grid that is used in all upcoming Dalton calculations. This ensures the homogeneous distribution of grid points over all subsystems involved and circumvents the occurrence of grid artefacts.<sup>[297, 298]</sup> Additionally, ADF calculation grids can be easily read by PyADF and therefore readily be made available for further Dalton calculations, whereas grids obtained from Dalton are not supported in the PyADF framework yet. PyADF then further partitions the supermolecule into subsystems and calls Dalton for ground-state DFT calculations of all obtained subsystems individually and saves the results on the supermolecular grid.

The PyADF module DensityEvaluator then uses the molecular orbital coefficients of the Dalton result files to generate the electron densities, Coulomb and nuclear potentials of all subsystems. These are then employed in the PyEmbed module of PyADF in order to determine the embedding potential for a specific active system involving the evaluation of the non-additive kinetic and XC terms (Eqs. (2.4.6) and (2.4.7)) on the same grid and adding it to the electrostatic terms (Eq. (2.4.10)). In case of the NOPOL model, this initial embedding potential is employed in a ground-state and response calculation of the main system to yield the desired results. It should be noted, that in the response calculation including the embedding potential, only electrostatic contributions enter the TD-DFT equations and the non-additive terms (Eq. (4.1.29)) are disregarded in the response kernel. They do, however, contribute in the performed ground-state calculations.

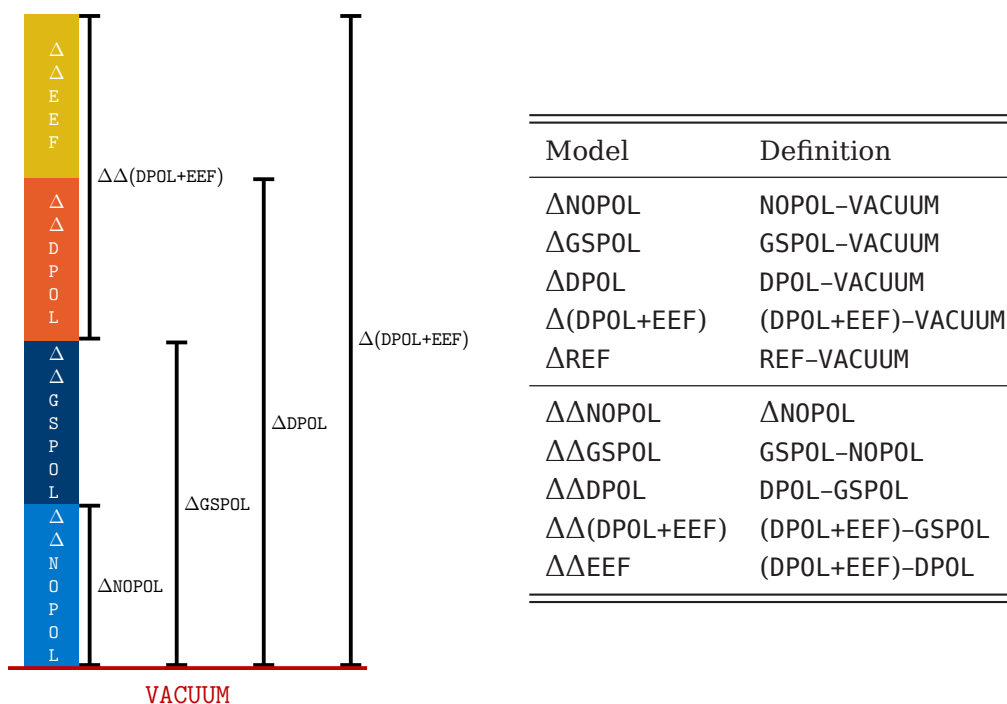
For a further mutual polarization according to the GSPOL model, the role of the active system is further exchanged between all subsystems. The new static embedding potential acting on the respective active subsystem of the current step is updated accordingly to the results from the step performed previously *via* the DensityEvaluator and PyEmbed modules. The embedding potential is then made available to Dalton for ground-state



**Figure 4.2.3.:** Workflow diagram of the computational setup for the FDE scheme. The elliptical nodes represent the input and output of the calculation. The nodes and bullet points in dashed boxes denote processes performed and called by PyADF<sup>[61, 296]</sup>. The dark blue boxes represent **programs** called in the process of the embedding calculation, whereas the white boxes represent called **modules** of PyADF. The arrows denote data exchanged in between modules and programs. Objects marked with \* are employed in all further calculations in all programs and models.

DFT calculations of that subsystem in the presence of the embedding potential. When all subsystems have taken an active role, and thus, have been optimized in the procedure, this is considered as one freeze-and-thaw cycle. The procedure can be continued until self-consistency of a desired property, however, a number of three freeze-and-thaw cycles has shown to be sufficient in many cases.<sup>[300]</sup> The updated data of all systems is then employed to perform a (uncoupled) response calculation of the main system with the updated ground-state embedding potential in order to obtain the desired properties. The further models DP0L and DP0L+EEF were not regarded in this implementation. However, the implementation that was to a large degree performed in Dalton for the NOPOL and GSPOL model enables a fair comparison with the PE scheme for these models.

## 4.2.2. Numerical Comparison



**Figure 4.2.4.:** Overview of solvation model definitions used in the results and their discussion. The upper part in the table denotes full  $\mathcal{S}/\mathcal{F}$ -shifts (in difference to vacuum). The lower part in the table denotes model-specific contributions of  $\mathcal{S}/\mathcal{F}$ -shifts (difference compared to a lower model).

In the numerical comparison, the excitation energies and oscillator strengths for FDE and PE with the different polarization models (NOPOL, GSPOL, DPOL and DPOL+EEF, Section 4.1) were obtained in the new computational setup for pNA in different water environment sizes (Fig. 4.2.1), different basis set sizes as well as pFTAA in a 3 Å water environment with an aug-cc-pVDZ basis set. For pNA and pFTAA the results were obtained for the seven and eight configurations (“snapshots”), respectively, extracted from MD simulations.

The results are generally presented here in the form of shifts, *i.e.* the difference between the excitation energy or oscillator strength of the solvated system in a specific polarization model to that of another calculation *e.g.* the vacuum calculation. These shifts are named  $\mathcal{S}$ -shifts for the changes in the excitation energy and  $\mathcal{F}$ -shifts for changes in the oscillator strength. When the difference is taken between results for the solvated system and the vacuum system, this is denoted  $\Delta$  (Fig. 4.2.4). Accordingly, the  $\mathcal{S}$  and  $\mathcal{F}$ -shifts of the reference calculations ( $\Delta$ REF) represent the excitation energies and oscillator strengths obtained by taking the difference between the respective calculated values and the corresponding vacuum values. Individual contributions of these shifts specific to a model are denoted as  $\Delta\Delta$  and refer to the difference between a value for a polarization model and a lower model, *e.g.* the  $\Delta\Delta$ GSPOL contribution is defined as the

shift caused by the GSP0L model by subtracting the value of the NOP0L model, *i.e.*, the GSP0L-specific shift part of the  $\mathcal{S}$  or  $\mathcal{F}$ -shifts.

#### 4.2.2.1. *para*-Nitroaniline

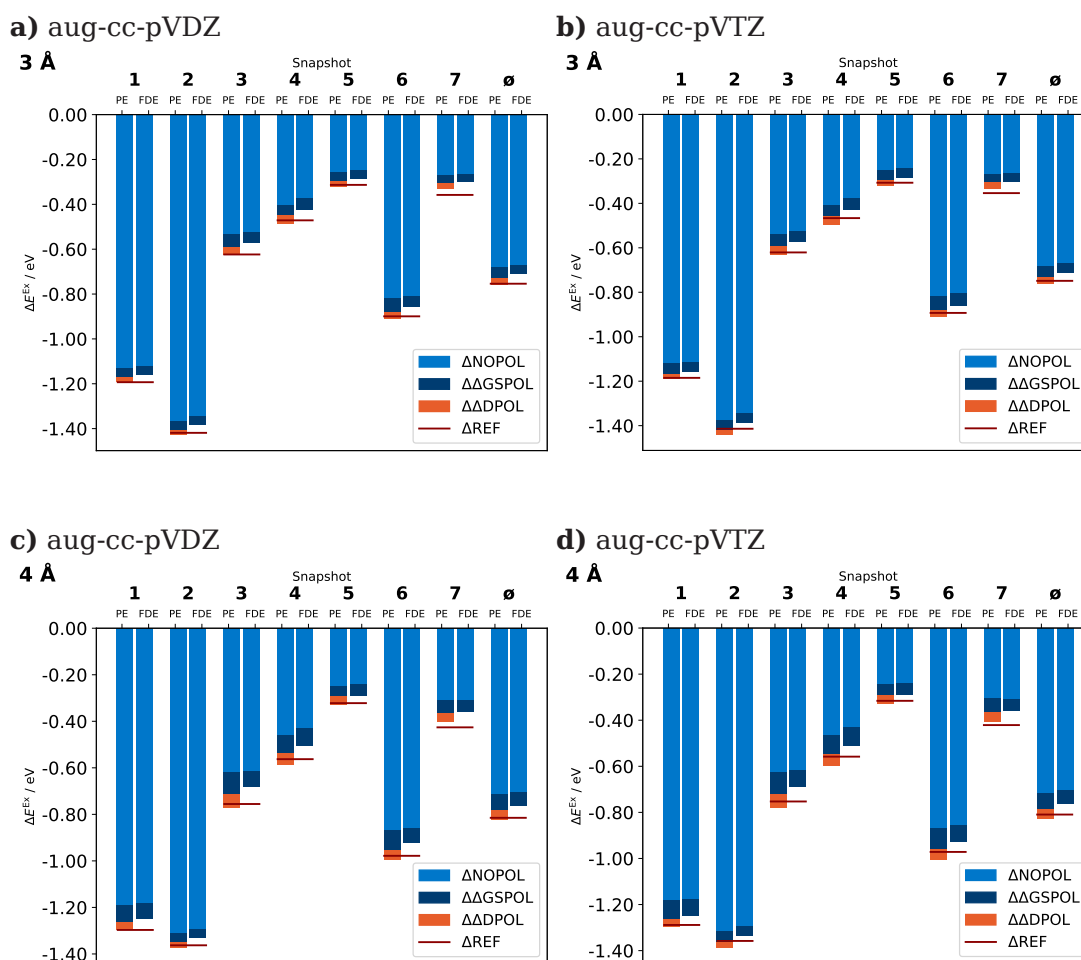
In the following, the results for the calculation *p*NA in a 3 Å water environment are presented and the parameters of the XC functional, that are to be used in the further calculations, are assessed. Subsequently, the results for the excitation energies and oscillator strengths are presented and discussed for different sizes of a water environment and different basis set sizes.

Initially, in the calculation of the supermolecular reference for *p*NA in a 3 Å environment, two different parametrizations of the CAM-B3LYP XC functional were compared. On the one hand, the default parametrization was employed ( $\mu = 0.33$ ), on the other hand, the slope of the error function introducing the Fock exchange in the range-separated XC functional is set to  $\mu = 1.0$ .<sup>[174, 357]</sup> The latter was examined to due the possibly better description of charge-transfer effects when introducing a higher amount of non-local Fock exchange since the default parameters do integrate 65% Fock exchange at long ranges.<sup>[357, 358]</sup> In the following this parametrization will be denoted CAM-B3LYP\*. The results for *p*NA in a 3 Å water environment for an aug-cc-pVDZ and aug-cc-pVTZ basis set are presented in Table A.1 and Table A.2, respectively. The excitation energies that were obtained with CAM-B3LYP\* are on average 0.33 eV higher than those obtained with CAM-B3LYP, independent of the basis set employed. The oscillator strengths of CAM-B3LYP and CAM-B3LYP\* on average differ by 0.019 and 0.018 for an aug-cc-pVDZ and aug-cc-pVTZ basis set, respectively. In comparison to an experimental reference value of the excitation energy for *p*NA of 3.26 eV<sup>[359, 360]</sup>, the excitation energies obtained with CAM-B3LYP deviate less than those obtained with CAM-B3LYP\* so that in the following calculations, the CAM-B3LYP XC functional is used further. Independent of the XC functional employed in the reference calculations, the strongest transition can be identified as the  $\pi \rightarrow \pi^*$  (HOMO $\rightarrow$ LUMO) transition expected from literature (Fig. 4.2.5).<sup>[76, 361]</sup>



**Figure 4.2.5.:** Orbitals involved in the transition with the highest oscillator strength obtained in a CAM-B3LYP/aug-cc-pVTZ calculation for snapshot 3. The occupied orbital is presented on the left (HOMO,  $\pi$ ) and the target orbital on the right (LUMO,  $\pi^*$ ).





**Figure 4.2.6.:**  $S$ -shifts and individual polarization model contributions for  $p\text{NA}$  in a 3 Å water environment (**a**) and **b**) and a 4 Å water environment (**c**) and **d**) for different snapshots and the average values over all snapshots obtained with an aug-cc-pVDZ (**a**) and **c**) and aug-cc-pVTZ basis set (**b**) and **d**). Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

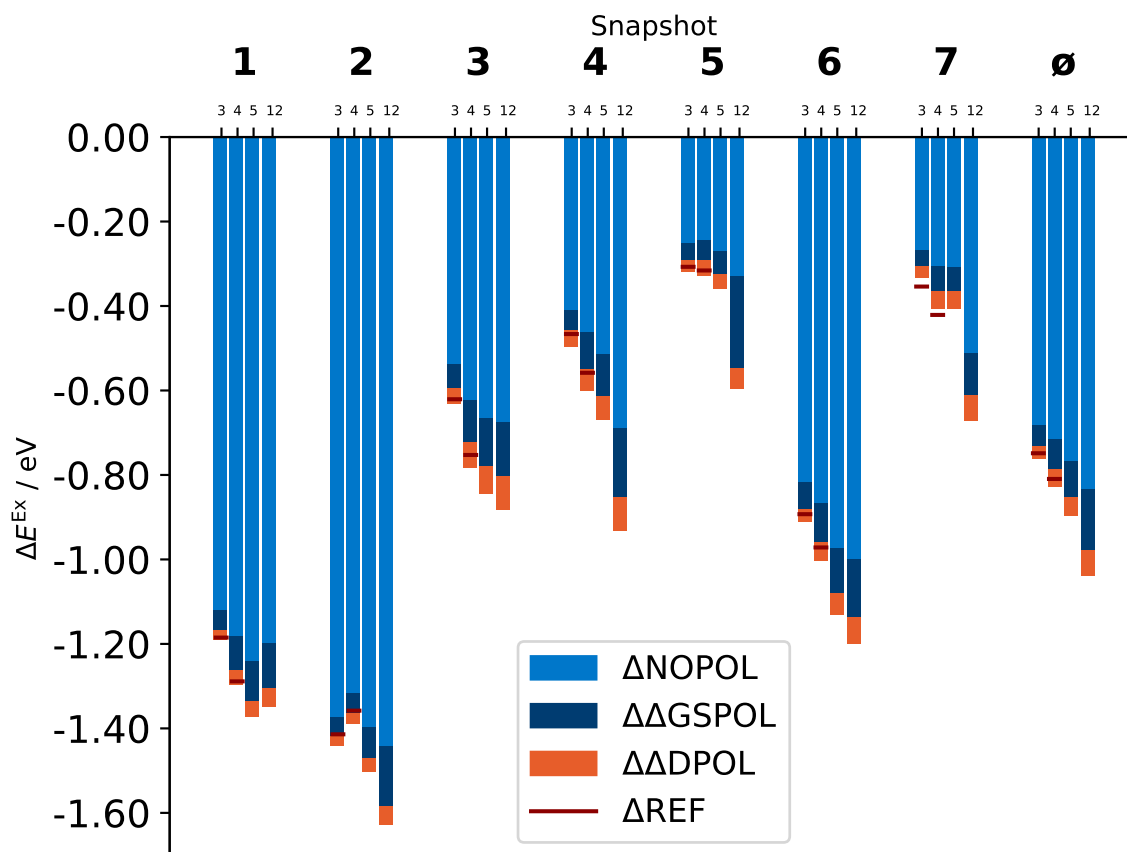
The excitation energies of *p*NA for the seven snapshots were calculated for a 3 and 4 Å environment for the PE and FDE scheme with an aug-cc-pVDZ and aug-cc-pVTZ basis set for all polarization models. The resulting shifts are presented as  $\mathcal{S}$ -shifts according to Fig. 4.2.4 in Fig. 4.2.6 **a)**-**d)** and Chapter A.2.2.1. Since the  $\mathcal{S}$ -shifts for FDE and PE are very similar for  $\Delta$ NOPOL and  $\Delta$ GSPOL for all environment and basis set sizes, some general observations can be made.  $\Delta$ NOPOL makes up the highest contribution of  $\Delta$ REF of the full  $\mathcal{S}$ -shift for all environment sizes and basis sets used, although the absolute values of the individual snapshots vary greatly.  $\Delta$  $\Delta$ GSPOL and  $\Delta$  $\Delta$ DPOL are significantly smaller, *i.e.* on average below 10% and 5% of  $\Delta$ REF, respectively, for both, PE and FDE.

Comparing  $\Delta$ NOPOL to  $\Delta$ REF, the overall shift is underestimated with this polarization model. In a 3 Å environment calculation with an aug-cc-pVDZ basis set, extending to a  $\Delta$ GSPOL model still underestimates the total shift and  $\Delta$ DPOL is necessary in order to achieve good agreement with  $\Delta$ REF (especially for snapshot 7). Increasing the basis set or the environment size,  $\Delta$ NOPOL still underestimates the  $\mathcal{S}$ -shift, but the inclusion of  $\Delta$ GSPOL recovers most polarization from  $\Delta$ REF.  $\Delta$ DPOL in many cases leads to a slight overestimation of the  $\mathcal{S}$ -shift, see *e.g.* snapshots 2, 3, 4 and 6 in Fig. 4.2.6 **d)**, however, for individual snapshots such as snapshot 7 it still shows to be indispensable for good agreement with  $\Delta$ REF. It should also be noted, that the  $\Delta$ DPOL overshooting effect is smaller than the effect expected from using B3LYP for the environmental fragments in the embedding potential calculations, in contrast to full CAM-B3LYP calculations for the supermolecular reference.

Focussing on the increase of the environment, a small difference between the embedding schemes is observed.  $\Delta$  $\Delta$ GSPOL becomes bigger for the PE scheme (on average from 7% to 9%), while for the FDE scheme it remains at 7% on average for the aug-cc-pVTZ basis set. It can also be seen, that the  $\Delta$  $\Delta$ GSPOL contribution is strongly dependent on the individual snapshot, for PE and FDE ranging between 3% and 11% in a 3 Å environment and between 4% and 9% in the 4 Å environment (all were obtained with an aug-cc-pVTZ basis set). However, all these changes are below a threshold of  $-0.1$  eV. Specifically snapshot 7 is very distinct here due to its overall low  $\mathcal{S}$ -shift confirming the overall strong snapshot dependence of the *p*NA calculation.

Due to the low computational effort of the PE scheme, bigger environment sizes of 5 and 12 Å were examined, however, without the possibility to compare to a supermolecular reference calculation. The results are shown in Fig. 4.2.7 (see Chapter A.2.2.1 for results obtained with an aug-cc-pVDZ basis set). In the discussion of these results, the individual relative contributions will be compared to the  $\Delta$ DPOL shift.

Generally,  $\Delta$ DPOL increases in the extension of the environment, where the increase of  $\Delta$  $\Delta$ NOPOL and  $\Delta$  $\Delta$ GSPOL on average is systematic with the system size, but for  $\Delta$  $\Delta$ DPOL it shows a small decrease changing from a 4 Å to a 5 Å environment and an increase from a 5 Å to a 12 Å environment. The contribution of  $\Delta$  $\Delta$ GSPOL in some cases considerably exceeds the average contribution. This specifically holds for snapshots, that show an overall small  $\mathcal{S}$ -shift, such as snapshots 5 or 7. For instance, for snapshot 5 in a 12 Å environment,  $\Delta$ NOPOL decreases to 55% of  $\Delta$ DPOL (average: 81%), while  $\Delta$  $\Delta$ GSPOL takes 37% of  $\Delta$ DPOL (average: 13%) and  $\Delta$  $\Delta$ DPOL takes 8% of  $\Delta$ DPOL (average: 6%).

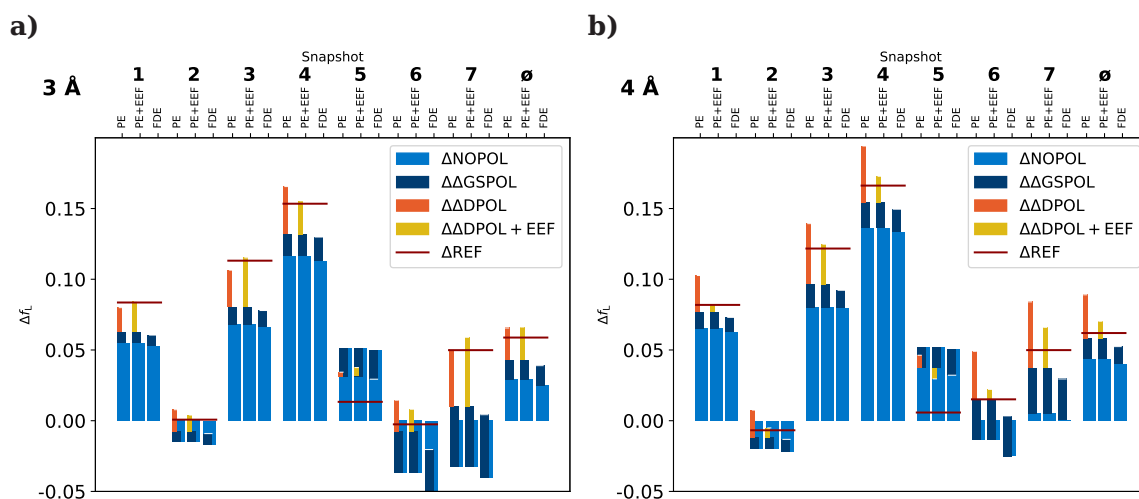


**Figure 4.2.7.:**  $S$ -shifts and individual polarization model contributions of the PE scheme for pNA in different environment sizes for different snapshots and the average values over all snapshots obtained with an aug-cc-pVTZ basis set. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

On the one hand, these effects once more confirm the snapshot-dependence of the calculations, on the other hand, it is in line with other studies having shown that environments of up to 10–15 Å can contribute to polarization effects.<sup>[18, 19, 362]</sup> Thus, an increase in the  $\mathcal{S}$ -shift with the environment size, specifically the  $\Delta\Delta\text{GSPOL}$  can be expected.

Another study with a PE scheme<sup>[77]</sup> and a high number of snapshots has shown higher individual contributions of  $\Delta\Delta\text{GSPOL}$  and  $\Delta\Delta\text{DPOL}$  for a 12 Å environment, however, this could also confirm the high snapshot-dependence, that is found in this study. In comparison to an EOM-CCSD/EFP study, the relative contributions are in good agreement with the results of this study and also showed an increase of  $\Delta\Delta\text{DPOL}$  when increasing the environment size, despite their use of microsolvation (2–6 water molecules in the environment).<sup>[76]</sup> Finally, Daday *et al.*<sup>[65]</sup> employed a state-specific embedding potential ansatz for the FDE scheme, yielding a description of  $\Delta\Delta\text{DPOL}$  that yields similar ranges to those achieved in this study with the PE scheme.

Subsequently to examining the  $\mathcal{S}$ -shifts of excitation energies, in the following the change of oscillator strengths, *i.e.* the  $\mathcal{F}$ -shifts, are presented (Fig. 4.2.8, see Chapter A.2.2.2 for aug-cc-pVDZ results showing similar trends) and discussed. Additionally,



**Figure 4.2.8.:**  $\mathcal{F}$ -shifts and individual polarization model contributions for pNA for different snapshots and the average values over all snapshots obtained with an aug-cc-pVTZ basis set in a) 3 Å and b) 4 Å environment. From left to right every bar shows the results for a PE, PE-EEF and FDE scheme for every snapshot. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

to a PE scheme, a PE-EEF scheme was introduced for the calculation of  $\mathcal{F}$ -shifts (see Section 4.1.3), since the introduction of an EEF influences the magnitude of the field, and thus, the oscillator strengths, which is in contrast to excitation energies, that remain unchanged in this treatment.

Compared to the results for  $\mathcal{S}$ -shifts, the  $\mathcal{F}$ -shifts show a higher sensitivity to different snapshots as well as environmental effects. Specifically, the inclusion of higher

models than NOPOL causes the change of sign in some snapshots and yield much higher contributions than for the  $\mathcal{S}$ -shifts, showing better agreement with  $\Delta_{\text{REF}}$  than  $\Delta_{\text{NOPOL}}$ . The influence of  $\Delta\Delta\text{DPOL}$  (and  $\Delta\Delta\text{EEF}$  effects) is showing a much greater impact than for the  $\mathcal{S}$ -shifts. It can be seen, that the values for PE and FDE for  $\Delta_{\text{NOPOL}}$  and  $\Delta_{\text{GSPOL}}$  on average are in very good agreement with each other for all environment sizes. While  $\Delta\Delta_{\text{NOPOL}}$  was the highest contribution in every case for the  $\mathcal{S}$ -shifts, here, this is often the case and also holds for the average values, however, for individual snapshots, even greater variations occur. Especially for  $\Delta_{\text{NOPOL}}$ , slightly higher differences between PE and FDE can be found. For instance, snapshot 6 shows absolute values for  $\Delta\Delta_{\text{NOPOL}}$  of -0.051 and -0.037 for FDE and PE, respectively, in a 3 Å environment (-0.025 for FDE and -0.014 for PE in a 4 Å environment). For  $\Delta\Delta_{\text{GSPOL}}$ , the values are once more in good agreement at 0.031 for FDE and 0.029 for PE for this snapshot. In comparison to the reference, the  $\Delta_{\text{GSPOL}}$  values excellently reproduce  $\Delta_{\text{REF}}$ .  $\Delta\Delta\text{DPOL}$  is showing high contributions of 0.022 and 0.034 for a 3 and 4 Å environment for PE. The addition of  $\Delta\Delta\text{DPOL}$  or  $\Delta\Delta\text{EEF}$ , do in fact improve the  $\mathcal{F}$ -shift in the correct direction but  $\Delta\text{DPOL}$  often shows an overestimation of this effect. It should, however, be noted that  $\Delta\Delta\text{EEF}$  slightly decreases this overshoot of  $\Delta\Delta\text{DPOL}$ . Additionally, this snapshot is an example for an opposing sign of the  $\Delta_{\text{NOPOL}}$  and  $\Delta\Delta_{\text{GSPOL}}$  contributions, which occurs in four of the snapshots for a 3 Å environment and three snapshots for a 4 Å environment.

Generally, it can be stated, that  $\Delta\Delta_{\text{GSPOL}}$  strongly improves agreement with  $\Delta_{\text{REF}}$  but in most cases underestimates the  $\mathcal{F}$ -shift. As observed for snapshot 6,  $\Delta\Delta\text{DPOL}$  corrects in the correct direction in all but one case (snapshot 5) and  $\Delta\Delta\text{EEF}$  decreases the overestimation occurring in three cases for the 3 Å environment and as many as six cases and the average value for the 4 Å environment. On average the  $\mathcal{F}$ -shift shows better agreement with  $\Delta_{\text{REF}}$  for  $\Delta(\text{DPOL}+\text{EEF})$  than for  $\Delta_{\text{GSPOL}}$  and specifically for the 4 Å environment  $\Delta\text{DPOL}$  yields results deviating more from  $\Delta_{\text{REF}}$  than the  $\Delta_{\text{GSPOL}}$  results. Here, the use of an EEF is indispensable, when introducing the DPOL model.

Overall, deviations between the two embedding models remain small, suggesting that QM effects do not contribute strongly for this system. For the calculation of  $\mathcal{S}$ -shifts mutual polarization yielded with the GSPOL model is often sufficient in order to reproduce the results for the supermolecular reference calculation for both models. The DPOL model did not obtain significantly better results for the  $\mathcal{S}$ -shifts, but showed a higher impact on the  $\mathcal{F}$ -shift. In contrast to the other models, it also did not show systematic increase with the system size. Additionally, the use of an EEF can decrease the overestimation of the DPOL model's  $\mathcal{F}$ -shifts in comparison to reference (specifically in a bigger environment) and – similarly to results for the GSPOL model – yield  $\mathcal{F}$ -shifts that are in excellent agreement with the reference values.

#### 4.2.2.2. Pentameric Formyl Thiophene Acetic Acid

Having performed a numerical comparison for pNA, the highly anionic system pFTAA [Fig. 4.2.1 **b**) and **c**)] was the next target system examined in this study. The quadruple negative charge in combination with sodium ions occurring in the 3 Å environment of four of the eight snapshots from MD simulations make this system more challenging for

QM/MM schemes such as PE due to their lack of repulsive terms in the environment. For FDE partial repulsion is expected to be recovered in the non-additive potential terms.

For snapshot 1 [Fig. 4.2.1 **b**], that does not show any sodium ions in close vicinity of pFTAA, the results and the orbitals involved in the transitions are presented in Table A.27 and Figs. A.5 and A.6. In all cases, one very strong  $\pi \rightarrow \pi^*$  transition is found, that is in line with literature.<sup>[359, 360]</sup> The excitation energy deviates by 0.03 eV at most from the reference value for all embedding schemes and models. The oscillator strengths show an increase for both embedding schemes when introducing higher polarization models.

Examining snapshot 2 [Fig. 4.2.1 **c**], where two sodium ions are present at an average distance of 2.3 Å to two of the carboxyl groups of pFTAA, problems occur. The results for this snapshot are exemplary presented in Table 4.2.1 and Fig. 4.2.9. The full data can be found in Figs. A.7 to A.9 and Table A.28. For FDE, especially the strongest transition is calculated correctly and as the first transition. There is some ESO effects occurring in the FDE GSPOL model, and the third strong transition (4.03 eV), which can be explained by lack of accuracy in the treatment of non-additive effects of the environment. However, the strongest transition, that is the centre of this discussion is not affected by this effect. For the PE model and specifically the NOPOL model, strong ESO effects occur, slightly decreasing for the higher polarization models. In conclusion, the first strong transition is calculated in the 10th or higher response vector. For the NOPOL model, the oscillator strength is also significantly lower than it would be expected from the reference calculation and shows a lower excitation energy. For the GSPOL and DPOL models, the excitation energy for the strongest transition is in good agreement with the strongest transition from reference calculations. The oscillator strengths improve as well, however, overestimate in comparison to reference and the FDE results. Also for other strong transitions ESO effects occur. The involved orbitals are shown in Fig. 4.2.9, presenting the characteristic accumulation on one of the carboxylic groups of pFTAA, that is in close vicinity to a sodium ion in the excitation on the example of the NOPOL model.

To alleviate the ESO effects for PE, atomic pseudopotentials are employed for the ions in the environment of pFTAA in order to introduce repulsion and obtain feasible results.<sup>[354]</sup> The results for the PE scheme including the potentials are presented in Table 4.2.2 for the excitation energies and  $\mathcal{S}$ -shifts and in Table 4.2.4 for the oscillator strengths and  $\mathcal{F}$ -shifts. Since the computational effort for the PE scheme is lower in this implementation, additional snapshots could be examined (Tables 4.2.1, 4.2.3 and 4.2.5).

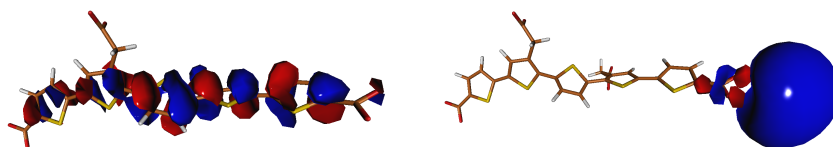
As can be seen in the overview for PE in Table A.29 and Fig. A.10, for snapshot 2 all strong transitions obtained by the inclusion of pseudopotentials lead to improved excitation energies compared to the reference calculation as well as meaningful transitions ( $\pi \rightarrow \pi^*$ ) for every model. Accordingly, the oscillator strength improved, specifically for the NOPOL model, but also for the GSPOL and DPOL model, where the values are in better agreement with the values obtained with the FDE scheme and reference calculation.

Comparing the PE and FDE scheme for the common snapshots 1 and 2 in Table 4.2.2, the differences between the models are small. For the  $\Delta$ NOPOL, a deviation of 0.01 eV and 0.03 eV is observed for snapshots 1 and 2, respectively. The values for  $\Delta\Delta$ GSPOL coincide for snapshot 1 and for the two embedding schemes deviate by 0.03 eV for

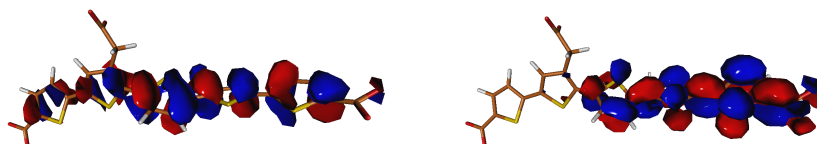
**Table 4.2.1.:** Orbitals involved in the strongest transitions of snapshot 2 of pFTAA for the PE NOPOL scheme with and without atomic pseudopotentials (PP) for sodium ions in the environment and the supermolecular reference calculation. Numbers in brackets denote the calculated contributions of involved orbitals in a transition if more than one strong contribution occurred. Orbitals affected by an ESO are denoted ESO.

Model	Transition	Excitation energy / eV	Oscillator strength	Orbitals
PE NOPOL	12	3.02	0.3080	$\pi \rightarrow \text{ESO}$
	15	3.22	0.4151	$\pi \rightarrow \pi^*$ (0.39) $\pi \rightarrow \text{ESO}$ (0.32) $\pi \rightarrow \text{ESO}$ (0.30)
	16	3.24	0.6427	$\pi \rightarrow \pi^*$
	20	3.72	0.2196	$\pi \rightarrow \pi^*$
PE NOPOL + PP	1	3.18	1.3077	$\pi \rightarrow \pi^*$
	2	3.59	0.1480	$\pi \rightarrow \pi^*$
	3	3.90	0.1999	$\pi \rightarrow \pi^*$
Reference	1	3.20	1.1573	
	2	3.27	0.1764	
	3	3.66	0.1564	
	5	4.01	0.1831	

PE NOPOL:  
(12)  
 $\pi \rightarrow \text{ESO}$



PE NOPOL:  
(15)  
 $\pi \rightarrow \pi^*$  (0.39)



$\pi \rightarrow \text{ESO}$  (0.32)



**Figure 4.2.9.:** Orbitals involved in the strongest transition of pFTAA in the PE NOPOL model for snapshot 2. Numbers in brackets after the model denote the calculated transition shown. Numbers in brackets after the orbital-type transition denote the contribution of the orbitals to the calculated transition.

snapshot 2. In total,  $\Delta\text{GSP0L}$  is the same for both embedding schemes and deviates by 0.03 eV for snapshot 1 and by 0.04 eV for snapshot 2. The introduction of the  $\Delta\Delta\text{DP0L}$  contribution in the PE scheme, that shows an opposing sign to  $\Delta\text{GSP0L}$ , gives an even better agreement with  $\Delta\text{REF}$  (deviation of 0.01 eV) in comparison to the already excellent agreement for  $\Delta\text{GSP0L}$ .

For the additional snapshots examined with the PE scheme (Table 4.2.3) the introduction of pseudopotentials alleviated occurring ESO effects and gave meaningful results independent of the number of sodium ions (one or two) in the close vicinity to pFTAA. On average  $\Delta\text{NOP0L}$  is the smallest contribution to the  $\mathcal{S}$ -shift, followed by  $\Delta\Delta\text{DP0L}$  and the highest contribution,  $\Delta\text{GSP0L}$ . It should be noted, that on average the  $\Delta\text{DP0L}$  contribution has an opposing sign to the other contributions. For individual snapshots, there is a great variation in contrast to the observed trends. For instance, the  $\Delta\text{NOP0L}$  contribution ranges between -0.14 eV and 0.16 eV,  $\Delta\text{GSP0L}$  ranges between 0 eV and 0.17 eV and  $\Delta\Delta\text{DP0L}$  ranges between -0.08 eV and 0.02 eV.

**Table 4.2.2.:** Contributions from the different models to the total  $\mathcal{S}$ -shift in eV for different configurations of pFTAA in a 3 Å environment of water obtained from an MD simulation and subsequently calculated in a PE and FDE framework and different orders of polarization contributions. All snapshots containing (two) sodium ions in close proximity are marked with \*\* and incorporate pseudopotentials in the PE calculations. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\mathcal{S}$	FDE			PE					$\Delta\text{REF}$
	$\Delta\text{NOP0L}$	$\Delta\text{GSP0L}$	$\Delta\text{GSP0L}$	$\Delta\text{NOP0L}$	$\Delta\text{GSP0L}$	$\Delta\Delta\text{DP0L}$	$\Delta\text{GSP0L}$	$\Delta\text{DP0L}$	
Snaps. 1	0.15	0.01	0.17	0.16	0.01	-0.03	0.17	0.14	0.14
Snaps. 2**	0.04	0.02	0.07	0.01	0.05	-0.08	0.07	0.04	0.03

The obtained  $\mathcal{F}$ -shifts for pFTAA comparing the FDE and PE scheme, shown in Table 4.2.4, show once more very similar results for  $\Delta\text{NOP0L}$  and  $\Delta\text{GSP0L}$  (deviation 0.026 or lower for both models). In comparison to  $\Delta\text{REF}$ , FDE  $\Delta\text{GSP0L}$  shows a deviation of 0.04 and 0.19 for snapshots 1 and 2, respectively, whereas PE  $\Delta\text{GSP0L}$  shows a deviation of 0.01 and 0.17. The inclusion of  $\Delta\Delta\text{DP0L}$  in the PE scheme corrects the  $\mathcal{S}$ -shift in the correct direction but overestimates the effect, worsening the agreement with  $\Delta\text{REF}$ .  $\Delta\Delta\text{EEF}$  improves the  $\Delta\text{DP0L}$  result but gives worse agreement with  $\Delta\text{REF}$  than the value for  $\Delta\text{GSP0L}$ .

Extending the PE calculations to more snapshots, the average results show similar contributions of the same sign for  $\Delta\text{GSP0L}$  and  $\Delta\Delta\text{DP0L}$ . The average  $\Delta\Delta\text{EEF}$  contribution shows an almost coinciding value with these models but is of opposing sign. The deviation of the reference on average is 0.073 for  $\Delta\text{GSP0L}$ , worsens for  $\Delta\text{DP0L}$  to 0.136 and gives the best results for  $\Delta(\text{DP0L}+\text{EEF})$  with a deviation of 0.071. Individual snapshots show a higher variation of  $\Delta\text{NOP0L}$  values (-0.125 to 0.497), than for  $\Delta\text{GSP0L}$  (0.0 to 0.150),  $\Delta\text{DP0L}$  (0.034 to 0.084) and  $\Delta\Delta\text{EEF}$  (-0.029 to -0.094). The general trend of  $\Delta\Delta\text{DP0L}$  worsening the results in comparison to  $\Delta\text{REF}$  and  $\Delta\Delta\text{EEF}$  improving them can



**Table 4.2.3.:** Excitation energies and contributions from the different polarization models to the  $\mathcal{S}$ -shifts and their average for different configurations of pFTAA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a PE embedding model (and optional EEF). All snapshots containing one or two sodium ions in close proximity are marked with \* and \*\*, respectively, and incorporate pseudopotentials in the PE calculations. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}$	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snaps. 1	2.92	3.08	0.16	3.09	0.01	3.06	-0.03	0.17	0.14	3.05	0.14
Snaps. 2**	3.17	3.18	0.01	3.24	0.05	3.16	-0.08	0.07	-0.02	3.20	0.03
Snaps. 3*	2.78	2.68	-0.09	2.85	0.17	2.82	-0.03	0.07	0.04	2.64	-0.13
Snaps. 4*	2.98	2.90	-0.08	2.97	0.07	2.95	-0.02	-0.01	-0.04	2.90	-0.09
Snaps. 5	3.20	3.31	0.11	3.32	0.02	3.29	-0.04	0.13	0.09	3.28	0.08
Snaps. 6	3.36	3.46	0.10	3.46	0.00	3.43	-0.03	0.10	0.07	3.42	0.06
Snaps. 7	2.74	2.80	0.06	2.83	0.03	2.79	-0.04	0.09	0.05	2.80	0.06
Snaps. 8**	3.05	2.91	-0.14	2.95	0.03	2.91	-0.03	-0.10	-0.14	2.95	-0.10
$\emptyset$	3.02	3.04	0.02	3.09	0.05	3.05	-0.04	0.07	0.03	3.03	0.01

**Table 4.2.4.:** Contributions from the different models to the total  $\mathcal{F}$ -shifts for different configurations of pFTAA in a 3 Å environment of water obtained from an MD simulation and subsequently calculated in a PE and FDE framework and different orders of polarization contributions obtained with an aug-cc-pVDZ basis set. All snapshots containing (two) sodium ions in close proximity are marked with \*\* and incorporate pseudopotentials in the PE calculations. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\mathcal{F}$	FDE			PE							$\Delta$ REF
	$\Delta$ NOPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ NOPOL	$\Delta\Delta$ GSPOL	$\Delta\Delta$ DPOL	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ (DPOL+EEF)	
Snaps. 1	0.294	0.053	0.348	0.320	0.060	0.063	-0.029	0.380	0.443	0.414	0.388
Snaps. 2**	0.140	0.044	0.183	0.143	0.024	0.052	-0.076	0.167	0.219	0.143	-0.007

**Table 4.2.5.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pFTAA in aug-cc-pVDZ quality in a 3 Å water environment obtained with a PE embedding model (and optional effective external potential (EEF)). All snapshots containing one or two sodium ions in close proximity are marked with \* and \*\*, respectively, and incorporate pseudopotentials in the PE calculations. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society..

$\Delta f_L$	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	DPOL+EEF	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ DPOL+EEF	REF	$\Delta$ REF
Snaps. 1	0.900	1.221	0.320	1.280	0.060	1.343	0.063	1.314	-0.029	0.380	0.443	0.414	1.288	0.388
Snaps. 2**	1.165	1.308	0.143	1.332	0.024	1.384	0.052	1.308	-0.076	0.167	0.219	0.143	1.157	-0.007
Snaps. 3*	1.150	1.076	-0.074	1.226	0.150	1.284	0.058	1.212	-0.072	0.076	0.134	0.0617	1.086	-0.064
Snaps. 4*	1.214	1.089	-0.125	1.197	0.108	1.274	0.078	1.214	-0.060	-0.017	0.061	-0.0001	1.065	-0.148
Snaps. 5	1.242	1.439	0.197	1.472	0.034	1.556	0.084	1.518	-0.038	0.231	0.315	0.276	1.504	0.263
Snaps. 6	0.802	1.299	0.497	1.350	0.000	1.422	0.071	1.348	-0.074	0.548	0.620	0.546	1.274	0.472
Snaps. 7	1.094	1.282	0.188	1.358	0.077	1.420	0.061	1.352	-0.068	0.265	0.326	0.258	1.340	0.246
Snaps. 8**	1.359	1.470	0.111	1.500	0.030	1.534	0.034	1.440	-0.094	0.141	0.175	0.081	1.416	0.057
$\emptyset$	1.116	1.273	0.157	1.339	0.060	1.402	0.063	1.338	-0.064	0.224	0.287	0.222	1.266	0.151

also be observed for the individual snapshots.

In conclusion, both embedding models were able to well reproduce the reference values. For this, ESO effects could be efficiently circumvented by the use of atomic pseudopotentials. A snapshot dependence could be shown specifically for the NOPOL model for both  $\mathcal{S}$  and  $\mathcal{F}$ -shifts. For pFTAA,  $\Delta$ GSPOL generally gives a good description of the  $\mathcal{S}$ -shifts and  $\mathcal{F}$ -shifts in both embedding schemes. For the  $\mathcal{F}$ -shifts in the case of using a PE DPOL model, the introduction of EEF effects is strongly recommended, since these improved the results in every examined case.

### 4.3. Summary

Overall, a common theoretical framework for FDE and PE was derived, successfully showing the commonalities and disparities of these embedding schemes on a theoretical basis targeting the same effects from two different philosophies. On the one hand, PE, a QM/MM scheme, accounts for environmental effects by an embedding potential obtained from a classical multipole expansion and anisotropic dipole polarizabilities on environmental sites. Dynamical response effects are partially accounted for by introducing multipoles and polarizabilities *via* the embedding potential into the linear-response framework. On the other hand, the original FDE formulation is based on real-space electron densities, which allows the approximate description of QM effects due to the QM/QM ansatz for constructing the embedding potential. In the response calculations, these QM terms also contribute, but no dynamical response of the environment in the excitation is accounted for in the uncoupled treatment. Therefore, the performance strongly depends on the occurrence of ESO effects and dynamical-response effects.

Comparing the embedding schemes on a numerical basis, the previously presented theoretical framework had to be implemented to enable a fair comparison. For this, the FDE scheme was implemented in a similar fashion to the available Dalton-based

implementation for PE extending the comparability in contrast to previous studies. On this fundament, the presented theoretical polarization models could be directly applied and dissected in the computational setup, and thus, in the following be employed in the numerical comparison for the two examined target cases: *p*NA and *p*FTAA in an explicit water environment.

In the comparison of  $\mathcal{S}$ - and  $\mathcal{F}$ -shifts (solvation shifts of the excitation energy and oscillator strength, respectively), the polarization models of the embedding schemes could be analyzed in detail: electrostatic, ground-state polarization, differential polarization as well as QM effects and EEF effects. An overall similar performance of PE and FDE schemes was found for the electrostatic and ground-state polarization models, the two main ingredients for reproducing polarization effects of a supermolecular calculation. For the  $\mathcal{S}$ -shifts the results including these models were sufficient to obtain good agreement with reference calculations for both systems, *p*NA and *p*FTAA. Differential polarization effects did not show a great contribution to the reproduction of the  $\mathcal{S}$ -shift.

In the systematic analysis of  $\mathcal{F}$ -shifts, a novel dissection of polarization effects, the impact of differential polarization is higher than for  $\mathcal{S}$ -shifts. Although not showing a big influence on the average results for *p*NA, a high dependence on individual snapshots was observed, that showed a partially sizeable influence of the differential polarization effects. In these cases, the introduction of EEF effects decreased the deviation from reference values. While the  $\mathcal{S}$ -shifts overall support the findings for *p*NA, for the  $\mathcal{F}$ -shifts good results could be obtained with ground-state polarization and with differential polarization, however, EEF effects showed to be indispensable for the reproduction of the polarization effects obtained in reference calculations.

For *p*FTAA it could further be shown, that in this highly-anionic system with sodium ions in close vicinity, ESO effects strongly impacted the PE results. These effects could greatly be alleviated by the use of atomic pseudopotentials placed on the sodium ions in the environment in order to account for the lack of repulsion effects in the classical embedding potential (obtained *via* QM calculations). Since for the FDE results in some transitions electronic-spill-out effects could be observed (but did not influence the results discussed here), this suggests *p*FTAA to be a more challenging test case for embedding schemes in general, aside from computational demands due to the sheer system size.

Altogether, extending previous studies, a thorough investigation of dissected polarization effects could be performed for a common theoretical and computational framework for both a neutral molecule exhibiting a charge-transfer effect under solvation as well as a LCO that proved to be highly challenging due to its size, high anionic charge and cations in close vicinity. This study confirms the necessity of employing atomic pseudopotentials in order to circumvent ESO effects in order to obtain feasible results for the PE scheme. On this basis, the performance of FDE and PE was similar for both target systems and ground-state polarization effects and well reproducing reference calculation results. Differential polarization strongly impacting  $\mathcal{F}$ -shifts in the *p*FTAA system only showed good agreement with reference calculations when introducing EEF effects. It is therefore recommended to either fully neglect differential polarization or incorporate EEF effects.



## 5 | Refractive Indices of Zeolitic Imidazolate Frameworks via Fragmentation

From this, *via* the Lorenz-Lorentz equation [Eq. (2.4.30)] the frequency-dependent refractive index can be calculated. Utilizing this approach, a fragmentation ansatz for MOFs was proposed by Treger *et al.* earlier in 2023.<sup>[129]</sup> Here, the refractive indices are calculated from the static polarizability on the example of several ZIFs, presenting in different topologies as well as Zr-based MOFs. In the presented BU approach with disjoint fragments only, the static polarizability was calculated with CP-SCF suggesting a MPW1X/def2-QZVP protocol. Refractive indices were then determined with the Lorenz-Lorentz equation [Eq. (2.4.30)], yielding overall good agreement with data from pDFT and experimental data<sup>[35]</sup>. Due to the non-occurrence of resonances in the visible region (proven by the transparency of MOFs viable for optical applications), they assume that the refractive index shows little dependence on the frequency and infer, that the calculated static refractive index can be compared to the experimental data obtained at a wavelength of 589 nm.<sup>[35]</sup> Moreover, in the fragmentation, solely monomer terms and no many-body terms were considered [*cf.* Eq. (2.4.28)].

In this study, the fragmentation approach by Treger *et al.* is extended to frequency-dependent polarizabilities and higher fragmentation orders on the example of ZIFs which present in a sodalite topology (ZIF-8, ZIF-71, ZIF-90 and ZIF-318, displayed in Fig. 1.3.2).<sup>[96, 102-104]</sup> Firstly, a sophisticated study on optimization and calculation parameters is performed for ZIF-8 and discussed in comparison to experimental data<sup>[35, 363]</sup> and data from pDFT<sup>[129]</sup>. For this, new fragmentation schemes and a computational protocol for the calculation of the refractive index of ZIFs are established. This includes calculation parameters and optimization schemes in order to maximize accuracy and computational efficiency and are discussed in detail in Sections 5.1.2 and 5.1.3. The established protocol is then transferred to the other ZIFs in Section 5.1.4. Subsequently, the influence of guest molecules loaded into the pore of ZIFs is incorporated in the fragmentation ansatz and computational protocol. The resulting tuning effect on the refractive index for ZIF-8 and ZIF-71 from these novel approaches is discussed in comparison to the reference data in Section 5.2.<sup>[35, 364]</sup>

Parts of the calculations discussed below were performed by students under the supervision of the author of this thesis at the Leibniz University Hannover. This *inter alia* refers to the XC functional and basis set benchmark on ZIF-8 in ORCA (Section 5.1.2

and Appendices B.2.1 and B.2.2 that was performed by Onno Strolka in 2022.<sup>‡</sup> Pascal Czember conducted calculations using ORCA and TURBOMOLE for different XC functionals and basis sets investigating different optimization and loading schemes for ZIF-8 and partially for ZIF-71 in 2022 (Sections 5.1.2, 5.1.3 and 5.2.2.1, Appendix B.2.2, Table B.3, and Figs. B.5, B.6 and B.8).<sup>||</sup> Daniel Bömke performed initial TURBOMOLE calculations on the unloaded ZIF-90 as well as the unloaded and loaded ZIF-71 in an internship in 2022 (Section 5.2.2.2 and Figs. B.5 and B.6).<sup>§</sup> Furthermore, Erona Shabani performed TURBOMOLE TD-DFT calculations and geometry optimizations with the previously established fragmentation and calculation schemes as well as new optimization schemes for ZIF-8, ZIF-71, ZIF-90 and ZIF-318 in 2023 (Sections 5.1.4, 5.2.2.1 and 5.2.2.2).<sup>♣</sup> In order to highlight the detailed students' contributions in the following chapters, results that involve their work are marked with their respective footnotes introduced here.

## 5.1. Refractive Indices for Zeolitic Imidazolate Frameworks

In the first part of this section, the fragmentation approaches and details on the calculation of the refractive index from these for empty ZIFs are introduced (Section 5.1.1). The resulting schemes are then tested on the calculation of polarizability volumes and refractive indices on the example of ZIF-8 and from this, a computational protocol is developed (Sections 5.1.2 and 5.1.3). The established protocol is then transferred to the calculation of refractive indices for other ZIFs with sodalite topology in Section 5.1.4. All results are discussed and compared to experimental and computational reference data.<sup>[35, 129, 363]</sup>

### 5.1.1. Fragmentation Schemes

As described in Section 2.4.2.1 for the energies, the framework of a MBE can be transferred to the calculation of the total isotropic polarizability volumes  $\alpha'$  in by a BU [*cf.* Eq. (2.4.28)] or TD formalism, respectively,

$$\begin{aligned}\alpha'^{\text{BU}} &= \sum_X \alpha'_X + \sum_{X<Y} \Delta\alpha'_{XY} + \sum_{X<Y<Z} \Delta\alpha'_{XYZ} + \dots, \\ \alpha'^{\text{TD}} &= \sum_X \alpha'_X - \sum_{X<Y} \alpha'_{XnY} + \sum_{X<Y<Z} \alpha'_{XnYnZ} - \dots\end{aligned}\quad (5.1.1)$$

<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.

<sup>§</sup> Daniel Bömke, Leibniz University Hannover, 2022.

<sup>♣</sup> Erona Shabani, Leibniz University Hannover, 2023.

with

$$\Delta_{XY}^{\alpha'} = \alpha'_{XY} - (\alpha'_X + \alpha'_Y) \quad (5.1.2)$$

$$\Delta_{XYZ}^{\alpha'} = \alpha'_{XYZ} - (\Delta_{XY}^{\alpha'} + \Delta_{XZ}^{\alpha'} + \Delta_{YZ}^{\alpha'}) - (\alpha'_X + \alpha'_Y + \alpha'_Z). \quad (5.1.3)$$

Here,  $\alpha'_{XY}$  and  $\alpha'_{XYZ}$  refer to the polarizability volumes calculated from the combined subsystems  $X, Y$  and  $X, Y, Z$ , respectively. In contrast to Eq. (2.4.28), these approaches include many-body effects. The Lorenz-Lorentz equation [Eq. (2.4.30)] can then be employed to determine the refractive indices from the total isotropic polarizability. The latter is obtained from the respective fragmentation scheme, and thus, the individual isotropic fragment polarizabilities. In the case of ZIFs, the total isotropic polarizability volume refers to that of a unit cell. Therefore, the other variable in this equation, the number density  $N$ , can be recognized as the reciprocal unit cell volume  $V^{-1}$ . The Lorenz-Lorentz equation is accordingly adapted to,

$$n = \sqrt{\frac{1 + 2\frac{\alpha'4\pi}{3V}}{1 - \frac{\alpha'4\pi}{3V}}}. \quad (5.1.4)$$

The unit cell volume for the different ZIFs can be obtained from single crystal X-ray diffraction (SC XRD) experiments at low temperatures (Table 5.1.1). This circumvents any potential temperature-dependent influence on the refractive index and accounts for the fact that energy-minimized fragments were employed in the calculation.<sup>[129]</sup> For some ZIFs, the synthesis is performed by the substitution of linkers from another ZIF. For ZIF-318 this results in the fact, that not the ideally assumed full substitution is yielded in the experiment and therefore the volume reported in Table 5.1.1 does not correspond to the phase-pure ZIF-318 structure.<sup>[104, 365]</sup> In this study, both, the measured volume for

**Table 5.1.1.:** Unit cell volume  $V$  of different ZIFs taken from experimental SC XRD structures at low temperatures.

ZIF	$V / \text{\AA}^3$	Temperature / K	Ref.
ZIF-8	4784.9	100	[366]
ZIF-71	4756.8	297.15	[102]
ZIF-90	5152.2	100	[103]
ZIF-318	4981.5	180	[104]
ZIF-318p	5178.1*	-	-

\* Extrapolated from the volumes of ZIF-8 for full incorporation with MIM and ZIF-318 for half incorporation with CF3IM.

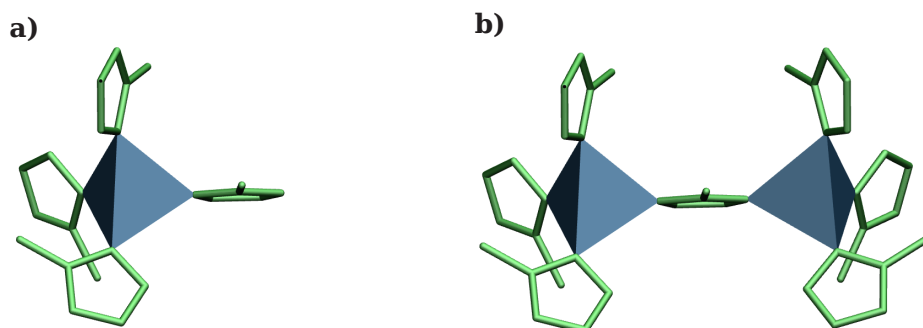
the partial linker incorporation as well as an extrapolated phase-pure product volume are employed. The latter is obtained from the two known data points of the phase-pure ZIF-8 volume and the experimental ZIF-318 volume, that is considered to be incorporated half with 2-trifluoromethyl-imidazolate (CF3IM) linkers. The examined ZIFs are composed of

different organic linkers, MIM, 4,5-dichloroimidazolate (DCIM), CF3IM and imidazolate-2-carboxyaldehyde (ICA) (Fig. 1.3.2). However, the fragmentation schemes and results are formulated with the general representative imidazolate fragment IM, that for the calculation for a specific ZIF refers to its respective linker.

Three fragmentation approaches are distinguished to ascertain the total polarizability volume that can be inserted in Eq. (2.4.30), of which the first is the BU approach reproduced from the formulation by Treger *et al.*<sup>[129]</sup>. It involves the heterolytic cutting of the coordinative metal-linker bonds yielding  $\text{Zn}^{2+}$  and the respective anionic IM. Since the unit cell of ZIFs in a sodalite topology is composed of 12 formula units, each consisting of one  $\text{Zn}^{2+}$  and two IM fragments<sup>[96, 97]</sup>, the unit cell polarizability volume can be obtained from the fragment's polarizability volumes as follows<sup>[129]</sup>

$$\alpha'^{\text{BU}} = 12(\alpha'_{\text{Zn}^{2+}} + 2\alpha'_{\text{IM}}). \quad (5.1.5)$$

Inserting the obtained unit cell polarizability volume  $\alpha'^{\text{BU}}$  into Eq. (5.1.4), yields the



**Figure 5.1.1.:** Representation of the **a)** TET und **b)** DI fragments with green IM linkers and blue tetrahedrons representing the coordination of  $\text{Zn}^{2+}$  ions.

respective refractive index.

Since this approach does not include any interactions between fragments and the TD approaches have shown beneficial properties and convergence<sup>[92, 93, 234]</sup>, while accounting for interactions in the bigger primary fragments, two novel fragmentation schemes for ZIFs are proposed that are denoted TD1 and TD2.

In the TD1 fragmentation scheme, a  $\text{Zn}^{2+}$  cation with its four tetrahedrally coordinated IM linkers is chosen as the primary fragment (Fig. 5.1.1 **a)**). Referring to the sodalite topology, the metal ions are assigned to nodes and the linkers are located on the edges of the structural representation (Fig. 1.3.1). These tetrahedral “TET” fragments overlap by a linker fragment (see Fig. 5.1.2 **a)**), yielding the following expression for the unit cell polarizability volume,

$$\alpha'^{\text{TD1}} = 12(\alpha'_{\text{TET}} - 2\alpha'_{\text{IM}}), \quad (5.1.6)$$

as a difference of the TET fragments and their overlapping IM fragments. Since the monomer energies are automatically included and beyond that partially intact metal-

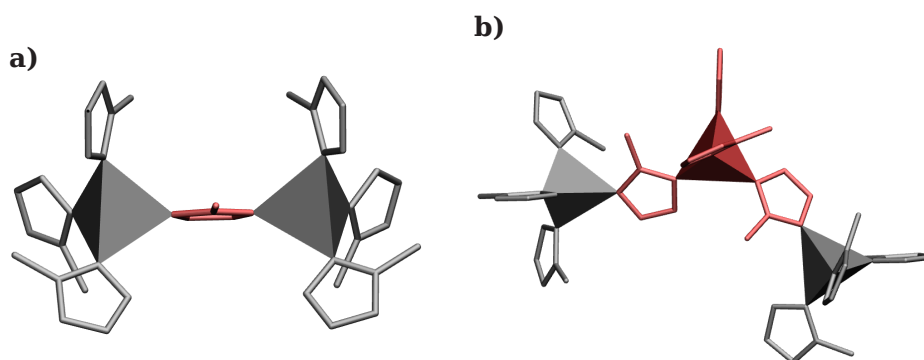


linker bonds contribute in this expression, this implies that this approach obtains superior results compared to the BU approach.

In the TD2 ansatz, this approach is advanced even further by choosing bigger primary fragments. The “DI” fragment consists of two  $\text{Zn}^{2+}$  ions and all their directly coordinated IM linkers (Fig. 5.1.1 **b**). The overlap region, therefore, consists of a full TET fragment (Fig. 5.1.2 **b**), resulting in the total polarizability volume as a difference between the polarizability volumes of these fragments,

$$\alpha'^{\text{TD2}} = 12(2\alpha'^{\text{DI}} - 3\alpha'^{\text{TET}}). \quad (5.1.7)$$

This approach involves the highest-order interactions and is therefore expected to yield the best results, however, is also expected to require the highest computational demand due to the fragments' size.



**Figure 5.1.2.:** Representation of the **a)** TD1 and **b)** TD2 fragmentation scheme. Tetrahedrons represent the coordinated  $\text{Zn}^{2+}$  ions. Red areas signify overlap regions of two fragments. The overlap is equal to one **a)** IM fragment or **b)** TET fragment.

## 5.1.2. Method Evaluation

In order to determine refractive indices for ZIFs, at first a computational protocol must be established. Initially, the polarizability volumes of the different introduced fragments (see Section 5.1.1) are calculated with different DFT approaches for the polarizability calculation (CP-SCF/TD-DFT), XC functionals and basis sets (see Section 3.2) on the example of ZIF-8. The calculations were performed for fragments directly extracted from the SC XRD structure or from a pDFT PBEsol-optimized structures.<sup>[129]</sup> Those performing best were then employed for further calculations of refractive indices comparing with the results from pDFT and experiment.<sup>[35, 129, 363]</sup>

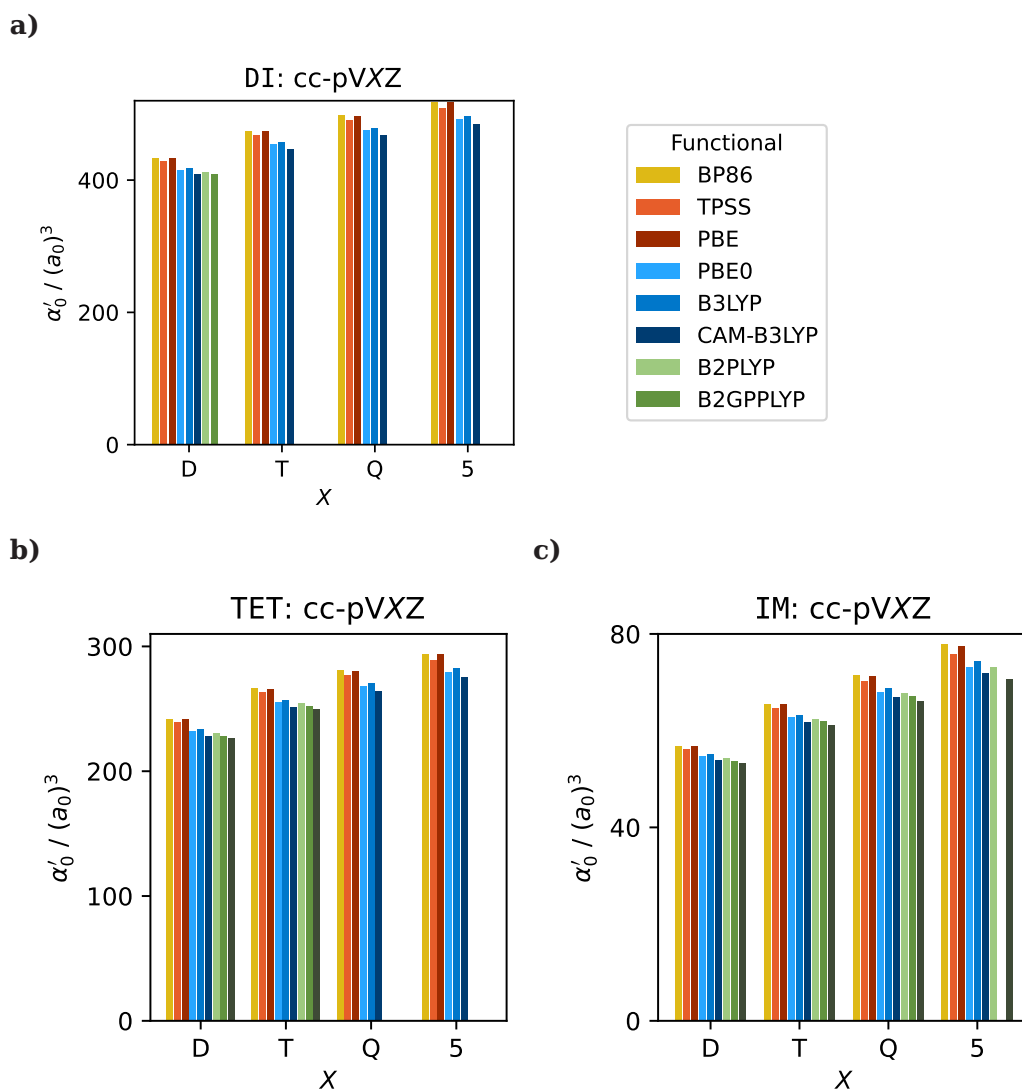
### 5.1.2.1. Exchange–Correlation Functional and Basis Set Evaluation

First calculations of the polarizability volume were performed in ORCA with the CP-SCF approach for the  $\text{Zn}^{2+}$ , IM, TET, and DI fragments. These were applied to SC XRD structures and structures optimized with the PBEsol XC functional in pDFT.<sup>[129, 366]</sup> In

the CP-SCF calculations a number of XC functionals (substantiated by previous studies) and basis sets were employed. For instance, in the study by Afzal *et al.*<sup>[253]</sup> BP86, TPSSh, B3LYP, B2PLYP, and PBE0 were applied for the calculation of polarizabilities, with PBE0 and B3LYP yielding the best results, outperforming the double-hybrid XC functional B2PLYP. Concluding from the functional benchmark of Hait *et al.*<sup>[367]</sup>, where static polarizabilities were calculated in a finite-field approach on small inorganic species, specifically double-hybrid functionals yielded excellent results. Even though in the past specifically GGA XC functionals were shown to incorrectly describe the decay of the XC potential leading to overestimated polarizabilities<sup>[368-370]</sup>, in the study from Hait *et al.* they partially gave good results.<sup>[367]</sup> Furthermore, a larger amount of Fock exchange can yield more accurate polarizabilities<sup>[368-370]</sup> so that hybrid functionals and several metaGGA functionals were found to give good results in this study. Especially range-separated functionals prove beneficial for polarizability calculations.<sup>[358, 368, 371-374]</sup> For these reasons and due to the availability and computational demand in ORCA, the set of examined XC functionals was chosen as follows: BP86, TPSS, PBE, PBE0, B3LYP, CAM-B3LYP, B2PLYP, B2GPPLYP, and DSD-PBEPBE-D3BJ. A large variation of basis sets was employed with Dunning (cc-pVXZ) and Ahlrichs-type basis sets (def2-X(Z)VP) from double- ( $X = D, S$ ) to quintuple- $\zeta$  ( $X = 5$ ) quality. The basis set availability was in some cases limited by the availability in ORCA. The results of this benchmark for the DI, TET, and IM fragments are shown in Fig. 5.1.3 and for the  $Zn^{2+}$  fragment in Fig. B.4 in the appendix. Further results are presented in the appendix in Figs. B.2 and B.3 for TET, Fig. B.1 for IM and Fig. B.4 for  $Zn^{2+}$ . Results for GGA-type functionals (BP86, TPSS, and PBE) are presented in yellow and red tones, for hybrid functionals (PBE0, B3LYP, and CAM-B3LYP) are shown in blue tones and double-hybrid functionals (B2PLYP, B2GPPLYP, and DSD-PBEPBE-D3BJ) are displayed in green tones.

The different XC functionals show the same trends for all basis sets for all fragments, DI, TET and IM: The GGA functionals seem to overestimate the polarizability as they show slightly higher values than all other XC functionals. This appears to be coherent due to the aforementioned lack of accuracy in describing the asymptotic decay of the XC potential in the GGA XC functionals.<sup>[368-370]</sup> Hybrid functionals give slightly lower values with the result from CAM-B3LYP being the lowest for every fragment and on a similar level as the double-hybrid functionals, specifically the DSD-PBEPBE-D3BJ functional, which was one of the best-performing functionals in the polarizability study by Hait *et al.*<sup>[367]</sup> This suggests, that CAM-B3LYP performs in a similar quality as the examined double-hybrid functionals, but with a smaller computational demand. For the DI fragment, the double-hybrid calculations proved difficult (numerical instabilities), so that no results could be obtained except for the cc-pVDZ basis set. For the IM and TET fragments, the results yielded with the double-hybrid functionals showed the lowest polarizability volumes for this basis set, with the CAM-B3LYP functional value closest to this. This implies, that CAM-B3LYP is the best-performing hybrid functional, however, deviations remain small.

† Onno Strolka, Leibniz University Hannover, 2022.



**Figure 5.1.3.:** Static polarizability volume  $\alpha'_0$  calculated with CP-SCF, different XC functionals and different basis sets in double-, triple-, quadruple- and quintuple- $\zeta$  quality for the **a)** DI **b)** TET and **c)** IM fragment extracted from the experimental ZIF-8 SC XRD structure that was additionally optimized with PBEsol (model **D1**).<sup>[129, 366]</sup> Results for GGA-type functionals (BP86, TPSS, and PBE) are presented in yellow and red tones, for hybrid functionals (PBE0, B3LYP and CAM-B3LYP) are shown in blue tones and double-hybrid functionals (B2PLYP, B2GPPLYP and DSD-PBEPBE-D3BJ) are displayed in green tones.<sup>‡</sup>

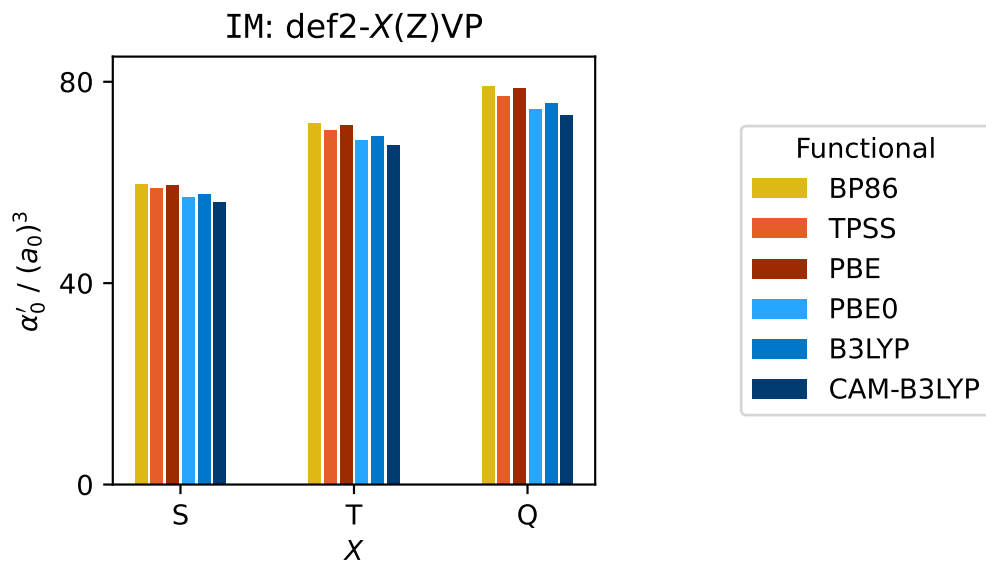
The results for the IM, TET and DI fragments give overall similar trends for the different basis sets. For instance, for DI the polarizability volume on average increases by  $79.6 a_0^3$  and  $40.3 a_0^3$  for all functionals from a cc-pVDZ to a cc-pV5Z basis set and from a cc-pVTZ to a cc-pV5Z basis set, respectively. Increasing the basis set size from double- to triple- $\zeta$  quality, thus, yield a majors step towards convergence. However, the further basis set extension, gives on average equal increases of the polarizability volume, suggesting that full convergence is not yet obtained with a cc-pV5Z basis set. Generally, when increasing the basis set size, the polarizability volume increases, however, not reaching a convergence in the range investigated. This is somewhat expected since it has been shown, that for accurate polarizabilities large basis set should be employed.<sup>[375-377]</sup> It suggests, that either bigger or different basis sets should be employed in the calculations.

Note that, the employment of different starting structures from SC XRD experiments as well as pDFT optimizations results in minor differences in the polarizability of around 5% for both the IM and TET fragment (Fig. B.1 for IM and Fig. B.2 for TET, in the appendix). The difference of the polarizability volumes for the different starting structures shows overall low fluctuations. The RMSD for all functionals and basis sets lies at  $0.27 a_0^3$  and  $0.69 a_0^3$  for the IM and TET fragment, respectively. This implies a low dependence of the calculation on the different examined structures for all XC functionals.

Additionally, a study with Ahlrichs basis sets (def2-X(Z)VP) was performed for the IM fragment with results displayed in Fig. 5.1.4 and Table 5.1.2. Further results for  $Zn^{2+}$ , IM and TET are shown in the appendix in Figs. B.1, B.2 and B.4. However, only basis sets up to quadruple- $\zeta$  quality were available. For the obtained results similar trends of the XC functional functionals and basis set increase could be obtained as for the Dunning basis sets. The Ahlrichs basis sets gave polarizability volumes that were in all cases higher than those obtained with the respective Dunning basis sets, *e.g.* with CAM-B3LYP for triple- and quadruple- $\zeta$  quality a difference of  $5.61 a_0^3$  and  $6.49 a_0^3$ , respectively, is observed. Although yielding higher polarizability volumes, the results obtained with the Ahlrichs basis sets do not show a convergence in the range up to the quadruple- $\zeta$  quality. This indicates, that basis sets with more basis functions are necessary and since the IM, TET and DI fragments are anionic, the addition of diffuse functions could be beneficial.

For the sake of employing the basis sets with diffuse functions as well as comparing the CP-SCF to a linear-response framework (yielding frequency-dependent rather than static polarizabilities), calculations were additionally performed with TURBOMOLE (see Table 5.1.2). As the hybrid XC functional CAM-B3LYP showed the closest agreement with double-hybrid values, while double-hybrid functionals did show numerical instabilities for the bigger basis sets and fragments, CAM-B3LYP was chosen for further calculations.

The polarizability volumes of the TURBOMOLE calculations with a response ansatz for the IM fragment are presented in Table 5.1.2 in comparison to those obtained with ORCA and the CP-SCF ansatz for different basis sets. The results for all other fragments are given in the appendix in Tables B.1 to B.3. It can be seen, that the results for IM obtained with TURBOMOLE and ORCA coincide in all cases. Differences can be found in the first or second decimal, which however could in part be caused by numerical accuracy deficiencies and are considered small deviations. Errors that occurred in the



**Figure 5.1.4.:** Static polarizability volume  $\alpha'_0$  calculated with CP-SCF, different XC functionals and different basis sets in double-, triple-, quadruple- $\zeta$  quality for the IM fragment extracted from the experimental ZIF-8 SC XRD structure (model A).<sup>[366]‡</sup>

**Table 5.1.2.:** Static polarizabilities in  $a_0^3$  calculated with ORCA and TURBOMOLE with CAM-B3LYP and different basis sets for IM.<sup>‡||</sup>

Basis set	ORCA	TM	Basis set	ORCA	TM
cc-pVDZ	53.78	-	def2-SVP	56.16	56.17
cc-pVTZ	61.75	61.75	def2-TZVP	67.36	67.31
cc-pVQZ	66.86	66.86	def2-QZVP	73.35	73.32
cc-pV5Z	71.92	71.92	-	-	-
aug-cc-pVDZ	77.57	-	def2-SVPD	71.97	71.97
aug-cc-pVTZ	78.40	78.36	def2-TZVPD	74.18	74.17
aug-cc-pVQZ	78.74	78.72	def2-QZVPD	76.05	76.04
aug-cc-pV5Z	-	78.91	-	-	-

ORCA calculations with an aug-cc-pV5Z basis set, did not occur in TURBOMOLE with a response calculation. For the other fragments, likewise minor to no differences can be found between the majority of the results for TURBOMOLE and ORCA (Appendix B.2.2). For instance, the average differences between the polarizability volumes obtained with the different programs are  $0 a_0^3$ ,  $0.1 a_0^3$  and  $0.01 a_0^3$  and below for the  $\text{Zn}^{2+}$ , TET and DI fragments, respectively, with the Dunning basis sets. Slightly higher differences can only be observed for the Ahlrichs basis sets without diffuse functions.

<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.

Moreover, it becomes evident that the polarizability volume obtained with the biggest Dunning basis set without diffuse functions (cc-pV5Z) gives as good results as the smallest Dunning basis set including diffuse functions (aug-cc-pVDZ). In previous studies, slow basis set convergence for Dunning basis sets and somewhat better convergence when including diffuse functions has been observed.<sup>[378, 379]</sup> This can be substantiated by the fact that augmented basis sets include a much larger number of basis functions. The aug-cc-pV5Z calculations show only a minor increase of the polarizability volume in comparison to the result for the aug-cc-pVQZ (an increase of  $0.19 a_0^3$ ) suggesting that either convergence or close to convergence of the polarizability volume is reached. This implies, that the calculation of polarizability volumes with an aug-cc-pVTZ is sufficient to obtain results close to convergence. This performance is much better than for the Ahlrichs basis sets. When increasing the basis set size for the def2-*X*(Z)VPD basis sets, the increase of the polarizability volume from double- to triple- and from triple- to quadruple- $\zeta$  quality lies at  $2.21 a_0^3$  and  $1.87 a_0^3$ , respectively. This implies that no convergence is reached in this range and bigger basis sets would be necessary for the Ahlrichs basis sets.

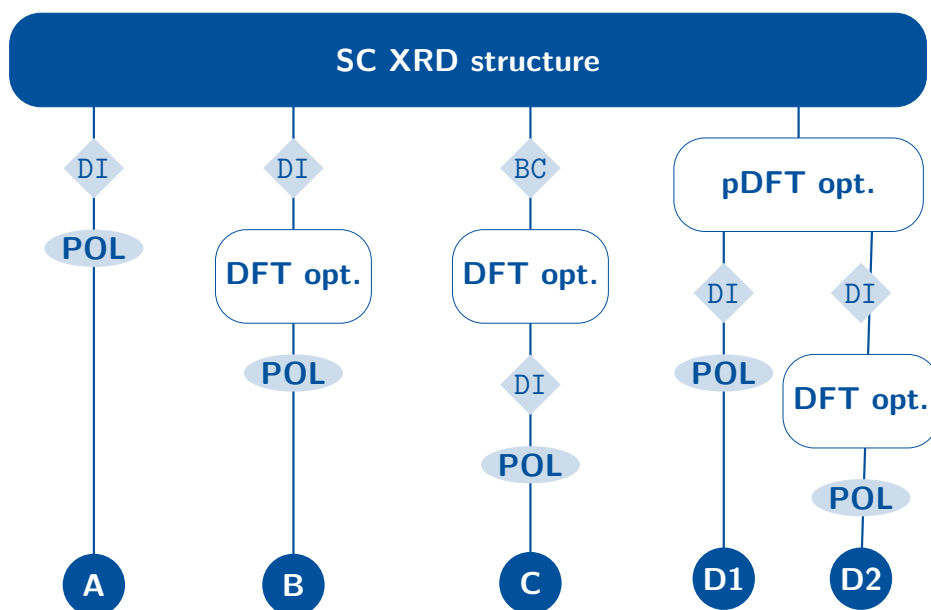
Since the IM, TET and DI fragments are anions, better performance of basis sets with diffuse functions is somewhat expected.<sup>[380]</sup> This is the case for all fragments examined, deeming these basis sets to be the most accurate for the calculation of polarizability volumes. Moreover, in all cases, the augmented Dunning basis sets lead to polarizability volumes that are equal to or higher than those of cc-pV5Z quality and show values close to convergence already at triple- $\zeta$  quality. Therefore, and because of the favourable scaling for bigger fragments such as DI, in further calculation of the dynamic polarizability an aug-cc-pVTZ basis set is employed.

In summary, the influence of the XC functional was proven small compared to basis set effects as long as functionals of hybrid or double-hybrid quality were employed. Additionally, the calculation with an aug-cc-pVTZ basis set yielded dynamic polarizabilities close to convergence. Therefore, the CAM-B3LYP/aug-cc-pVTZ protocol was determined as the most favourable option for further calculations.

### 5.1.3. Refractive Indices for ZIF-8

With CAM-B3LYP/aug-cc-pVTZ emerging as the most favourable calculation scheme for polarizability volumes, the refractive indices for ZIF-8 were calculated for the BU, TD1 and TD2 fragmentation schemes and on the basis of differently geometry-optimized structures (Fig. 5.1.5). Ensuing from the geometries taken from the SC XRD experiments, the following optimization schemes were pursued. For scheme **A** all fragments for the polarizability calculations were directly extracted from the experimental structure. For scheme **B** the extracted DI fragment was further optimized with either a PBE XC functional or a B97-3c method and all other fragments extracted from this optimized geometry for polarizability calculations. In a more extensive optimization (scheme **C**) the full  $\beta$ -cage was extracted from the experimental structure and fully optimized with B97-3c/def2-mTZVP. Then further fragments were extracted from this geometry for the polarizability calculation. Scheme **D** involves the most extensive optimization with

pDFT and a PBEsol XC functional.<sup>[129]</sup> The obtained geometries were then either used for fragment extraction and polarizability calculations (**D1**) or a further optimization of an extracted DI with B-97-3c/def2-mTZVP was performed before the polarizability calculation of the obtained fragments (scheme **D2**).

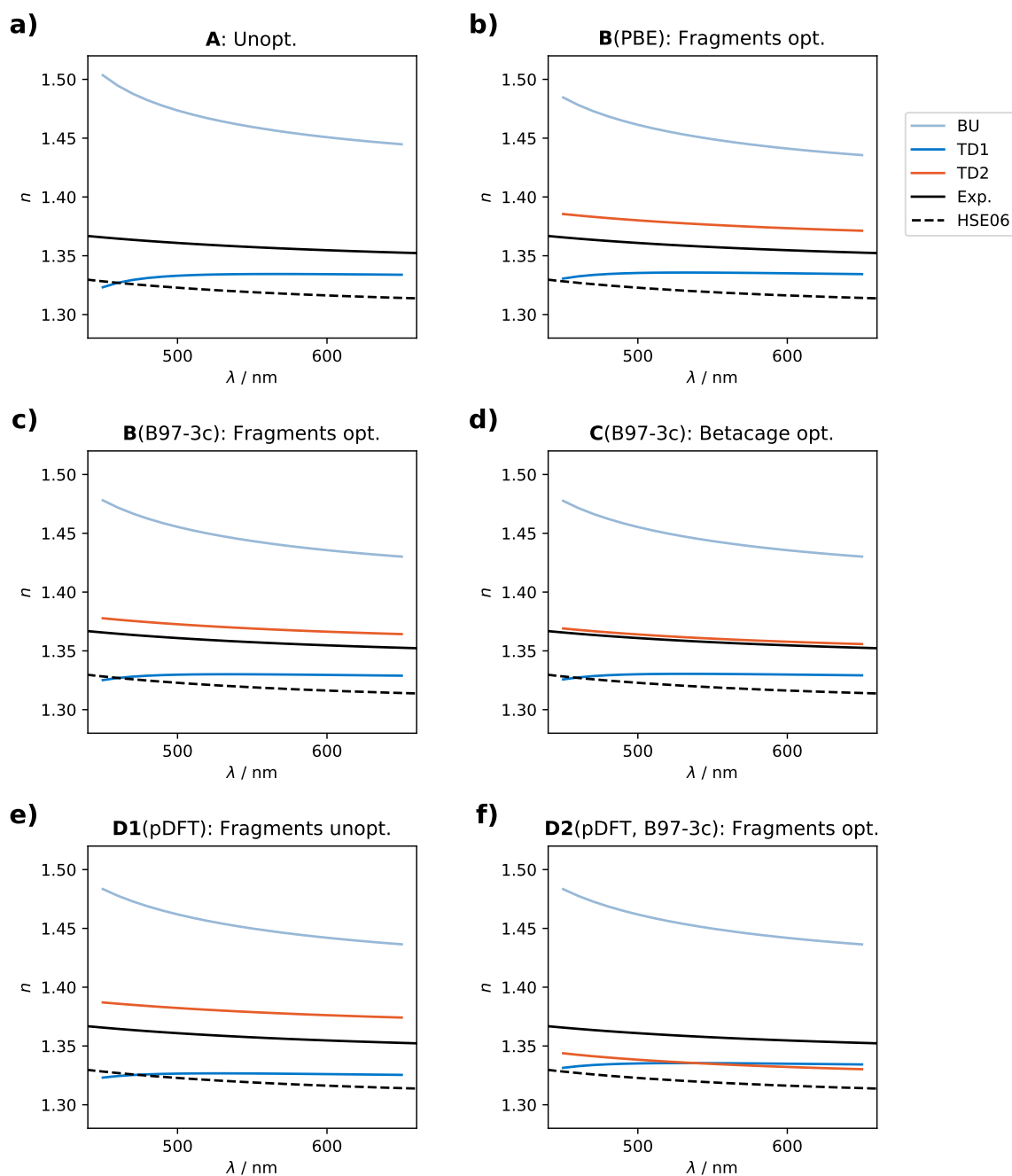


**Figure 5.1.5.:** Schematic representation of employed geometries and their origin for the calculation of polarizabilities and refractive indices. The dark blue rectangular box represents the original starting structure. White rectangular boxes refer to performed optimization processes, diamond shapes refer to the extraction of substructures and ellipses refer to polarizability calculations (POL). Optimization models (**A**, **B**, **C**, **D1** and **D2**) are defined in blue circular nodes.

The results for these schemes in comparison to full pDFT calculations with HSE06<sup>[129]</sup> and experimental data<sup>[35]</sup> are presented in Fig. 5.1.6.

A few general statements can be made for the overall results. Firstly, the BU fragmentation scheme does in every case strongly overestimate the refractive index in comparison to all reference data. For instance, the difference between the refractive index from the BU ansatz and from the experiment for Fig. 5.1.6 **a**) (model **A**) and Fig. 5.1.6 **e**) (model **D1**) on average lies at 0.11 and 0.10, respectively. For model **A**, for lower wavelengths the difference between the refractive indices for BU and experiment is slightly higher (on average 0.13 at 450 nm) than for higher wavelengths (on average 0.09 at 650 nm). For the polarizability volume from the BU approach low differences between the optimization schemes are consistent with the fact, that for the BU ansatz only changes in the IM fragment contribute. This is substantiated by the fact that changes to the structure in the different optimizations are marginal for all optimization models.

The results from the TD1 and TD2 schemes are in all cases closer to the reference data than for the BU approach, specifically the TD1 results being closest to the HSE06 reference and the TD2 approach being closest to the experimental reference in most



**Figure 5.1.6.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure (model **B**) for **a**)–**d**) and additional optimization with PBEsol (model **D1**) for **e**)–**f**).<sup>[129, 366]</sup> Different fragmentation schemes are compared to the experimental and pDFT data.<sup>[35, 129]</sup> The refractive indices for TD2 could not be obtained for **a**) (model **A**) due to non-real instabilities in the SCF.<sup>||</sup>

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.



cases [a), c), d) and e)]. It should also be noted that solely the TD1 results give an incorrect curve progression for small wavelengths for all optimization models. The reason for this probably originates in a difficult description of the isolated IM fragment. This also in part explains the overestimation of refractive indices in the BU approach and the wrong frequency dependence in the TD1 approach. Since the linkers show resonance structures, a partial multireference character can be found, and in a test calculation this was estimated to be of low to medium influence. Nevertheless, it can be speculated to be large enough to overestimate the polarizability specifically for small wavelengths and therefore change the curve progression in the TD1 approach.

Focussing on the data obtained with the **D1** and **D2** schemes (pDFT-optimized structures, Fig. 5.1.6 e)-f)), it can be seen, that major differences occur in the refractive indices when comparing the results for **D1** without further optimization and **D2** with further optimized DI fragments. Leaving out the previously discussed results for the BU scheme here, the results for the TD1 and TD2 schemes are much closer to reference data than the BU results. For the TD1 scheme, the refractive indices for the **D1** model are closest to those for the HSE06 reference, which is coherent with both calculations being obtained from the same optimized structure. For the **D2** model higher deviations of up to 0.02 at 650 nm occur. The refractive indices for the TD2 range between those for both reference datasets for the **D2** model (difference to exp. around -0.02) and above the experimental reference (difference on average 0.02) for **D1**. However, the general curve progression is in good agreement with experiment in both cases for the TD2 scheme.

For the HSE06 data, a very small deviation in comparison to experiment can be observed for lower wavelengths showing a slightly steeper decay of the refractive index. Overall the HSE06 data was expected to underestimate the refractive index<sup>[129]</sup>, however, the experimental data could possibly also show slight effects of overestimating the refractive index due to remaining guest molecules in the porous structure, that are not accounted for in the fragmentation calculations. Nonetheless, the TD2 approach gives the closest results to experiment for these structures and is expected to be closest to the actual refractive index.

Fig. 5.1.6 additionally shows the results for the schemes **A**, **B** and **C**. The refractive indices for the TD1 fragmentation scheme are in all optimization models overall very close to those obtained with the **D1** scheme, showing refractive indices close to the HSE06 reference with differences of up to 0.02 at 650 nm. In contrast, the refractive indices for the TD2 fragmentation scheme are overall closest to the experimental data implying a better performance of this scheme for model **A**, **B** and **C**. However, the refractive indices for TD2 could not be obtained for model **A** due to non-real instabilities in the SCF even when extensively tightening convergence criteria and increasing the quadrature grid size. A further problem analysis is described in Section 5.1.3.1. For model **B** it is found, that the optimization with B97-3c gave results closer to the experimental reference deviating by 0.19 and 0.12 at 650 nm with TD2 for the PBE and B97-3c optimized structure, respectively. This implies that the optimization with B97-3c leads to refractive indices in better agreement with experiment for the TD2 fragmentation scheme. This is substantiated by the results for model **C** with a TD2 fragmentation scheme, that

are in excellent agreement with the experimental reference with a slight deviation to experiment of 0.003 at both 450 and 650 nm.

Overall the refractive indices for TD2 are closer to the experimental reference in the **A**, **B** and **C** optimization schemes and for all wavelengths. In further calculations, mainly the **B**(B97-3c) and **C** optimization schemes are suggested in order to obtain the best agreement with the experimental reference. Moreover, the TD2 fragmentation involving the highest-order of interaction terms is recommended as long as a stable SCF minimum can be obtained. This also accounts for the fact that the individual IM fragment is not part of this scheme, avoiding the observed overestimation of the polarizability volume for small wavelengths.

### 5.1.3.1. Electrostatic Embedding

In the calculation of the DI fragment, SCF instabilities occurred and several approaches, tightening the convergence criteria to find a true minimum, were performed unsuccessfully. Therefore, electrostatic embedding (*cf.* Section 2.4 and Chapter 4) of the full structure was considered *via* a continuum solvation model, since electrostatically embedded fragmentation approaches have been employed successfully in the past.<sup>[82, 228, 236, 381–384]</sup> Therefore, the Clausius–Mossotti equation relating the dielectric permittivity to the polarizability,

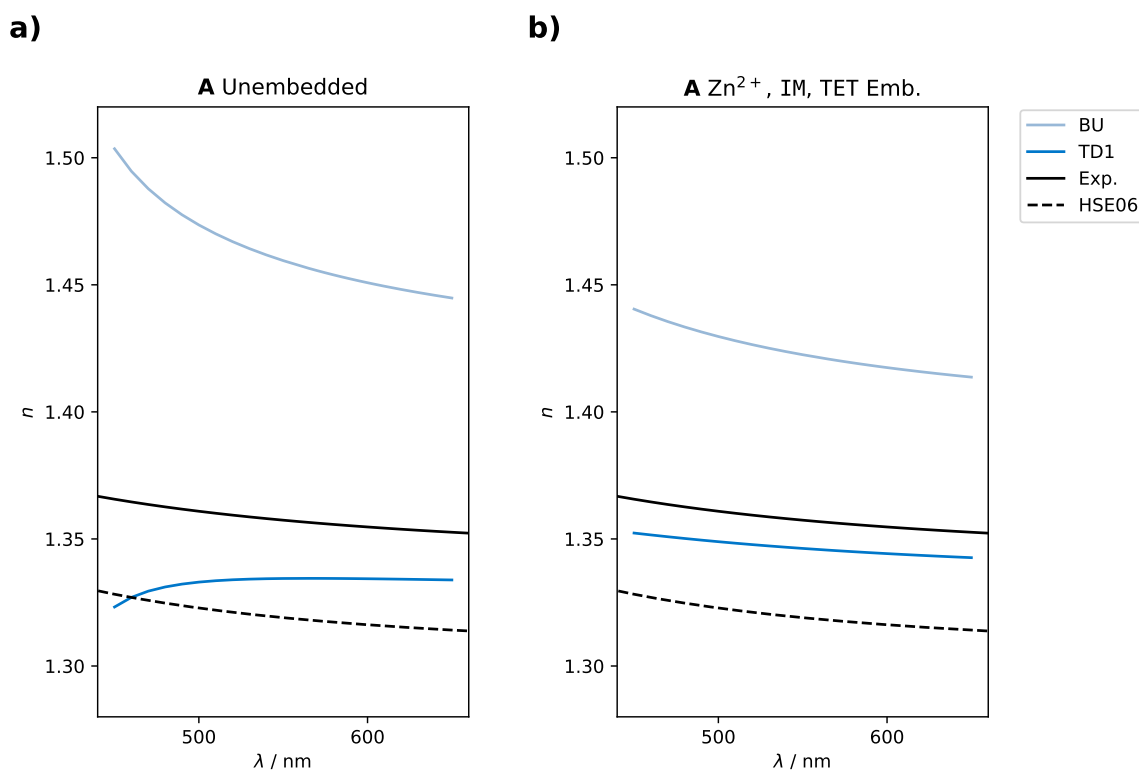
$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{N\alpha}{3\epsilon_0}, \quad (5.1.8)$$

was employed. In order to determine an average permittivity for the ZIF surrounding the fragment, it was solved for  $\epsilon_r$  making use of the reciprocal unit cell volume for the number density and the definition for the polarizability volume, giving

$$\epsilon_r = \frac{\frac{8\alpha'\pi}{3V} + 1}{1 - \frac{4\alpha'\pi}{3V}}. \quad (5.1.9)$$

The polarizability volume was then taken from the polarizability calculation of the TD1 approach for the same model (**A**, see Fig. 5.1.6) yielding value of  $\epsilon_r = 2.263 e^2/a_0E_h$  when averaged over all wavelengths, while the values for the highest and lowest wavelength only deviated by  $0.02 e^2/a_0E_h$ . The calculations for all fragments were then repeated embedded in the continuum with the obtained average permittivity. In the subsequent TD-DFT calculation, no embedding was employed for the orbitals in the perturbation.<sup>[385, 386]</sup> Unfortunately, stable SCF results were obtained for all fragments except DI, therefore no results for the TD2 scheme are displayed in Fig. 5.1.7. The results in comparison to those obtained without electrostatic embedding are displayed in Fig. 5.1.7. Modified fragmentation schemes employing embedded fragments are denoted with the prefix “m”, *i.e.* mBU and mTD1.

Although not obtaining results for TD2, the stabilization of the fragments in the continuum leads to the improvement of the curve progression at small wavelengths for both



**Figure 5.1.7.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure (model **A**).<sup>[366]</sup> Results for model **A** are obtained **a)** without and **b)** with electrostatic embedding for Zn<sup>2+</sup>, IM and TET with a relative permittivity continuum with  $\epsilon_r = 2.263 e^2/a_0 E_h$ . The refractive indices for TD2 could not be obtained due to non-real instabilities in the SCF for **a)** and **b)**.

BU and TD1. For BU a strong decrease of the refractive indices due to the polarizability volume decrease for the IM fragment are the consequence, leading to polarizability volume differences of  $10.3 a_0^3$  and  $5.26 a_0^3$  at 450 nm and 650 nm, respectively. For the TET fragment the difference is  $10.2 a_0^3$  and  $7.4 a_0^3$  at 450 nm and 650 nm. The curve for the mTD1 approach did show a shift to lower refractive indices. These results substantiate the conclusion drawn before deeming the electronic structure description of the IM fragment in vacuum partially insufficient. In the following calculations, the results for embedded  $\text{Zn}^{2+}$ , IM and TET fragments are therefore additionally evaluated.

#### 5.1.4. Refractive Indices for Sodalite-Type Zeolitic Imidazolate Frameworks

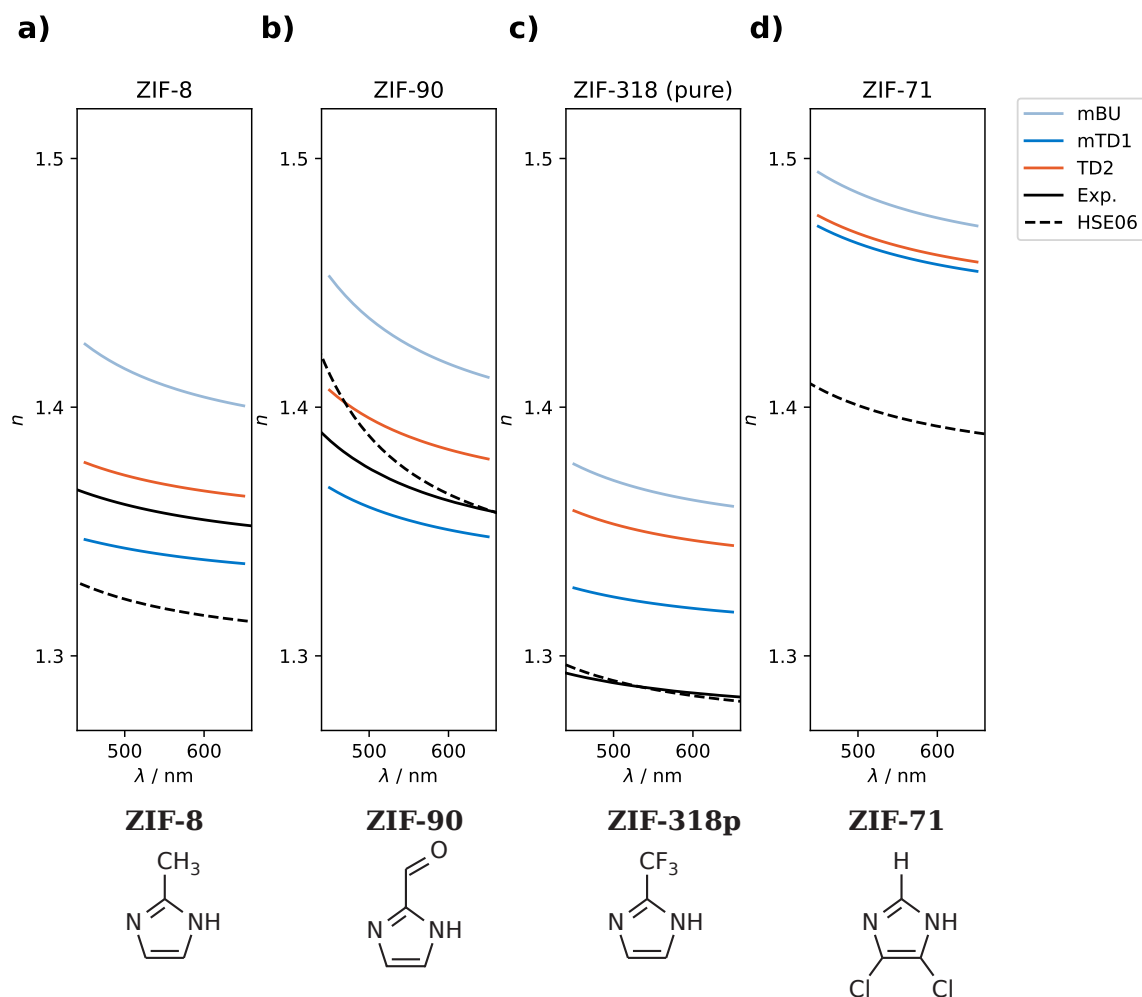
In this section, the refractive indices for ZIF-8, ZIF-90, ZIF318p and ZIF71 are compared to reference data.<sup>[35, 129, 363, 364]</sup> ZIF-318 presents phase-pure in the thin film experiments. Therefore, the refractive indices for ZIF-318 are calculated with the extrapolated phase-pure volume and phase-pure linkers, denoted ZIF-318p. All fragmentation schemes were calculated for the optimization models **A**, **B** and **C**. Since models **B** and **C** previously performed best, the results are presented in this section. For model **B**  $\text{Zn}^{2+}$ , IM and TET fragments solvated in a continuum were employed (Fig. 5.1.8). Results obtained for model **C** and embedded  $\text{Zn}^{2+}$ , IM and TET fragments are displayed in Fig. 5.1.9. Additional results for model **A** and **B** are presented in the appendix in Figs. B.5 and B.6.

Focussing on the comparison of the refractive indices of ZIFs obtained with model **B** and the  $\text{Zn}^{2+}$ , IM and TET fragments solvated in a continuum with the permittivity calculated from the TD1 results, the obtained refractive indices are displayed in Fig. 5.1.8.

Comparing the refractive indices for the embedded fragments (Fig. 5.1.8) and unembedded fragments (Fig. B.6 in the appendix), the results strongly improved for all ZIFs with the BU ansatz in comparison to reference data. However, overall the BU approach gives results strongly deviating from the reference data in all cases. The results for TD1 for all ZIFs are shifted towards higher refractive indices when solvating the fragments, decreasing the deviation to the curves for the TD2 scheme.

For ZIF-8 (Fig. 5.1.8 **a**)), the results for a TD1 fragmentation scheme are in good agreement to those obtained without further optimization. The curve progression is improving for TD1, however, in worse agreement than for the TD2 scheme.

For ZIF-90 (Fig. 5.1.8 **b**)), the refractive indices for a TD1 scheme are underestimated with a difference to experiment of 0.02 and 0.01 at 450 nm and 650 nm, respectively. The results are, however, improved in contrast to the calculations for unembedded fragments. The curve progression is less steep than for the experiment at small wavelengths. The pDFT reference for ZIF-90 is coinciding with reference data for the higher wavelengths, however, it shows an even steeper curve progression than the TD1 results for small wavelengths. For 450 nm the difference to experiment is 0.03, for 650 nm it is  $8 \cdot 10^{-5}$ . In contrast, the results of TD2 for ZIF-90 show good agreement of the curve progression with experiment, but overestimate the refractive index by 0.02 on average.



**Figure 5.1.8.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure and optimized with B97-3c/def2-mTZVP (model **B**).<sup>[366]</sup> Results are obtained for **a)** ZIF-8, **b)** ZIF-90, **c)** ZIF-318 with an extrapolated volume and **d)** ZIF-71.  $\text{Zn}^{2+}$ , IM and TET fragments were electrostatically embedded in a relative permittivity continuum with **a)**  $\epsilon_r = 2.263 e^2/a_0E_h$ , **b)**  $\epsilon_r = 1.825 e^2/a_0E_h$ , **c)**  $\epsilon_r = 1.787 e^2/a_0E_h$ , **d)**  $\epsilon_r = 2.119 e^2/a_0E_h$ . The results are compared to experimental and computational reference.<sup>[35, 129, 363]</sup> Respective linkers are shown at the bottom.<sup>||\*</sup>

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.

<sup>\*</sup> Erona Shabani, Leibniz University Hannover, 2023.

**Table 5.1.3.:** Unit cell volume  $V$  of different ZIFs obtained from pDFT (PBEsol) and SC XRD structures at low temperatures (exp.) in  $\text{\AA}^3$ .

MOF	$V_{\text{exp.}} / \text{\AA}^3$	$V_{\text{pDFT}} / \text{\AA}^3$
ZIF-8	4784.9 <sup>[366]</sup>	4781.7 <sup>[129]</sup>
ZIF-90	5152.2 <sup>[103]</sup>	5163.9 <sup>[129]</sup>
ZIF-71	4756.8 <sup>[102]</sup>	4765.2 <sup>[129]</sup>
ZIF-318	5178.1 <sup>*</sup>	5168.1 <sup>[129]</sup>

\* Extrapolated from the experimental volumes of ZIF-8 for full incorporation with MIM and ZIF-318<sup>[104]</sup> for half incorporation with CF3IM.

The refractive index for TD2 at 450 nm lies at 1.407, which is 0.006 lower than for the pDFT reference. The TD2 refractive indices therefore give closest agreement in comparison to the experiment for the curve progression, while showing a constant overestimation for the refractive index of ZIF-90. Aiming at higher wavelengths, the pDFT reference performs best. Furthermore, in comparison to ZIF-8 a slightly higher polarizability of the linker, that is dominating the optical properties<sup>[35, 127, 128, 387, 388]</sup>, is anticipated, leading to an increase of refractive indices, which can be reproduced in all calculations except for a BU scheme with unsolvated fragments.

In contrast, in ZIF-318 (Fig. 5.1.8 c)), the linker is substituted with a fluorinated methyl group and expected to be less polarizable, leading to an overall lower refractive index than for ZIF-8. This trend is reproduced for all calculations except for the unsolvated TD1 fragments for small wavelengths (Figs. B.6 and B.7 in the appendix). The refractive indices were calculated for the phase-pure ZIF-318 concerning the involved linkers and extrapolated volume used in the Lorenz-Lorentz equation. Additionally, the same calculation was performed employing the SC XRD volume (Table 5.1.1). Results from the SC XRD volume slightly shifted the curves to higher refractive indices than for ZIF-8 and are displayed in the appendix in Fig. B.7. These deviate further from the experimental refractive indices, that are obtained for phase-pure ZIF-318 films and are, therefore, omitted in the further discussion.

The calculations with embedded fragments lead to different curves than the unsolvated fragments, decreasing the deviation of the BU curve and the experimental curve by 0.01 on average. The TD1 refractive indices change minorly upon solvation (mTD1), however, the mTD1 results slightly deviate more from experiment than those obtained with the TD1 approach (deviation increases on average by 0.003).

The pDFT reference shows very low refractive indices, that almost coincide with the experimental data but deviate largely from the results for the fragmentation schemes. Solely the curve progression is deviating from experiment but with deviations of only  $1.81 \cdot 10^{-6}$  on average. The difference between TD2 and pDFT at 450 nm and 650 nm is 0.062 and 0.063, respectively. The curve progression in the results from the TD1 approach gives the best agreement with experiment concerning solely the curve progression. Moreover, the deviation of the TD1 and TD2 refractive indices from the experimental data is on average 0.034 and 0.063, respectively. Thus, in this case, the results for TD1

deviate less from experiment than for the TD2 approach. The deviation between TD2 and mBU (0.017) is on average lower than for TD2 and TD1 (0.029). However, it should be noted, that this could be improved by an additional solvation of the fragments in the TD2 approach.

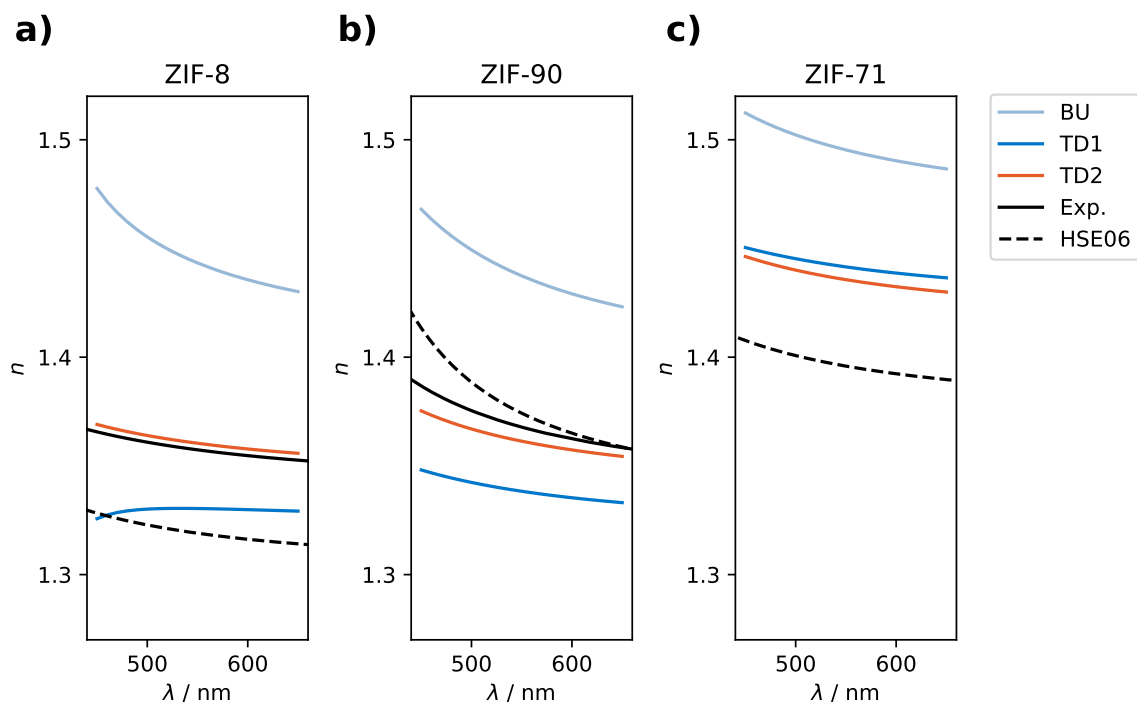
The pDFT calculations were performed with PBEsol optimized structures and thus with an adapted unit cell volume. It is possible, that the performed extrapolation to determine the volume for the phase-pure ZIF-318 unit cell in the fragmentation-based calculations is insufficient in this case. Incorporating an increased volume into the Lorenz-Lorentz equation could result in downward shifts within the refractive index curves (*cf.* Fig. B.7 in the appendix). In Table 5.1.3, the unit cell volumes from experiment and pDFT are listed. It is evident, that the volumes from a pDFT optimization are very close to the SC XRD volumes for ZIF-8, ZIF-71 and ZIF-90. Therefore, the extrapolated volume, that is in good agreement with the volume obtained in pDFT is expected to be close to the real-phase pure volume. However, this assertion remains subject to experimental verification. Given this assumption, the fragmentation approaches appear to exhibit an apparent overestimation of the refractive indices concerning this particular ZIF. The origin of this factor remains a subject for subsequent investigation, which could include the use of a different optimization scheme or examining embedding not only for the fragments in the polarizability calculation but also within the different optimization schemes, as an example.

For ZIF-71 (Fig. 5.1.8 **d**)), that incorporates a 4,5-dichloroimidazolate linker, much higher refractive indices than for ZIF-8 are expected. This effect is reproduced by all fragmentation calculations. For the pDFT reference the values are higher than for ZIF-8, but not than all ZIF-90 values where the curve is much steeper. For ZIF-71 the curve progression only deviates slightly for all calculations. Unfortunately, no experimental reference is available for comparison. The results for the mTD1 and TD2 approaches are very close (average deviation 0.004) and closer than for the unembedded TD1 and TD2 approaches (average deviation 0.019), *cf.* Fig. 5.1.8. The result for the mBU approach is closer to the curve for TD2 than that for the BU scheme as well, where the average deviation is 0.016 (average deviation of BU and TD2: 0.030) For this ZIF, the results for the fragmentation approaches are very close when involving embedding, but deviate strongly from pDFT (deviation of 0.069 on average).

In Fig. 5.1.9 the results for ZIF-8, ZIF-90 and ZIF-71 calculated on the structure from model **C** ( $\beta$ -cage optimized) are compared for the different fragmentation schemes without embedded fragments. For ZIF-8 (Fig. 5.1.9 **a**)) the results were previously discussed. Specifically, the TD1 curve is in very good agreement with TD2 and experimental data.

For ZIF-90 (Fig. 5.1.9 **b**)) the optimization of the  $\beta$ -cage improved the results of the BU and TD2 ansatz in comparison to the experimental reference, however, not for the TD1 ansatz. The refractive indices for a TD2 ansatz shifted strongly and for this optimization model are overall closest to the experimental reference showing the highest deviation at small wavelengths (0.01 at 450 nm) and lowest for high wavelengths (0.004 at 650 nm) and a curve progression that is in good agreement with the experiment.

For ZIF-71 (Fig. 5.1.9 **c**)), the results for BU slightly shift to higher refractive indices from model **B** to **C** and for TD1 only slightly shifts to higher values, whereas for TD2 the



**Figure 5.1.9.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure and optimized  $\beta$ -cage with B97-3c/def2-mTZVP (model **C**). Results are obtained for **a)** ZIF-8, **b)** ZIF-90, **c)** ZIF-71.<sup>[366]</sup> Results are obtained for **a)** ZIF-8, **b)** ZIF-90, **c)** ZIF-71. The results are compared to experimental and computational reference.<sup>[35, 129, 363]</sup>\*

\* Erona Shabani, Leibniz University Hannover, 2023.



curve shifted to lower refractive indices (-0.027) showing a curve progression in good agreement with the pDFT reference. The shifting leads to the TD1 and TD2 curves to be extremely close with a deviation of 0.027 and 0.021 for 450 nm and 650 nm, respectively.

In summary, the results for the optimized  $\beta$ -cage in combination with a TD2 approach yielded results in excellent agreement with experimental reference data. BU and TD1 approaches in most cases were found insufficient and specifically BU yielded major shifts of the refractive indices even when modified with a solvation model. For mTD1 some improvements mainly for the curve progressions could be observed. The trends expected due to the polarizability of substituents could be fully reproduced in TD2 results with structures from model **A** and **B** and **C**. The refractive index spanned a range of 1.354 and 1.446 for the different ZIFs in the TD2 scheme for model **C**. The results for model **B** and **C** were in excellent agreement with experimental data. In this case the refractive index is predicted to be in close agreement with the TD2 result for ZIF-71. A recommendation for the use of continuum solvation models can be given if values for  $\epsilon_r$  are available.

## 5.2. Loading of Zeolitic Imidazolate Frameworks in Fragmentation Schemes

As shown in Chapter 1, ZIFs with a sodalite topology exhibit a pore in every  $\beta$ -cage with a volume of approximately  $2500 \text{ \AA}^3$ .<sup>[107]</sup> These large pores are connected by so-called pore windows, six 4MR and eight 6MR, marked in the sodalite-type unit cell in Fig. 1.3.1 **c)** and **d)**, respectively. While the former present with a Zn-Zn diagonal of approximately  $8.5 \text{ \AA}$ , the 6MR shows a Zn-Zn diameter of circa  $12.2 \text{ \AA}$ .

In Sections 5.1.3 and 5.1.4 energy-based fragmentation methods were successfully applied in order to calculate the refractive index of unloaded ZIFs. Therefore, it is feasible to adapt these methods to account for the loading of these MOFs and possibly gain insight on the number of introduced molecules and favoured positions in the pore. The adaptation of the fragmentation ansatz accounting for guest molecules is presented in Section 5.2.1. Refractive indices are calculated with these novel schemes on the example of ZIF-8 and ZIF-71 (Section 5.2.2) and compared to reference data.<sup>[35, 129, 364]</sup>

### 5.2.1. Fragmentation Schemes

Building on the theoretical basis set in Section 5.1.1, the total polarizability of the loaded ZIF can be formulated as,

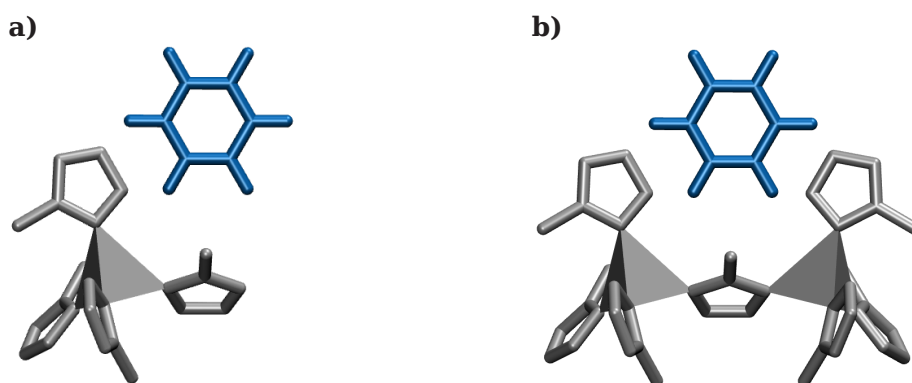
$$\alpha'^{\text{ZIF}} = \alpha'^{\text{ZIF,empty}} + \Delta^{\text{L}}, \quad (5.2.1)$$

with  $\Delta^{\text{L}}$  being the terms arising due to the introduction of the guest molecule. These are added to the terms of the empty ZIF as mentioned above. This means, that in the process of loading the ZIF, the unit cell volume is considered constant and solely the polarizability of the guest itself and the interaction with that guest modifies the total polarizability.

Generally, different approaches for representing the empty ZIFs can be employed in this equation but will be limited to TD2 schemes here due to their performance for empty ZIFs. According to the fragmentation schemes introduced before,  $\Delta^L$  can be defined in a different orders. The zeroth order scheme is presented as a BU approach, where only the polarizability of the guest molecule but no guest-host interactions are considered. For the example of a single benzene molecule (B) per pore as the guest, the total polarizability of the loaded ZIF becomes

$$\Delta^{L0} = \alpha'^{L,BU} = 2\alpha'^{\text{guest}}. \quad (5.2.2)$$

This approach is denoted as L0 and takes into account, that the guest is in close vicinity to more than one window in the pore. In this approach, multiple fragments can easily



**Figure 5.2.1.:** Representation of the **a)** TETB and **b)** DIB fragments. Zn<sup>2+</sup> ions are represented as tetrahedrons to highlight the coordination. The benzene guest molecule is depicted in blue.

be incorporated by simple addition.

Introducing the TETB fragment (Fig. 5.2.1), involving both TET and B, the interactions between the guest and every TET fragment can be considered as follows,

$$\Delta^{L1} = 2\alpha'^B + 12 \cdot (\alpha'^{\text{TETB}} - \alpha'^{\text{TET}} - \alpha'^B). \quad (5.2.3)$$

This is the first-order approach, denoted L1, which is scaled to account for the full cell. This scheme requires the symmetry of the guest molecule so that no additional terms occur due to different interactions with respective fragments in the ZIF. The pore size available for incorporation of guests is experimentally approximated at 11.6 Å in diameter and the benzene molecule spans approximately 4.9 Å in diameter. Thereby, incorporating the guest molecule into the pore, at minimum minor interactions are anticipated for all fragments in the pore. Additionally, fragments are part of more than one pore and fragments that are not in close vicinity to the guest molecule in one pore are closer to that in a neighbouring pore. Increasing the number of benzene molecules in this approach could mean, that *e.g.* two benzene molecules are handled as a single fragment. Therefore, the TETB and B fragments would be replaced by TETBB and BB fragments, respectively, where the TETBB fragment is a TETB fragment with the additional B guest and the BB fragment is the two incorporated B guests combined as one

fragment. Alternatively, more than one TETB fragment could be introduced, interacting with either of the two guest molecules. This would require additional consideration of the B-B interactions either *via* the introduction of a new BB fragment or (additionally) a TETBB fragment.

In the second-order scheme, additional interaction terms of the guest molecule with the DI fragment are introduced. The DIB fragment (Fig. 5.2.1 **b**) includes these, however, also introduces double-counting of interactions obtained with the TETB fragment. Therefore, subtracting the double-counted terms, the total loading polarizability, denoted L2, can be obtained with,

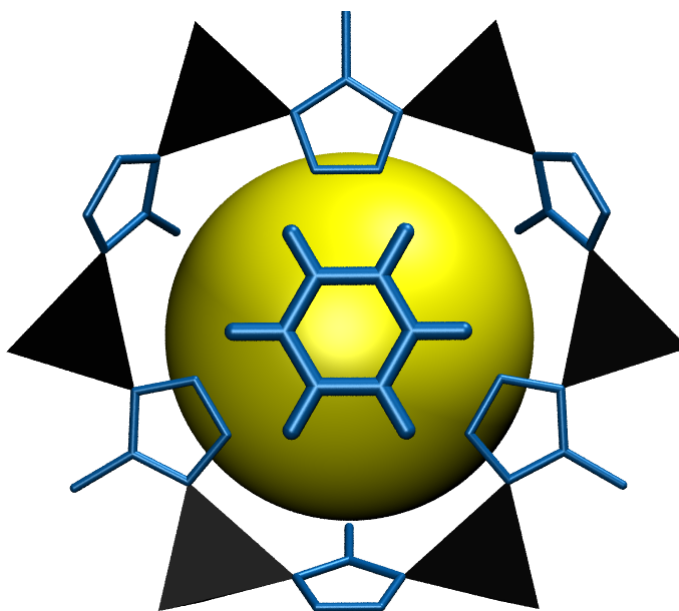
$$\begin{aligned}\Delta^{L2} &= 2\alpha'^B + 12 \cdot (\alpha'^{TETB} - \alpha'^{TET} - \alpha'^B) \\ &\quad + 12 \cdot (\alpha'^{DIB} - \alpha'^{DI} - \alpha'^B) - 24 \cdot (\alpha'^{TETB} - \alpha'^{TET} - \alpha'^B) \\ &= 2\alpha'^B - 12 \cdot (\alpha'^{TETB} - \alpha'^{TET} - \alpha'^B) + 12 \cdot (\alpha'^{DIB} - \alpha'^{DI} - \alpha'^B).\end{aligned}\tag{5.2.4}$$

Similar to the L1 scheme, a general interaction of DI fragments in the whole cell with the B fragment is introduced so that the introduction of asymmetrical molecules would require an adaptation. Similar to L1, the incorporation of two guest molecules can be followed by the addition of DIB terms for interactions with the other B fragment as well as terms to account for the B-B interactions and possibly DIBB terms (where a DIBB fragment is a DIB fragment with the additional B guest). These fragmentation schemes for the total loaded polarizabilities are applied to the calculation of different ZIFs loaded with benzene and refractive indices obtained with the Lorenz-Lorentz equation in the following sections.

### 5.2.2. Refractive Indices for Loaded Zeolitic Imidazolate Frameworks

To retrieve fragments according to the novel fragmentation schemes without repeating the calculations for the empty ZIF, the guest molecule (B) is brought into a single pore of the system in strategically relevant positions, for instance, directly below pore windows or the centre of the pore. Using this approach, the lowest possible local minima can be found in the optimization. The ZIF structure is kept frozen in the optimization process so that solely the guest molecule relaxes inside the pore. From the obtained position of energetic minima further considerations can be made to add further guests. In the fragmentation process, the same fragments as for the empty ZIF can be employed and solely fragments involving the guest molecule need to be calculated. Since the unit cell is kept frozen in this process, the same unit cell volume as for the calculation of the empty ZIF can be applied in the calculation of refractive indices. However, this also means, that not all optimization schemes are applicable unless a new unit cell volume is determined from the recombination of optimized fragments (models **B** and **D2**) to a full unit cell.

Technically, the placement of the B guest molecule is performed as follows with the MOF guest placer script (Appendix B.1). First, a pore window is chosen, in whose close vicinity the guest is supposed to be placed. The coordinates of the Zn<sup>2+</sup> ions of that window are extracted to form a plane and determine the centre of mass, *i.e.* the centre



**Figure 5.2.2.:** Representation of a benzene (B) guest molecule placed inside a 6MR pore window of ZIF-8 at a  $0^\circ$  rotation. The tetrahedrons represent  $\text{Zn}^{2+}$  ions that are coordinated to the blue IM linkers. The yellow sphere highlights the empty pore inside the  $\beta$ -cage that was incorporated with the benzene guest molecule.

of the pore window. Repeating the plane creation and centre of mass determination for B, the molecular plane can be placed on the window plane with their centres of mass coinciding. From here, the guest molecule plane can be rotated in plane if necessary and subsequently shifted into the pore parallel to the window plane. The distance between the parallel window and molecular planes is defined as the distance between their centres of mass ( $d_{\text{Zn}}$ ). An example of the benzene placement in the plane of the 6MR is displayed in Fig. 5.2.2 with the hydrogen atoms pointing towards the  $\text{Zn}^{2+}$  ions, which is considered to be a  $0^\circ$  rotation.

#### 5.2.2.1. ZIF-8

Keppler *et al.*<sup>[35]</sup> measured the tuned refractive indices with different guest molecules for ZIF-8, however, not for benzene. Nevertheless, due to the extensive simplification of the fragmentation approach because of the molecule's symmetry benzene is utilized for this initial loading study. Moreover, the refractive indices of toluene and benzene differ by 0.0014 on average<sup>[106, 389]</sup> and the refractive index of loaded films was described as a blend from the individual component's refractive indices.<sup>[35]</sup> Therefore, some similarities in the loading behaviour can be expected.<sup>[37]</sup> However, this is a crude approximation since the interactions of toluene with the ZIF could vary and due to its substitution, the length and volume of toluene could lead to greatly different loading positions as well as a different number of molecules per pore. Therefore, the comparison yields only a rough estimate.

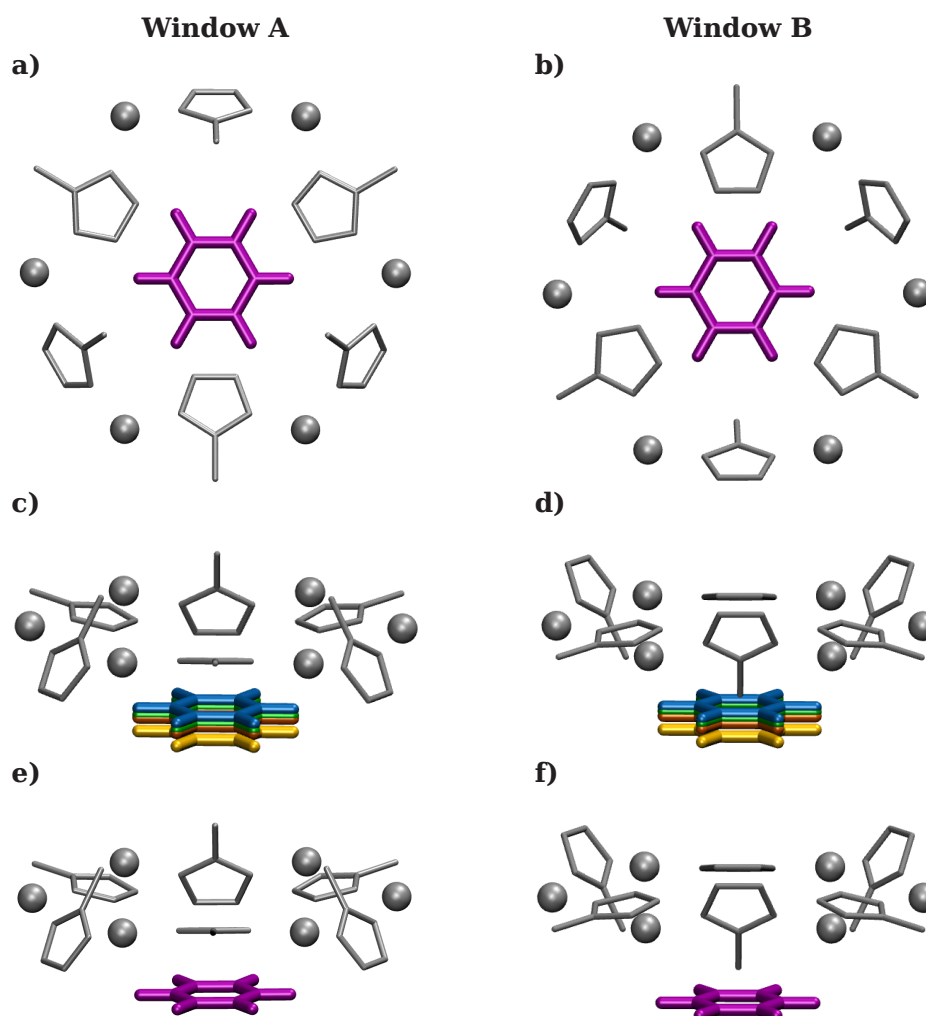
Because for ZIF-8, unfortunately, no results for the DI fragment for model **A** could be obtained, the empty ZIF from the calculation for model **B** was taken in this case. This is an approximation since the DI fragment is also necessary for the calculation of the loading using the L2 scheme. Ideally, the results for model **C** will replace these in further calculations.

Introducing B into the pore below a 4MR and performing the optimization in the frozen  $\beta$ -cage showed, that the molecule migrated towards the 6MR (Fig. B.8 in the appendix). Thus, only approaches with the placement of the guest below these windows were further investigated. Due to the topology of ZIF-8, two different 6MR pore windows could be identified. Window A (Fig. 5.2.3 **a**), **c**) and **e**)), where all methyl substituents of the linker point outwards of the pore (top of Fig. 5.2.3 **c**)) and window B, (Fig. 5.2.3 **b**), **d**), **f**)), where every other linker is rotated with a methyl group pointing inwards to the pore centre. Therefore, two sets of calculations were performed for placing B below either of these pore windows.

Four different distances between the  $\text{Zn}^{2+}$  plane of the respective window and the B plane  $d_{\text{Zn}}$  were examined: 2.7 Å, 2.85 Å, 3.00 Å and 3.45 Å that are displayed in blue, green, red and orange, respectively, in Fig. 5.2.3 **c**) for window A and **d**) for window B. For every distance, additionally, the 30° in-plane rotated guest molecules were also introduced and optimized. After the geometry optimization all starting positions independent of rotation parameters lead to minima with a single distance (but maintaining the rotation). For window A a distance of  $d_{\text{Zn}} = 3.26$  Å and for window B of  $d_{\text{Zn}} = 3.45$  Å was obtained.

Additionally, B was placed in the exact middle of the pore and the total energy for that system is displayed in Table 5.2.1 in comparison to the other obtained positions after the optimization process. Interestingly, the position of B in the pore centre is the highest in energy and therefore, the positions closer to the 6MR are more favourable. For window A, the 30°-rotated molecule leads to a lower energy, than for the unrotated benzene, however, just with a low deviation. For window B, the opposite trend can be observed, with the unrotated system giving a lower total energy. This can be explained by the orientation of the linker in this window with half of the methyl substituents pointing inwards of the pore (Fig. 5.2.3). Therefore, when the hydrogen atoms of the benzene molecule point in the same direction, this leads to a higher observed energy. Thus, in the following calculations, the rotated benzene was employed for window A and the unrotated benzene for window B. The results for these calculations are presented in Fig. 5.2.4.

The results for both windows are in very good agreement. The results for L0 are additionally shown for the introduction of up to four benzene molecules. This leads to a simple shift to higher refractive indices of 0.035 per molecule on average. This is in good agreement with pDFT predicting a shift of 0.033 for every added toluene molecule. Following the L0 approach, in theory, approximately four molecules would be necessary to reproduce the refractive index shift of toluene in the experiment, that is assumed to be close to those for benzene. However, since no interactions are represented whatsoever, in reality, this number of guests per pore probably deviates from this result.

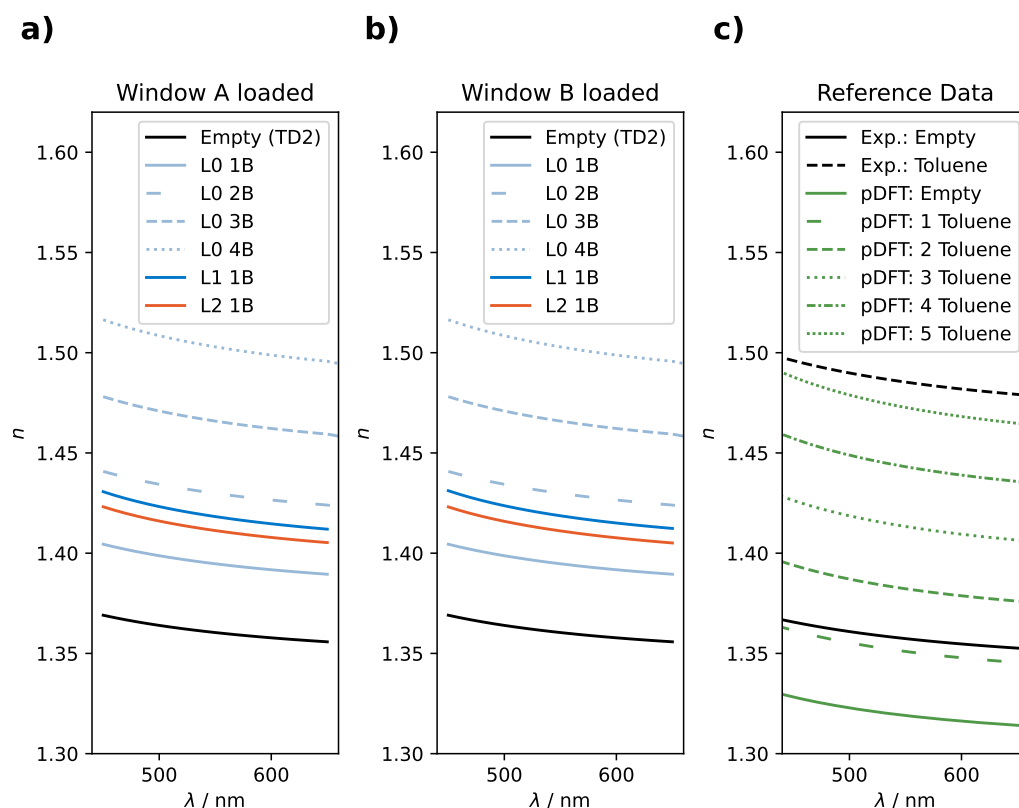


**Figure 5.2.3.:** **a)** and **b)**: Top view and **c)**-**f)** side view of the 6MR Zn<sup>2+</sup> plane and the loaded benzene guest molecule. Colours in **c)** and **d)** represent a  $d_{\text{Zn}}$  of 2.7 Å (blue), 2.85 Å (green), 3.00 Å (red) and 3.45 Å (orange). The pre-optimization positioning of benzene for window A is shown in **a)** and **c)** and for window B in **b)** and **d)**. The post-optimization position is presented in **e)** for window A and **f)** for window B.<sup>||</sup>

**Table 5.2.1.:** Total energies for benzene in the ZIF-8 unit cell placed at different positions and rotated in comparison to the Zn<sup>2+</sup> ions.<sup>||</sup>

Position	Total energy / $E_{\text{h}}(0^\circ)$	Total energy / $E_{\text{h}}(30^\circ)$
<b>6MR - A</b> (3.26 Å)	-58841.492845	-58841.493379
<b>6MR - B</b> (3.45 Å)	-58841.491268	-58841.489526
<b>Middle</b>	-58841.482470	-

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.



**Figure 5.2.4.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure (model A). Results are obtained for the different loading schemes L0, L1, and L2 for **a)** window A and **b)** window B. For L0 different numbers of B fragments were calculated (1B–6B). Reference data from experiment and pDFT is presented in **c)**.<sup>[35, 129, 364]†\*</sup>

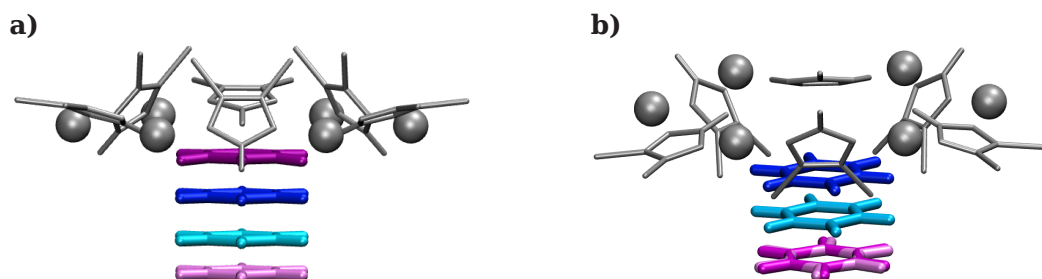
Focussing on the L1 and L2 schemes, the refractive indices are significantly higher than for L0. The average shift from the empty ZIF is 0.058 for the TD1 scheme of both window A and B. The L1 approach leads to even higher refractive indices than the L2 approach, that deviates from the empty ZIF by 0.051 on average for both window A and B. When comparing to the toluene reference, this leads to the assumption, that the number of molecules per pore should be lower than expected from the L0 results. It should, however, be noted, that a simple additive shift is not sufficient in this fragmentation scheme. Despite this limitation, it is expected to outperform the L0 results. As mentioned in Section 5.2.1, this would also neglect the B–B interaction as well as the higher-order interactions of the other fragments with more than one B fragment. Therefore, the refractive index is expected to be somewhat lower for two incorporated benzene molecules than double the shift of a single benzene molecule. In consequence, the observed shift equals a number of approximately 4–6 guest molecules per pore, which would have to be confirmed with more sophisticated fragmentation schemes. This

† Pascal Czember, Leibniz University Hannover, 2022.

\* Erona Shabani, Leibniz University Hannover, 2023.

estimate is in good agreement with the results from pDFT that predicts approximately five toluene molecules per pore, however, for the empty ZIF the refractive index is found to be underestimated, influencing the loading prediction.

#### 5.2.2.2. ZIF-71



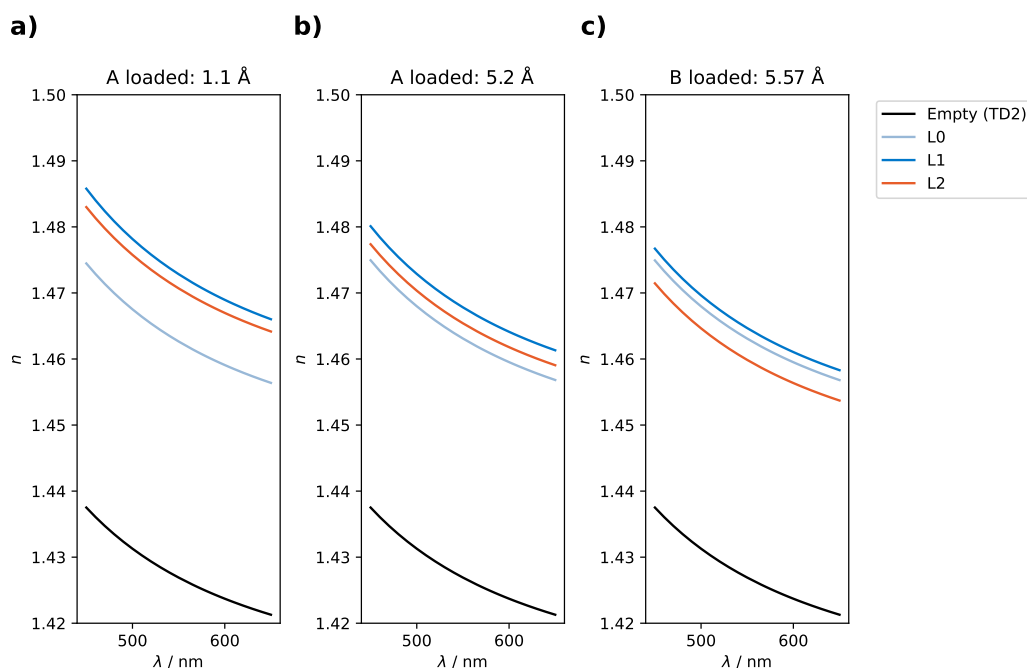
**Figure 5.2.5.:** Representation of benzene positioning in **a)** window A and **b)** window B before (blue tones) and after (pink tones) geometry optimization. The optimization of the light blue benzene position ( $d_{Zn} = 4.00 \text{ \AA}$ ) leads to the light pink optimized position. The optimization of the dark blue benzene position ( $d_{Zn} = 2.5 \text{ \AA}$ ) leads to the dark pink optimized position. **a)** Light pink:  $d_{Zn} = 1.1 \text{ \AA}$ , dark pink:  $d_{Zn} = 5.2 \text{ \AA}$ . **b)** Light and dark pink:  $5.57 \text{ \AA}$ .<sup>§</sup>

For ZIF-71, the investigation of possible guest molecule positioning was repeated for the 6MR on the structure from model **A**, since all fragments lead to feasible results in the empty ZIF. Two window types can be distinguished for ZIF-71. Window A exhibits a linker orientation, where all chlorine substituents point outwards of the pore. In window B, half of the chlorine substituents point towards the pore centre, whereas the other half is oriented inside the pore but pointing to the neighbouring window. The guest molecule was placed in  $d_{Zn} = 2.5 \text{ \AA}$  and  $d_{Zn} = 4.0 \text{ \AA}$  (Fig. 5.2.5). Only  $30^\circ$  rotated benzene molecules were investigated since these are visibly more favourable due to the placement of the chlorine substituents of the DCIM in ZIF-71 (see Fig. 5.2.5). In contrast to ZIF-8 in window A, two favourable positions can be found in the optimization. For window B, a single position is found for both starting geometries, however, the distance to the  $Zn^{2+}$  plane is higher than for ZIF-8, caused by the inwards pointing chlorine substituents on the IM. The refractive indices were calculated for all three single-loaded positions and are presented in Fig. 5.2.6.

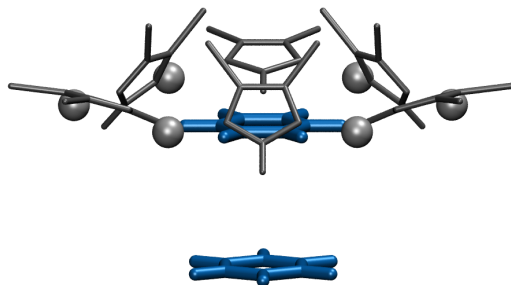
For window A the different benzene positions yield significantly different results. While the L0 curve does not differ much between the positions, the L1 and L2 curves are significantly higher for the  $1.1 \text{ \AA}$  position than for the  $5.2 \text{ \AA}$  position with a deviation of 0.005 for both L1 and L2. For window B the refractive indices for L1 and L2 are even lower than for window A at the  $5.2 \text{ \AA}$  position with a deviation of 0.003 and 0.006 for L1 and L2, respectively, suggesting that interactions have to be accounted for to perform reliable calculations. Unfortunately, no reference data for the loaded ZIF-71 is available for comparison at this point. However, it can be seen, that there is a strong shift when

<sup>§</sup> Daniel Bömke, Leibniz University Hannover, 2022.





**Figure 5.2.6.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure (model A). Results are obtained for the different loading positions **a)** window A at  $d_{\text{zn}} = 1.1 \text{ \AA}$ , **b)** window A at  $d_{\text{zn}} = 5.2 \text{ \AA}$  and **c)** window B of ZIF-71 at  $d_{\text{zn}} = 5.57 \text{ \AA}$  and different fragmentations schemes L0, L1 and L2 in contrast to the results for the empty ZIF.<sup>§\*</sup>



**Figure 5.2.7.:** Window A of ZIF-71 loaded with two benzene molecules in T0B30 rotation.<sup>\*</sup>

introducing a single guest molecule. Therefore, a wide tuning range is expected for this ZIF in possible further experimental studies.

Due to the occurrence of two favoured positions below the 6MR, a double-loading study was performed. The benzene molecules were incorporated in either the same ( $0^\circ$ ) or  $30^\circ$  rotation to each other. For instance, in Fig. 5.2.7, the top benzene molecule is in a  $0^\circ$  rotation (T0), while the bottom benzene molecule is in a  $30^\circ$  rotation (B30). All four possible rotation combinations were produced and optimized giving the total

<sup>§</sup> Daniel Bömke, Leibniz University Hannover, 2022.

<sup>\*</sup> Erona Shabani, Leibniz University Hannover, 2023.

**Table 5.2.2.:** Total energies for two benzene molecules in the ZIF-71 unit cell with the top and bottom rotation of 0° and 30° in comparison to the Zn<sup>2+</sup> ions.\*

Position	Total energy / $E_h$
<b>T0B0</b>	-111861.460439
<b>T30B0</b>	-111861.459849
<b>T0B30</b>	-111861.462661
<b>T30B30</b>	-111861.462293

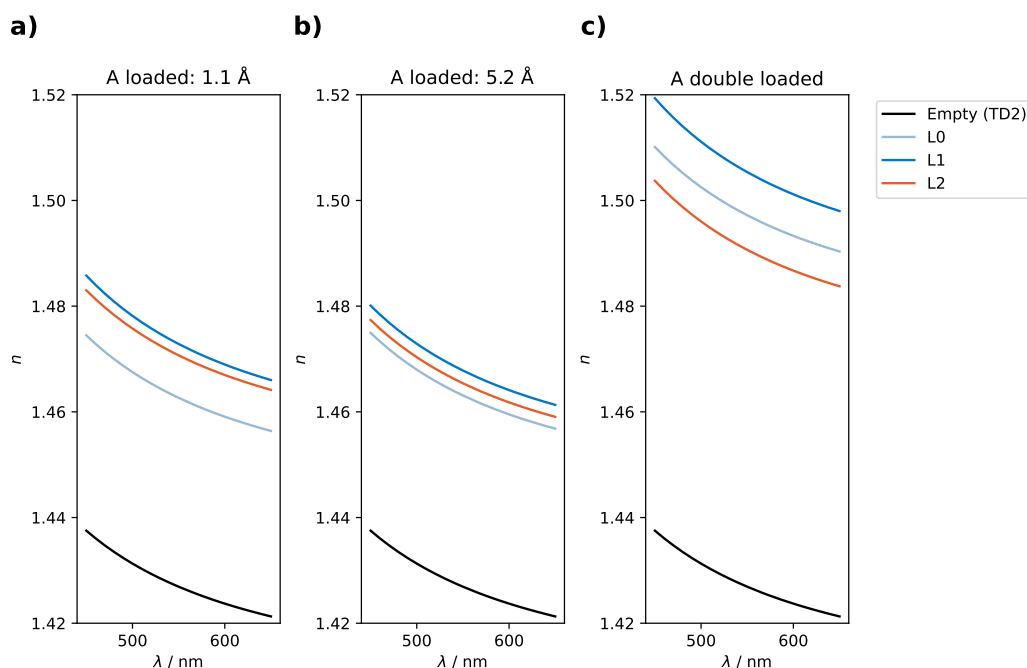
energies listed in Table 5.2.2 with T0B30 emerging as the most favourable position, where the top molecule is shifted towards the bottom molecule by 0.48 Å, whereas the bottom molecule remained at the same position. The refractive indices for this fragment have been calculated by simply replacing the single B to a double BB fragment and, accordingly, replacing DIB to DIBB and TETB to a TETBB fragment, while keeping the equations as shown in Section 5.2.1. The results are depicted in Fig. 5.3.1 c).

The refractive indices for the L0 approach are shifted by 0.07 in comparison to the empty ZIF, which is approximately double the L0 shift of two single guest molecules added although interactions between the molecules were introduced in the double-loading. The L1 and the L2 curve are both shifted by 0.008 and -0.006 in comparison to the L0 curve, respectively. The refractive indices for L1 and L2 are therefore significantly shifted towards high refractive indices, showing a deviation of 0.079 and 0.064, respectively, in comparison to the empty ZIF. For L1, this is found to be a multiple of 1.7 and 1.9 of the loading shift for a single benzene molecule for the position at 1.1 Å and 5.2 Å distance to window A, respectively. For the L2 approach, for the position at 1.1 Å distance to window A, this is a multiple of 1.5 and for the position at 5.2 Å distance to window A, the shift is a multiple of 1.7 in comparison to the loading shift for a single benzene molecule. This roughly coincides with the observation made for ZIF-8. It should be noted, that the shifts between the curves of the fragmentation schemes in the double-loading scheme (Fig. 5.3.1 c)) are significantly higher than in the single-loaded approaches (Fig. 5.3.1 a) and b)). This might be traced back to insufficiencies of the fragmentation ansatz and requires a more sophisticated study of such effects. Additional reference data is indispensable for further determining these schemes for an arbitrary number of incorporated guest molecules.

### 5.3. Summary

Overall, the ansatz for the calculation of refractive indices for ZIFs *via* fragmentation proposed by Treger *et al.* was successfully extended to account for interactions of the fragments. For this purpose, two new TD fragmentation schemes were established and extensively tested, yielding a computational protocol for empty ZIFs with the highly efficient CAM-B3LYP/aug-cc-pVTZ scheme for the fragment calculations. In the employment of different starting structures, the best performance was obtained with an

\* Erona Shabani, Leibniz University Hannover, 2023.



**Figure 5.3.1.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-8 SC XRD structure (model **A**). Results are obtained for the different loading positions **a**) window A at  $d_{Zn} = 1.1 \text{ \AA}$  with single loading, **b**) window A at  $d_{Zn} = 5.2 \text{ \AA}$  with single loading and **c**) window A at both positions (double loading) Different fragmentations schemes L0, L1 and L2 are presented in contrast to the results for the empty ZIF. For **c**) this refers to fragmentation schemes obtained from BB fragments.\*

optimization model (**A**, **B** or **C**), completely independent from pDFT optimized structures, thus, directly applicable to experimentally yielded structures for future calculations.

Furthermore, the application of CP-SCF and linear-response approaches could be compared and gave coinciding results. Therefore, for a simple screening of static polarizabilities, the CP-SCF approach is sufficient, however, for further comparison to experiment at common wavelengths, such as  $\lambda = 589 \text{ nm}$ , the linear-response ansatz enables the calculation of a full frequency-dependent set of refractive indices. Furthermore, the evaluation of the established protocol could be evaluated on the popular representative ZIF-8 and gave results in excellent agreement with experimental reference data and especially for small wavelengths and overall curve progression in some cases outperforming data obtained with HSE06 in pDFT.

Since some SCF instabilities occurred, electrostatic embedding was applied to enhance the description of  $Zn^{2+}$ , IM and TET fragments. Specifically shifted for the BU fragmentation scheme and curve progressions for the TD1 fragmentation scheme consequently improved and can be recommended for the  $Zn^{2+}$  and IM fragments. However, the calculations of the refractive index for the TET fragment resulted in no major improvement and could not be finished for the DI fragment for ZIF-8. Moreover, the refractive indices

\* Erona Shabani, Leibniz University Hannover, 2023.

for ZIF-71, ZIF-90 and ZIF-318 could be easily determined with the developed protocol and partially compared to reference data to yield overall good agreement. ZIF-318 showing an overestimation of the refractive indices from fragmentation schemes is subject to further investigation, for instance, with electrostatic embedding in TD2, different optimization schemes or possibly experimental phase-pure unit cell volumes. Approximate predictions for the refractive index range could be given for ZIF-71, however, further confirmation *via* experiment is necessary. Overall a recommendation for the second-order fragmentation ansatz (TD2) in an optimized model (**B** or **C**) can be given for further calculations of refractive indices in ZIFs. For this purpose, the fragmentation scheme can be easily adapted to account for different ZIFs possibly even in different topologies.

Further adaptations to the established fragmentation schemes allowed for the introduction of guest molecules into the pore of ZIF-8 and ZIF-71. The calculation of the refractive indices for the loaded systems was successfully performed for a single and double-loaded system with a benzene guest molecule. The comparison to reference data for toluene allowed a crude prediction of the molecule number per pore. First trends resulting from the calculations showed, that for ZIF-8 the number of benzene/toluene molecules per pore should be in the range of 5–7 molecules which is supported by results from pDFT calculations for toluene. For ZIF-71 loading calculations were performed and more than one favourable position for the guest molecule below the 6MR pore windows could be identified. A first approach for the double-loading was performed and builds the fundament for higher loading studies. However, further studies are indispensable in order to refine the proposed loading schemes, specifically for unsymmetrical molecules. Further reference data is imperative for the evaluation of the obtained data.

## 6 | Summary, Conclusions and Outlook

Accurate theoretical calculations of optical properties can be challenging, specifically when dealing with extended system sizes. The required quantum-mechanical (QM) treatment greatly exceeds the possibilities of common quantum-chemical calculations in terms of computational demand. However, the quantum nature of optical processes makes this treatment indispensable even when dealing with large systems. Subsystem-based methods open the possibility to apply the quantum-chemical methods to extended systems by partitioning the full system, yielding computationally more feasible subsystems, while maintaining the necessary accuracy. Depending on the system type, either local embedding schemes or fragmentation approaches can be applied. The first yielding localized optical properties for a dominant subsystem, whereas the second recombines individual subsystem contributions yielding the global property.

Polarizable local embedding schemes were employed to calculate optical excitations in extended systems, that are located on one subsystem. In contrast to most other studies, a comparison of two different major representatives of advanced embedding schemes was performed in the same consistent, theoretical framework. These representatives are the quantum-mechanical/molecular-mechanical (QM/MM)-based polarizable embedding (PE) and the quantum-mechanical/quantum-mechanical (QM/QM)-based frozen-density embedding (FDE). The implementation of said framework allows for the observation of dissected interaction effects introduced by the respective embedding schemes and a classification of the importance of the occurring interaction effects.

In Chapter 4, it has been investigated how these dissected interaction contributions for the different embedding approaches influence the calculation of properties for optical excitations. Before addressing the distinct characteristics in a numerical comparison, the aforementioned theoretical framework was established for the QM/MM-based as well as the QM/QM-based methods, putting their characteristics in a common context. Based on this framework, the methods were implemented within a unified computational setup based on Dalton software, ensuring maximum comparability and incorporating the formulation of the established interaction models. Building upon this implementation, optical properties were calculated for *para*-nitroaniline (*p*NA) and pentameric formyl thiophene acetic acid (pFTAA) using both approaches. The obtained framework allows for the segregation of the underlying individual interaction effects, thereby enabling the formulation of models with comparable interactions for both systems and highlighting distinctions between the two approaches. The results for the target molecules were subdivided into the arising interaction models. This comprehensive analysis allowed for an understanding of the individual contributions of the various interaction models and their effects on the overall optical properties of the target system. By separating

the different interaction effects, valuable insights were gained into how each model influenced the optical properties, and a comparison of the performance of the two approaches in accurately capturing the behaviour of *p*NA and pFTAA was made. This study paved the way for a deeper understanding of the interactions and their effects on the optical response of these molecules.

For both embedding schemes, a classification of the polarization interactions into two models could be performed. The first model considers electrostatic interactions between subsystems and includes solely the ground-state polarization of the main subsystem. The second model takes into account mutual polarization in the ground state. These models were identified within the standard formulations of both embedding approaches. In the calculations for the PE scheme calculations of pFTAA erroneous electron spill-out (ESO) effects occurred. These could be diminished by the application of atomic pseudopotentials. Thereafter, good agreement has been obtained for the calculated excitation energies and oscillator strengths for both embedding schemes including mutual ground-state polarization in comparison with a full supermolecular calculation and neglecting differential polarization. This is in line with results from previous studies, expecting ground-state polarization to give a major contribution to the full interaction.<sup>[76, 77]</sup> For both embedding schemes exceedingly small deviations were found for the electrostatic treatment of the environment polarizing solely the main subsystem as well as for the mutual ground-state polarization. This suggests, that QM effects did not significantly influence the interaction of these systems, however, the introduced repulsion is necessary in order to counteract the ESO effect for pFTAA.

In general, excitation energies could be shown to be less sensitive to polarization effects, than oscillator strengths for both molecules. Therefore, the impact of differential polarization is lower for excitation energies. For oscillator strengths, the effects were sizeable in part of the results but very dependent on the specific structural configurations (snapshot). In the cases, where differential polarization showed a big influence, agreement of the oscillator strengths with the reference values was in many cases worsened. It was demonstrated that in these calculations the use of an effective external field (EEF) improved the results in all cases showing to be essential for the reproduction of polarization obtained with reference calculations. For *p*NA the effects of differential polarization can be sizeable for individual snapshots with and without including EEF, but it is small on average. Specifically for pFTAA, a system that is challenging to describe due to its size and highly anionic charge, differential polarization had a significant impact. However, the agreement with the reference calculation was only good, when including EEF effects or overall neglecting differential polarization. Altogether, this effect boils down to the dependence on the individual snapshots, where the dynamic response of the environment might be sizeable, but on average does not strongly impact the optical properties, which is in line with previous studies.<sup>[76, 77]</sup>

Overall, it was demonstrated that the established theoretical framework and computational setup allow the one-to-one comparison of PE and FDE on the same fundament and is a valuable tool in the detailed analysis of the different contributions to the solvent shift. This facilitates the characterization and deeper understanding of solvent effects on these molecules. This methodology could possibly be extended to further

---

systems dissecting their interaction contributions. Overall, while general statements regarding ground-state polarization are applicable, specific attention must be given to the impact of differential polarization in PE, and the incorporation of EEF effects can be recommended for accurately reproducing polarization effects on oscillator strengths with PE. For challenging systems, that are prone to ESO effects, the use of atomic pseudopotentials can be recommended without any preservation. The present framework is deemed promising for the calculation and interpretation of optical spectra of solvated chromophores. Furthermore, the framework can be extended to account for more complex environments such as the protein environment of a chromophore. The investigation of additional effects such as averaged environmental contributions or differential polarization effects for FDE as seen in state-specific approaches logically follow.<sup>[66]</sup> Moreover, more excited state properties could be considered due to their applicability to fluorescent biomarkers, however, requires the introduction of a FDE optimization scheme to allow for the minimization of excited states.

For the calculation of global optical properties, the desired characteristic is not calculated localized on a main subsystem, but distributed over all subsystems. An energy-based molecular fragmentation approach was developed for zeolitic imidazolate frameworks, that in contrast to the computational advances in other studies on metal-organic frameworks (MOFs) combines a significant decrease in computational demand with the preservation of essential accuracy. Moreover, contrary to other molecular fragmentation studies on MOFs, the presented approach includes higher-order many-body effects, that are indispensable for the accurate reproduction of optical properties. The performed evaluation of this method is applicable for the rationalization and furthermore the prediction of optical properties for differently substituted ZIFs. Therefore, it has a high potential to assist the characterization and design of further ZIFs. Moreover, a strength of the proposed fragmentation scheme is the possibility to examine the influence of guest molecules incorporated into the porous network of the investigated systems and draw approximate conclusions for the tuning range of the calculated optical properties. However, in order to reliably predict this loading behaviour a benchmark with existing systems has yet to be performed.

In global subsystem-based methods the property of interest is calculated from the individual subsystem contributions, that can then be recombined to give the global optical property. A many-body type ansatz can be employed to account for the individual subsystem properties and the many-body effects between them. In previous studies such approaches have been applied to many systems, however, for MOFs, only one approach neglecting all many-body effects is known.<sup>[129]</sup> Specifically the determination of refractive indices has gained attention in optical applications<sup>[35]</sup> and the periodically repetitive pattern of MOFs allows for the optimal application of molecular fragmentation methods. For this, a sufficient number of many-body terms has to be included to obtain feasible results. Additionally, the convergence of the property of interest can be increased by providing overlapping fragments as primary fragments in the fragmentation scheme.<sup>[92, 93, 234]</sup> Since these schemes have not been applied to ZIFs yet, a computational protocol has to be established and evaluated to reliably predict refractive indices. Moreover, in experimental studies, a tuning of the refractive index range could

be obtained by incorporating guest molecules into the porous structure of the ZIFs.<sup>[35]</sup> In order to reflect this in the calculations, the development of novel fragmentation schemes is necessary allowing for approximate predictions of the loading process.

In chapter Chapter 5 fragmentation schemes including many-body effects have been developed for the calculation of refractive indices in ZIFs. It was evaluated on the example of ZIF-8, in an extensive computational study involving the calculation of polarizability volumes for a variety of XC functionals and basis sets calculated with a coupled-perturbed self-consistent field (CP-SCF) and a linear-response ansatz. The latter allows the frequency-dependent determination of polarizability volumes, while solely static polarizabilities are determined for the former. The calculation of refractive indices from the polarizability volumes *via* the Lorenz-Lorentz equation has been employed to further evaluate the approaches on different geometries for ZIF-8. The obtained computational protocol has been transferred to the calculation of refractive indices for ZIFs with the same topology but differently substituted organic linkers. The comparison with reference data allowed further evaluation and prediction for the other empty ZIFs. Moreover, in this study, the influence of continuum solvation on the individual fragments was examined.

Subsequently, the established framework was extended to account for the incorporation of guest molecules in the ZIF's pore. Energetic minima were found for benzene guest molecules placed inside the pore. Considerations concerning symmetry and placement of the guest molecule and interactions with the pore's subsystems allowed for the calculation of refractive indices on ZIF-8 and ZIF-71 while incorporating a single benzene molecule and up to two benzene molecules per pore, respectively. The results from this approximate scheme could then be employed to give rough estimates about the number of molecules incorporated per pore.

It has been found, that in contrast to other studies<sup>[253, 367]</sup> on the calculation of polarizabilities, a range-separated hybrid CAM-B3LYP XC functional yielded the best results for all fragment types. Double-hybrid functionals performed similarly, however, could often not be utilized due to the occurrence of numerical instabilities. Although it has been shown, that the employment of sizeable basis sets is essential for the accurate calculation of polarizabilities<sup>[375-377]</sup>, the property of interest has been close to convergence for an aug-cc-pVTZ basis set in this study. The computational protocol was then applied to the calculation of refractive indices based on differently optimized structural geometries. Structures were obtained directly from single crystal X-ray diffraction (SC XRD) structures, that then were further optimized for some models. Additionally, periodic density-functional theory (pDFT)-optimized structures<sup>[129]</sup> were employed, that were further optimized in one model. It was found, that for ZIF-8, those optimization schemes gave the best results in comparison to experiment, that were directly extracted from experimental structures and subsequently optimized with the B97-3c/def2-mTZVP approach, highlighting the independence from pDFT geometries in these calculations. The optimization was either performed for the extracted fragments or the extracted unit cell of the respective ZIF. Specifically the latter yielded excellent results for the fragmentation scheme incorporating the highest (second) order of many-body effects albeit the approximation made in this order. Generally, that fragmentation scheme (TD2)



---

gave the best results in all calculations, outperforming the schemes with lower-order and without many-body effects, TD1 and BU, respectively, as well as the results for HSE06.<sup>[129]</sup> Due to self-consistent field (SCF) instabilities for some fragments the electrostatic embedding of the fragments in a dielectric continuum was performed for all fragments of ZIF-8. The relative permittivity was estimated from previous calculations. For the BU approach major improvements have been obtained, however, still giving results far from the experimental reference. The curve progressions as well as the curve shift could be improved in comparison to experiment.

The transfer of the computational protocol for the different fragmentation schemes to other ZIFs has confirmed the good performance of the TD2 fragmentation scheme and allows for the estimation of possible refractive index ranges for those ZIFs where no experimental data is available at this point. The predictions made here should, however, be checked against upcoming results in order to further advance the proposed fragmentation scheme. The continuum solvation of the fragments similarly to the ZIF-8 results improved the curve shift for the BU approach, but only slightly improved the results for the TD1 approach. For the TD2 approach no results are available for ZIF-8 due to reoccurring SCF instabilities. Moreover, embedding of the fragments for the TD2 approach should be subject of future investigations of the remaining ZIFs. Specifically, for ZIF-318, where overestimation of the refractive indices were observed for all fragmentation schemes, this approach could result in improvements when comparing to reference data. Furthermore, extending the electrostatic embedding to the employment in the different optimization schemes could improve the agreement with the experimental refractive indices and eliminate possible occurring error cancellation. For this, the volumes for the  $\beta$ -cage-optimized structures should be critically examined in comparison to the experimentally obtained volumes.

For the purpose of incorporating guest molecules the fragmentation schemes were adapted to account for the polarizability of a single benzene molecule introduced into the ZIF-8 pore. For this, optimization of the guest molecule in the frozen unit cell of the ZIF was performed to retrieve favoured positions inside the pore. The placement below the six-membered ring (6MR) is clearly favoured in contrast to a placement below the pore center or four-membered ring (4MR). The latter can be attributed to the small pore window opening that exhibits close organic linkers interacting with the guest molecule. Therefore, the calculation of refractive indices was performed for the obtained favourable position for the two 6MR pore windows. Feasible results were obtained, however, no experimental reference data was available. Generally, incorporation of benzene leads to an overall shift of the frequency-dependent refractive index curves, while the frequency dependence remains similar. The curve shift obtained for the loading process from the BU approach is smaller than for the other fragmentation schemes. This can be traced back to the lack of many-body effects for the introduced guest molecules. The results from the top-down approaches (TD1 and TD2) suggest an approximate number of 4–6 guest molecules per pore, which is in good agreement with the reference data for toluene.<sup>[35, 364]</sup>

The calculations were then transferred to the loaded ZIF-71, where the optimization of the benzene molecule located inside the pore yielded two favourable positions for one of

the 6MR pore windows. In a single-loading study, the results for both positions in that window have shown only small deviations for the TD1 and TD2 schemes, with the results for the benzene molecule placed deeper inside the pore giving lower refractive indices than for molecules located in closer vicinity to pore windows. This trend is also obtained for the other 6MR window, where the guest molecule is placed even further into the pore. Solely for the BU approach only slight deviations for the different pore positions have been observed. This highlights the importance of including the higher-order many-body effects in the calculations.

In a double-loading study, orientations of the two incorporated benzene molecules were examined for the two favoured positions below the pore window. The most favoured position was then examined in further calculations. Great deviations between the results for all fragmentation schemes have been obtained, however, giving a shift that is greater than the double of two single-loading events. This suggests, the influence of the interaction between the two guest molecules is impacting the refractive index. These results are rough estimates of the possible refractive indices for this ZIF. They suggest, however, that for ZIF-71 a wide tuning range by the incorporation of guest molecules can be obtained.

In order to make reliable predictions, further reference data for the benzene loading would greatly facilitate the evaluation of the proposed scheme. Alternatively, a more extensive fragmentation scheme taking into account the lower symmetry of toluene and other (multiple) guest molecules could be introduced to allow comparison to available reference data. The proposed scheme did, however, lay the theoretical and computational foundation for the application of fragmentation schemes for these systems and allowed for predictions that could be further extended to structural isomorphs. Incorporation of these schemes in the in-house fragmentation script FragPy<sup>[82]</sup> potentially opens up the possibility to examine higher-order fragmentation schemes potentially also in multilevel schemes as well as the effects of geometry optimization within the established fragmentation schemes.

In summary, this thesis provides insights into the accurate theoretical calculation of optical properties in challenging systems using subsystem-based methods. The proposed computational protocols allowed for a detailed analysis of the individual interaction effects and their contributions to the overall optical properties. The comparison of local embedding methods facilitated a better understanding of the strengths and limitations of each approach and the importance of described effects. The development of fragmentation schemes allowed for the prediction of optical properties for several structural isomorphous ZIFs. Furthermore, the extension of the computational protocol to incorporate guest molecules in the porous structure of ZIFs opened up the possibility for an approximate prediction of the tuning range of the calculated optical properties in loaded systems. The results from this thesis lay the foundation for further investigations in the field of optical properties in complex systems, allowing for a more comprehensive and accurate characterization of these systems and their interactions with the environment.

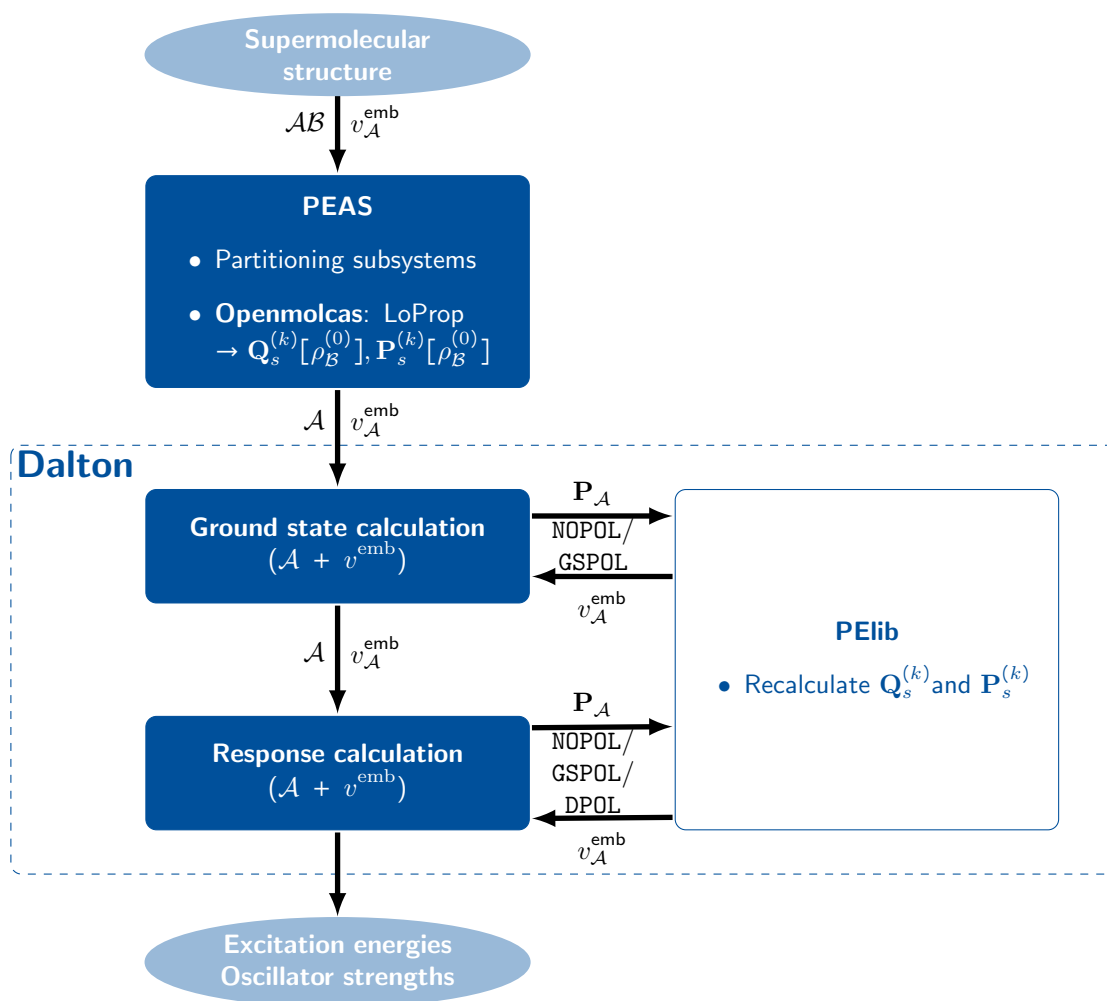
**Part III.**  
**Appendix**



# A | Local Optical Properties

## A.1. Computational Setup

### A.1.1. Polarizable Embedding



**Figure A.1.:** Workflow diagram of the PE scheme. The elliptical nodes represent input and output of the calculation. Dark blue boxes denote major work steps in the workflow. The dashed box represents the Dalton part of the workflow and its substeps inside. The white box represents submodules. Arrows denote data exchanged in between programs and models.

### A.1.2. Frozen-Density Embedding

The workflow presented in Section 4.2.1.2 was implemented in PyADF.<sup>[61, 296]</sup> The implementation (mainly for calling Dalton) is presented in listings A.1 and A.2. Subsequently, the PyADF input performing the FDE cycles is shown on the example of *pNA* and *pFTAA* in listings A.5 to A.9 and listings A.10 to A.16, respectively.

**Listing A.1:** PyADF code: Extracts from adapted classes in DaltonSinglePoint module

---

```
class daltonsettings:
    """
    Class that holds the settings for a Dalton calculation..

    @group Initialization:
        __init__,
        set_method, set_functional
    @group Input Generation:
        get_wavefunction_block
    @group Other Internals:
        __str__
    """

    def __init__(self, method='DFT', functional='LDA', dftgrid=None, freeze_occ=0,
                iterations=None, nexci=None, freeze_virt=0, memory=None):
        """
        Constructor for daltonsettings.

        All arguments are optional, leaving out an argument will choose
        default settings.

        @param method: the computational method, see L{set_method}
        @type method: str
        @param functional:
            exchange-correlation functional for DFT calculations,
            see L{set_functional}
        @type functional: str
        @param dftgrid: the numerical integration grid for the xc part in DFT,
            see L{set_dftgrid}
        @type dftgrid: None or str
        @param freeze_occ: number of occupied orbitals to freeze, see L{set_freeze}.
        @type freeze_occ: int
        @param freeze_virt: number of virtual orbitals to freeze, see L{set_freeze}.
        @type freeze_virt: int
        @param memory: the maximum total memory to use (in MB)
        @type memory: integer
        """
```

---

**Listing A.2:** PyADF code: Extracts from adapted classes in DaltonSinglePoint module Part 2.

---

```
self.method = None
self._functional = None
self.dftgrid = None
self.freeze_occ = None
self.freeze_virt = None
self.memory = None
self.iterations = None

self.set_method(method)
if self.method == 'DFT':
    self.set_functional(functional)
self.set_dftgrid(dftgrid)
self.set_freeze(freeze_occ, freeze_virt)
self.set_memory(memory)
self.set_iterations(iterations)
[.....]

def set_iterations(self, iterations):
    """
    Set number of iterations in SCF (default = 60).

    @param iterations: the maximum number of SCF iterations to use in SIRIUS module
    @type iterations: integer
    """
    self.iterations = iterations

def set_nexci(self, nexci):
    """
    Set the number of excitations to calculated.

    @param nexci: the number of excitations to calculate
    @type nexci: int
    """
    self.nexci = nexci

def get_wavefunction_block(self):
    block = "***WAVE FUNCTIONS\n"
    if self.method == 'HF':
        block += ".HF\n"
    elif self.method == 'DFT':
        block += ".DFT\n"
        block += " " + self.functional + "\n"
        if self.dftgrid is not None:
            block += "*DFT INPUT\n"
            block += "." + self.dftgrid.upper() + "\n"
```

---

**Listing A.3:** PyADF code: Extracts from adapted classes in DaltonSinglePoint module Part 3.

---

```
    elif self.method in ['CC', 'CCSD', 'CCSD(T)']:
        block += '.CC\n'
    else:
        raise PyAdfError('Unknown method in Dalton job')
    if self.iterations:
        block += "*SCF INPUT\n"
        block += ".MAX DIIS\n"
        block += " " + str(self.iterations) + "\n"
    return block

def __str__(self):
    """
    Returns a human-readable description of the settings.
    """
    s = f' Method: {self.method} \n'
    if self.method == 'DFT':
        s += f' Exchange-correlation functional: {self.functional} \n'
    elif self.method.startswith('CC'):
        s += f" Number of frozen occupied orbitals: {self.freeze_occ:d} \n"
        s += f" Number of frozen virtual orbitals: {self.freeze_virt:d} \n"
    return s
```

---



---

**Listing A.4:** PyADF code: Extracts from adapted classes in DaltonSinglePoint module Part 4.

---

```

class daltonsinglepointjob(daltonjob):
    """
    A class for Dalton single point runs.

    See the documentation of L{__init__} and L{daltonsettings} for details
    on the available options.

    Corresponding results class: L{daltonsinglepointresults}

    @Note: Right now, HF, DFT, and CC jobs are supported.

    @Note: Importing of embedding potential requires a modified Dalton version.

    @group Initialization:
        set_restart
    @group Input Generation:
        get_dalton_block, get_integral_block, get_molecule, get_options_block,
        get_other_blocks, get_properties_block
    @group Other Internals:
        print_extras, print_molecule, print_settings

    """
[.....]

    def before_run(self):
        super().before_run()
        if isinstance(self.fdein, GridFunctionPotential):
            self.fdein.get_xyzwvfile('EMBPOT', add_comment=False, endmarker=True)
        elif self.fdein is not None: # adffragmentsresults
            self.fdein.export_embedding_data('EMBPOT','FRZDNS')

    def after_run(self):
        super().after_run()
        if isinstance(self.fdein, GridFunctionPotential):
            os.remove('EMBPOT')
        elif self.fdein is not None:
            os.remove('EMBPOT')
            os.remove('FRZDNS')

```

---

**Listing A.5:** PyADF input code for pNA Part 1.

---

```
import numpy
import xcfun
from pyadf.Plot.FileReaders import GridFunctionReader
import copy

input = 'in.pdb'
cycles = 3 # How many FDE cycles
main_functional = 'CAMB3LYP' # Which functional for the active fragment
env_functional = 'B3LYP' # Which functional for the environment fragments
dalton_basis = 'aug-cc-pVDZ' # Which basis set in Dalton calculations
response_sp = True # Do a response calculation for the Dalton/Dirac SP calculation
                    # of main fragment, too? (For comparison)

#=====#
# READ FILE #
#=====#
molfile = os.path.join(pyadfenv.outdir, input)
supermol = molecule(molfile, 'pdb')
supermol.set_symmetry('NOSYM')
frags = supermol.separate()

#=====#
# IDENTIFY PNA (must be first) #
#=====#
pna = frags[0] # identify pna as first molecule in fragment list
pna.write(os.path.join(pyadfenv.outdir, 'pna.xyz')) # write out pna coordinates
pna.set_symmetry('NOSYM')

#=====#
# SUPER MOL GRID #
#=====#
settings = adfsettings(functional=main_functional, zlmfit=True)
settings.set_integration(6.0, becke='good') # grid settings
settings.set_ncycles(1) # do not converge SCF, we only need the grid
settings.set_save_tapes([21, 10])

res_super_adf = adfsinglepointjob(supermol, basis='QZ4P', settings=settings, options=
['TOTALENERGY', 'NumericalQuality good']).run() # adf calculation
agrid = adfgrid(res_super_adf) # save supermoelcular adf grid in variable
```

---

**Listing A.6:** PyADF input code for pNA Part 2.

---

```

#####
# SINGLE POINT ON ADF GRID IN DALTON          #
#####
env_frgs_res = fragmentlist()
updated_frgs = fragmentlist()

#####
# a) MAIN FRAGMENT                            #
#####
#start dalton calculation
settings = daltonsettings(method='DFT',functional=main_functional)
if response_sp:
    response = ['*LINEAR','.SINGLE RESIDUE','.DIPLN','.ROOTS','5',
               '.NSTART','10','.PRINT','4']
    res_pna = daltonsinglepointjob(pna, dalton_basis, settings=settings,
                                   response=response).run() # response calculation of main fragment
else:
    # ground-state calculation of main fragment
    res_pna = daltonsinglepointjob(pna, dalton_basis, settings=settings).run()
# update fragment data in list
updated_frgs.append(fragment(res_pna, pna,isfrozen=False))

#####
# b) ENVIRONMENT FRAGMENTS                    #
#####
settings = daltonsettings(method='DFT', functional=env_functional)
settings.set_iterations(200)
#do calculations for all fragments
for frag in frgs:
    if not frag == frgs[0]: # all fragments except main fragment
        frag.set_symmetry('NOSYM')
        # ground-state calculation for environment fragment
        frag_res = daltonsinglepointjob(frag, dalton_basis, settings=settings).run()
        env_frgs_res.append(frag_res)
        # update fragment in list
        updated_frgs.append(fragment(frag_res, frag, isfrozen=True,
                                    fdeoptions={"RELAXsave":""}))

#####
# Embedding potential                          #
#####
dens_pna = res_pna.get_density(grid=agrid, order=2) # get density for pna
pot_pna = res_pna.get_potential(grid=agrid, pot='elstat') # get pot elstat for pna
pot_pna.get_xyzwvfile(filename='PNA-ELSTAT-POT') # write out elstat potential
dens_env = None
pot_env = None
old_dens = None

```

---

**Listing A.7:** PyADF input code for pNA Part 3.

---

```
for res in env_fragments_res:
    dens_f = res.get_density(grid=agrid, order=2)
    pot_f = res.get_potential(grid=agrid, pot='elstat')
    # add up environment densities and potentials
    if dens_env is None:
        dens_env = dens_f
    else:
        dens_env = dens_env + dens_f
    if pot_env is None:
        pot_env = pot_f
    else:
        pot_env = pot_env + pot_f
#####
# EMB POT SETTINGS: #
#####
embed_settings = EmbedXCFunSettings()
embed_settings.set_fun_nad_xc({'BP86': 1.0})
embed_settings.set_fun_nad_kin({'PW91k': 1.0})
embed_settings.show_functionals()
embed_eval = EmbedXCFunEvaluator(settings=embed_settings)
embpot_pna = None
# CREATE EMBPOT FOR MAIN FRAGMENT FROM DENSITIES AND POTENTIAL
embpot_pna = embed_eval.get_emb_pot(dens_pna, dens_env, pot_env) # generate emb. potential
embpot_pna.get_xyzvwfile(filename='EMBPOT-FIRST-EMBPOT') # write out emb. potential

#####
#####
# Starting FDE cycles #
#####
#####
for cycle in range (1,cycles+1):
    print ('ENTER FDE CYCLE')
    # SP CALC FOR MAIN FRAG WITH EMBPOT FOR MAIN FRAG
    settings = daltonsettings(method='DFT',functional=main_functional)
    settings.set_iterations(200)
    res_pna = daltonsinglepointjob(pna, dalton_basis, settings=settings,
                                   fdein=embpot_pna).run()
    # update active density (MAIN) and put it in fragment list
    # make all of them frozen for next loop
    for f in updated_fragments:
        if f.get_total_molecule() == pna:
            f.results = res_pna
            f.isfrozen = True
```

---

**Listing A.8:** PyADF input code for pNA Part 4.

---

```

#####
#loop over env fragments, unfreeze one at a time, create new emb pot and #
# do sp calculation with new embpot, freeze #
#####
settings = daltonsettings(method='DFT',functional=env_functional)
settings.set_iterations(200)
# IF FRAG IS TO BE RELAXED, UNFREEZE, THEN GET DATA FOR EMBPOT IN LOOP
# CREATE EMBPOT_FRAG AND DO SP CALC FOR FRAG
for f in updated_fragments:
    if f.has_fdeoption("RELAXsave"):
        # unfreeze -> new active fragment
        f.isfrozen = False
        # get data for embpot of main frag
        dens_main = f.results.get_density(grid=agrid, order=2)
        dens_env = None
        pot_env = None
        settings = daltonsettings(method='DFT',functional=env_functional)
        settings.set_iterations(200)
        # check out embpot data for all frags that are not active
        # in this cycle
        for frag in updated_fragments:
            if not frag == f:
                dens_f = frag.results.get_density(grid=agrid, order=2)
                pot_f = frag.results.get_potential(grid=agrid, pot='elstat')
                if dens_env is None:
                    dens_env = dens_f
                else:
                    dens_env = dens_env + dens_f
                if pot_env is None:
                    pot_env = pot_f
                else:
                    pot_env = pot_env + pot_f

        # get emb pot for fragment
        embpot_f = embed_eval.get_emb_pot(dens_main, dens_env, pot_env)
        print ('got embpot data')

        # update results for fragment in list
        f.results = daltonsinglepointjob(f.get_total_molecule(), dalton_basis,
                                         settings=settings, fdein=embpot_f).run()

        # freeze for next cycle
        f.isfrozen = True
        print ('1 ENV FRAGMENT IN CYCLE DONE')

```

---

**Listing A.9:** PyADF input code for pNA Part 5.

---

```

#####
#create new embpot for MAIN #
#####
#none so they don't add up to the existing ones
dens_env = None
pot_env = None
res_pna = None
embpot_pna = None
#####
#get necessary densities and potential for emb pot #
#####
for f in updated_frags:
    if not f.has_fdeoption("RELAXsave"): # all frozen fragments
        res_pna = f.results
        f.isfrozen = False
    if f.has_fdeoption("RELAXsave"): # unfrozen fragment
        dens_f = f.results.get_density(grid=agrid, order=2)
        pot_f = f.results.get_potential(grid=agrid, pot='elstat')
        # add up environment density and potential
        if dens_env is None:
            dens_env = dens_f
        else:
            dens_env = dens_env + dens_f
        if pot_env is None:
            pot_env = pot_f
        else:
            pot_env = pot_env + pot_f
dens_pna = res_pna.get_density(grid=agrid, order=2)
#####
# CREATE EMBPOT FOR MAIN FRAG #
#####
embpot_pna = embed_eval.get_emb_pot(dens_pna, dens_env, pot_env)
# unfreeze main fragment
for f in updated_frags:
    if not f.has_fdeoption("RELAXsave"):
        f.isfrozen = False
    print ("=====CYCLE ", cycle, " DONE.=====")
#####
# End of cycles #
# Final calculation for main fragment #
#####
#start dalton response calculation
settings = daltonsettings(method='DFT',functional=main_functional)
settings.set_iterations(200)
response = ['*LINEAR', '.SINGLE RESIDUE', '.DIPLN', '.ROOTS', '5', '.NSTART',
            '10', '.PRINT', '4']
results = daltonsinglepointjob(pna, dalton_basis, settings=settings,
fdein=embpot_pna, response=response).run()

```

---

**Listing A.10:** PyADF input code for pFTAA.

---

```

import numpy
import xcfun
from pyadf.Plot.FileReaders import GridFunctionReader
from pyadf.Molecule.OBFreeMolecule import OBFreeMolecule

frag_dict = 'Frag_Dict.py'
import os, sys
sys.path.append(os.path.dirname(os.path.expanduser(frag_dict)))
from Frag_Dict import *    # read in charges from Frag_Dict.py

#=====#
#INPUT DATA                #
#=====#
input = 'in.pdb'
cycles = 3 # How many FDE cycles
response_sp = True #
main_functional = 'CAMB3LYP' #
env_functional = 'B3LYP' #
dalton_basis = 'aug-cc-pVDZ' # Which basis set in Dalton calculations
dalton_mem = 2000

#=====#
# READ FILE                #
#=====#
filename = os.path.join(pyadfcenv.outdir, input)
supermol = molecule(filename, 'pdb')
supermol.set_symmetry('NOSYM')

#=====#
# TAKE NA ATOMS OUT AND READ BACK IN AFTER SUPERMOL SEPARATION #
#=====#
na_list = []
# Check for sodium ions in supermolecule and get out.
na_in_supermol = False
print("Check for Na in supermol.")
na_in_supermol = any(atom.symbol == "Na" for atom in supermol.atoms)
while na_in_supermol:
    # find next item index with symbol Na in supermol atoms
    idx = next((i for i, atom in enumerate(supermol.atoms) if atom.symbol == 'Na'))
    # remove that item from list and add to na list
    na_list.append(supermol.atoms.pop(idx))
    # still na in supermol?
    na_in_supermol = any(atom.symbol == "Na" for atom in supermol.atoms)
    print("Still Na in supermol? " , na_in_supermol)
# Reset bonds and separate fragments
supermol.guess_bonds()
frags = supermol.separate()

```

---

**Listing A.11:** PyADF input code for pFTAA Part 2.

```
# Readd sodium ions to supermol
for na in na_list:
    na.bonds = []
    supermol.add_atom(na)

for na in na_list:
    mol = OBFreeMolecule()
    mol.add_atom(na)
    mol.delete_all_bonds()
    frags.append(mol)

#=====#
# IDENTIFY PFTAA (must be first) AND SET CHARGE #
#=====#
total_charge = 0
# set charge of env and add up total charge
for i, frag in enumerate(frags):
    frag.set_charge(frag_charge[frag.get_formula()])
    print("DICT ", frag_charge[frag.get_formula()])
    total_charge += int(frag.get_charge())
    print("FRAG", frag.get_formula(), " CHARGE: ", frag.get_charge())
    frag.write(os.path.join(pyadfenv.outdir, 'frag'+str(i)+'.xyz'))
# set total charge
supermol.set_charge(total_charge)
print("TOTAL ", supermol.get_charge())
# identify pftaa
pftaa = frags[0]
pftaa.write(os.path.join(pyadfenv.outdir, 'pftaa.xyz'))
pftaa.set_symmetry('NOSYM')

#=====#
# SUPER MOL GRID #
#=====#
settings = adfsettings(functional=main_functional, zlmfit=True)
settings.set_integration(6.0, becke='good')
settings.set_ncycles(1) # do not converge SCF, we only need the grid
settings.set_save_tapes([21, 10])
res_super_adf = adfsinglepointjob(supermol, basis='DZP', settings=settings,
                                options=['TOTALENERGY', 'NumericalQuality normal']).run()
agrid = adfgrid(res_super_adf)
num_calc -= 1

#=====#
# SINGLE POINT ON ADF GRID IN DALTON #
#=====#
env_frags_res = fragmentlist()
updated_frags = fragmentlist()
```

---



**Listing A.12:** PyADF input code for pFTAA Part 3.

---

```

#####
# MAIN FRAGMENT                                #
#####
# start dalton calculation
settings = daltonsettings(method='DFT',functional=main_functional)
settings.set_memory(dalton_mem)
if response_sp:
    response = ['*LINEAR','.SINGLE RESIDUE','.DIPLN','.ROOTS','5',
               '.NSTART','10','.PRINT','4']
    res_pftaa = daltonsinglepointjob(pftaa, dalton_basis, settings=settings,
                                     response=response).run()
else:
    res_pftaa = daltonsinglepointjob(pftaa, dalton_basis, settings=settings).run()
num_calc -= 1
print(num_calc, " calculations left.")
updated_frags.append(fragment(res_pftaa, pftaa,isfrozen=False))
#####
# ENVIRONMENT FRAGMENTS                        #
#####
settings = daltonsettings(method='DFT', functional=env_functional)
settings.set_memory(dalton_mem)
for frag in frags:
    if not frag == frags[0]:
        frag.set_symmetry('NOSYM')
        frag_res = daltonsinglepointjob(frag, dalton_basis, settings=settings).run()
        num_calc -= 1
        print(num_calc, " calculations left.")
        env_frags_res.append(frag_res)
        updated_frags.append(fragment(frag_res, frag,
                                     isfrozen=True,fdeoptions={"RELAXsave":""}))
#####
# PyEmbed part                                #
# calculate FDE embedding potential with PyEmbed and import it in Dirac #
#####
dens_pftaa = res_pftaa.get_density(grid=agrid, order=2)
pot_pftaa = res_pftaa.get_potential(grid=agrid, pot='elstat')
pot_pftaa.get_xyzwvfile(filename='PNA-POT')
dens_env = None
pot_env = None
old_dens = None
for res in env_frags_res:
    dens_f = res.get_density(grid=agrid, order=2)
    pot_f = res.get_potential(grid=agrid, pot='elstat')
    if dens_env is None:
        dens_env = dens_f
    else:
        dens_env = dens_env + dens_f

```

---

**Listing A.13:** PyADF input code for pFTAA Part 4.

---

```

if pot_env is None:
    pot_env = pot_f
else:
    pot_env = pot_env + pot_f
#####
# EMB POT SETTINGS:                                     #
#####
embed_settings = EmbedXCFunSettings()
embed_settings.set_fun_nad_xc({'BP86': 1.0})
embed_settings.set_fun_nad_kin({'PW91k': 1.0})
embed_settings.show_functionals()
embed_eval = EmbedXCFunEvaluator(settings=embed_settings)
embpot_pftaa = None
#CREATE EMBPOT FOR MAIN FRAGMENT FROM DENSITIES AND POTENTIAL
embpot_pftaa = embed_eval.get_emb_pot(dens_pftaa, dens_env, pot_env)
embpot_pftaa.get_xyzvwfile(filename='EMBPOT-PNA-FINAL')
print ('ENTER FDE CYCLE')

#####
# Starting FDE cycles                                     #
#####
for cycle in range (1,cycles+1):
    #SP CALC FOR MAIN FRAG WITH EMBPOT FOR MAIN FRAG
    settings = daltonsettings(method='DFT',functional=main_functional)
    settings.set_memory(dalton_mem)
    res_pftaa = daltonsinglepointjob(pftaa, dalton_basis, settings=settings,
                                     fdein=embpot_pftaa).run()

    num_calc -= 1
    print(num_calc, " calculations left.")
    # update active dens and put it in fraglist, make all frozen for next loop
    for f in updated_frags:
        if f.get_total_molecule() == pftaa:
            f.results = res_pftaa
            f.isfrozen = True

#####
# loop over env fragments, unfreeze one at a time,      #
# create new emb pot and do sp calculation with new embpot, freeze      #
#####
settings = daltonsettings(method='DFT',functional=env_functional)
settings.set_memory(dalton_mem)
# IF FRAG IS TO BE RELAXED, UNFREEZE, THEN GET DATA FOR EMBPOT IN LOOP
# CREATE EMBPOT_FRAG AND DO SP CALC FOR FRAG
for f in updated_frags:
    if f.has_fdeoption("RELAXsave"):
        f.isfrozen = False # unfreeze -> new active fragment
        dens_main = f.results.get_density(grid=agrid, order=2) # get data of main frag

```

---

**Listing A.14:** PyADF input code for pFTAA Part 5.

```

dens_env = None
pot_env = None
settings = daltonsettings(method='DFT',functional=env_functional)
settings.set_memory(dalton_mem)
# check out embpot data for all frags not active in this cycle
for frag in updated_frags:
    if not frag == f:
        dens_f = frag.results.get_density(grid=agrid, order=2)
        pot_f = frag.results.get_potential(grid=agrid, pot='elstat')
        if dens_env is None:
            dens_env = dens_f
        else:
            dens_env = dens_env + dens_f
        if pot_env is None:
            pot_env = pot_f
        else:
            pot_env = pot_env + pot_f
    embpot_f = embed_eval.get_emb_pot(dens_main, dens_env, pot_env) # get frag embpot
    print ('got embpot data')
# update results for fragment in list
f.results = daltonsinglepointjob(f.get_total_molecule(), dalton_basis,
                                settings=settings, fdein=embpot_f).run()

num_calc -= 1
print(num_calc, " calculations left.")
f.isfrozen = True # freeze for next cycle
print ('1 ENV FRAGMENT IN CYCLE DONE')

#####
#create embpot for pftaa #
#####
# none so they don't add up to the existing ones
dens_env = None
pot_env = None
res_pftaa = None
embpot_pftaa = None
#####
#get necessary densities and potential for emb pot #
#####
for f in updated_frags:
    if not f.has_fdeoption("RELAXsave"):
        res_pftaa = f.results
        f.isfrozen = False
    if f.has_fdeoption("RELAXsave"):
        dens_f = f.results.get_density(grid=agrid, order=2)
        pot_f = f.results.get_potential(grid=agrid, pot='elstat')
        if dens_env is None:
            dens_env = dens_f
        else:
            dens_env = dens_env + dens_f

```

**Listing A.15:** PyADF input code for pFTAA Part 6.

---

```
    if pot_env is None:
        pot_env = pot_f
    else:
        pot_env = pot_env + pot_f

dens_pftaa = res_pftaa.get_density(grid=agrid, order=2)
#####
# CREATE EMBPOT FOR MAIN FRAG #
#####
embpot_pftaa = embed_eval.get_emb_pot(dens_pftaa, dens_env, pot_env)
for f in updated_frags: # unfreeze main fragment
    if not f.has_fdeoption("RELAXsave"):
        f.isfrozen = False
    print ("====CYCLE ", cycle, " DONE.====")

#####
# End of cycles #
#####
# Final calculation for main fragment #
#####

settings = daltonsettings(method='DFT',functional=main_functional)
settings.set_memory(dalton_mem)
response = ['*LINEAR', '.SINGLE RESIDUE', '.DIPLN', '.ROOTS', '5',
            '.NSTART', '10', '.PRINT', '4']
results = daltonsinglepointjob(pftaa, dalton_basis, settings=settings,
                               fdein=embpot_pftaa, response=response).run()
```

---

**Listing A.16:** Charge dictionary for pFTAA "Frag\_Dict.py"

---

```
frag_charge={'01H2' : '0' ,
            'Na1' : '1' ,
            'S508C26H12' : '-4' ,
            'S509Na1C26H13' : '-3' ,
            '02H6C6N2' : '0'
            }
```

---

## A.2. Numerical Comparison

### A.2.1. Functional Evaluation

#### Supermolecular Reference Calculation for *p*N<sub>A</sub> in a 3 Å Environment with a aug-cc-pVDZ Basis Set

**Table A.1.:** Excitation energies and oscillator strengths and their average obtained for the different snapshots of *p*N<sub>A</sub> in a 3 Å water environment calculated with CAM-B3LYP/aug-cc-pVDZ. On the left 100% Fock exchange was employed in the calculations, on the right the default CAM-B3LYP was employed. The presented values are presented for the transition with the highest oscillator strength.

Snap.	Trans.	$E^{\text{ex}} / \text{eV}$	$f_{\text{L}}$	Snap.	Trans.	$E^{\text{ex}} / \text{eV}$	$f_{\text{L}}$
1	1	3.82	0.2352	1	1	3.48	0.2109
2	1	3.42	0.1727	2	1	3.09	0.1598
3	1	3.84	0.4503	3	1	3.56	0.4101
4	1	3.88	0.2944	4	1	3.54	0.2729
5	1	4.06	0.1341	5	2	3.94	0.2749
6	1	3.99	0.2885	6	1	3.64	0.2666
7	1	4.51	0.3695	7	1	4.17	0.3529
∅	-	3.96	0.2972	∅	-	3.63	0.2783

#### Supermolecular Reference Calculation for *p*N<sub>A</sub> in a 3 Å Environment with a aug-cc-pVTZ Basis Set

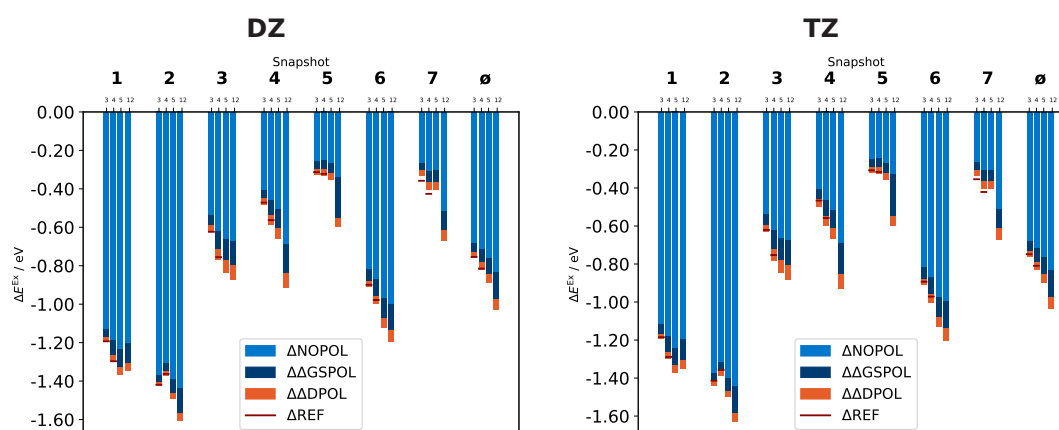
**Table A.2.:** Excitation energies and oscillator strengths and their average obtained for the different snapshots of *p*N<sub>A</sub> in a 3 Å water environment calculated with CAM-B3LYP/aug-cc-pVTZ. On the left 100% Fock exchange was employed in the calculations, on the right the default CAM-B3LYP was employed. The presented values are presented for the transition with the highest oscillator strength.

Snap.	Trans.	$E^{\text{ex}} / \text{eV}$	$f_{\text{L}}$	Snap.	Trans.	$E^{\text{ex}} / \text{eV}$	$f_{\text{L}}$
1	1	3.84	0.2371	1	1	3.50	0.2125
2	1	3.44	0.1735	2	1	3.11	0.1606
3	1	3.86	0.4496	3	1	3.57	0.4102
4	1	3.89	0.2940	4	1	3.55	0.2729
5	2	4.26	0.2619	5	2	3.95	0.2689
6	1	4.00	0.2899	6	1	3.65	0.2680
7	1	4.52	0.3693	7	1	4.18	0.3534
∅	-	3.97	0.2965	∅	-	3.64	0.2781

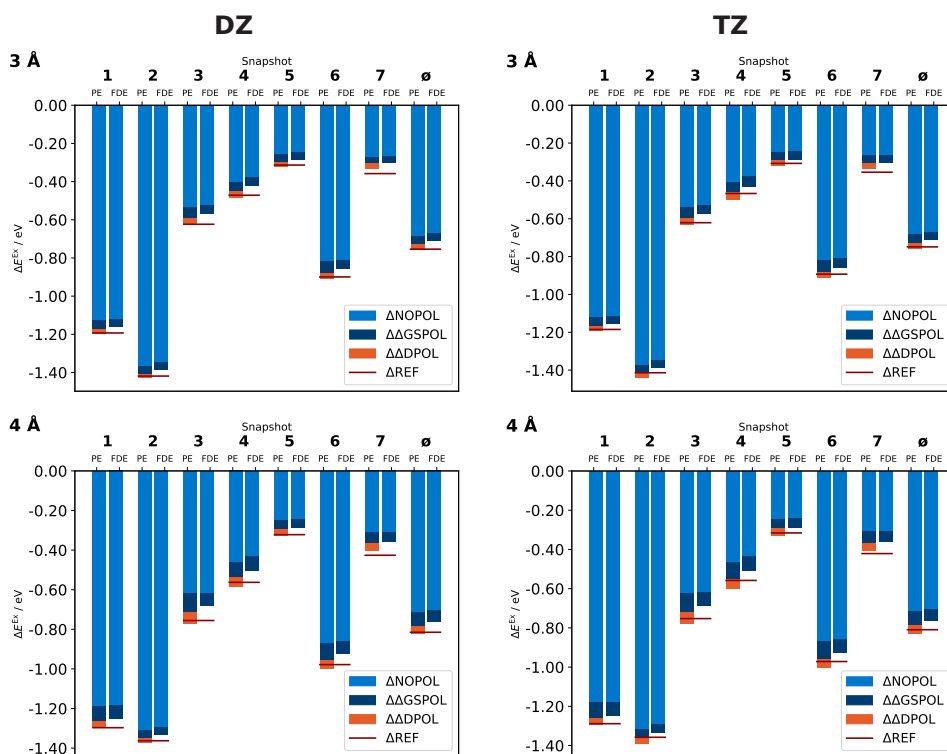
## A.2.2. *para*-Nitroaniline

The calculated excitation energies (in eV) and oscillator strengths for the models NOPOL, GSPOL and PE DPOL of the strongest  $\pi \rightarrow \pi^*$  transition are given. Calculations were done with PE and FDE embedding models for different configurations of *p*NA and *p*FTAA. The definition of the different shifts ( $\Delta$ NOPOL,  $\Delta$ GSPOL,  $\Delta$ DPOL,  $\Delta$  DPOL+EEF, and  $\Delta$ REF) as well as contributions ( $\Delta\Delta$ NOPOL,  $\Delta\Delta$ GSPOL,  $\Delta\Delta$ DPOL,  $\Delta\Delta$ DPOL+EEF, and  $\Delta\Delta$  EEF) are described in Fig. 4.2.4.

### A.2.2.1. Excitation Energies and $\mathcal{S}$ -Shifts for *para*-Nitroaniline



**Figure A.2.:** Contributions to the  $\mathcal{S}$ -shifts and their average for different configurations of *p*NA in 3, 4, 5 and 12 Å environments of water obtained from a MD simulation and subsequently calculated in a PE framework and different orders of polarization contributions obtained in calculations with aug-cc-pVDZ and aug-cc-pVTZ basis sets. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.



**Figure A.3.:** Contributions to the  $S$ -shifts and their average for different configurations of  $pNA$  in 3 and 4 Å water environments of water obtained from a MD simulation and subsequently calculated in a PE and FDE framework and different orders of polarization contributions obtained in calculations with aug-cc-pVDZ and aug-cc-pVTZ basis sets. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. J. *Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

### Calculations in a 3 Å water environment

**Table A.3.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689-5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{FDE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.56	-1.12	3.52	-0.04	-1.16	-1.16	3.48	-1.19
Snap 2	4.51	3.17	-1.35	3.13	-0.04	-1.38	-1.38	3.09	-1.42
Snap 3	4.19	3.66	-0.52	3.62	-0.05	-0.57	-0.57	3.56	-0.62
Snap 4	4.01	3.64	-0.37	3.59	-0.05	-0.42	-0.42	3.54	-0.47
Snap 5	4.25	4.01	-0.24	3.97	-0.04	-0.29	-0.29	3.94	-0.31
Snap 6	4.54	3.73	-0.81	3.68	-0.05	-0.86	-0.86	3.64	-0.90
Snap 7	4.53	4.26	-0.27	4.23	-0.03	-0.30	-0.30	4.17	-0.36
$\emptyset$	4.39	3.72	-0.67	3.67	-0.04	-0.71	-0.71	3.63	-0.75

**Table A.4.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689-5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.55	-1.13	3.50	-0.04	3.48	-0.02	-1.17	-1.19	3.48	-1.19
Snap 2	4.51	3.14	-1.37	3.11	-0.04	3.09	-0.02	-1.41	-1.43	3.09	-1.42
Snap 3	4.19	3.65	-0.53	3.60	-0.06	3.56	-0.04	-0.59	-0.63	3.56	-0.62
Snap 4	4.01	3.61	-0.40	3.56	-0.04	3.53	-0.04	-0.45	-0.49	3.54	-0.47
Snap 5	4.25	3.99	-0.26	3.95	-0.04	3.93	-0.03	-0.30	-0.32	3.94	-0.31
Snap 6	4.54	3.72	-0.82	3.66	-0.06	3.63	-0.03	-0.88	-0.91	3.64	-0.90
Snap 7	4.53	4.26	-0.27	4.22	-0.04	4.19	-0.03	-0.31	-0.33	4.17	-0.36
$\emptyset$	4.39	3.70	-0.68	3.66	-0.05	3.63	-0.03	-0.73	-0.76	3.63	-0.75



**Table A.5.:** Excitation energies and contributions from the different polarization models to the  $\mathcal{S}$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVTZ quality and a 3 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{FDE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.57	-1.11	3.53	-0.04	-1.16	-1.16	3.50	-1.19
Snap 2	4.52	3.18	-1.34	3.14	-0.04	-1.39	-1.39	3.11	-1.41
Snap 3	4.19	3.67	-0.52	3.62	-0.05	-0.57	-0.57	3.57	-0.62
Snap 4	4.02	3.64	-0.37	3.59	-0.05	-0.43	-0.43	3.55	-0.47
Snap 5	4.26	4.01	-0.24	3.97	-0.04	-0.29	-0.29	3.95	-0.31
Snap 6	4.54	3.74	-0.80	3.68	-0.05	-0.86	-0.86	3.65	-0.89
Snap 7	4.53	4.27	-0.26	4.23	-0.04	-0.30	-0.30	4.18	-0.35
$\emptyset$	4.39	3.73	-0.67	3.68	-0.05	-0.71	-0.71	3.64	-0.75

**Table A.6.:** Excitation energies and contributions from the different polarization models to the  $\mathcal{S}$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVTZ quality and a 3 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.56	-1.12	3.52	-0.05	3.49	-0.02	-1.17	-1.19	3.50	-1.19
Snap 2	4.52	3.15	-1.37	3.10	-0.04	3.08	-0.02	-1.42	-1.44	3.11	-1.41
Snap 3	4.19	3.65	-0.54	3.60	-0.06	3.56	-0.04	-0.60	-0.63	3.57	-0.62
Snap 4	4.02	3.61	-0.41	3.56	-0.05	3.52	-0.04	-0.46	-0.50	3.55	-0.47
Snap 5	4.26	4.00	-0.25	3.96	-0.04	3.94	-0.03	-0.29	-0.32	3.95	-0.31
Snap 6	4.54	3.72	-0.82	3.66	-0.07	3.63	-0.03	-0.88	-0.91	3.65	-0.89
Snap 7	4.53	4.26	-0.27	4.22	-0.04	4.20	-0.03	-0.31	-0.33	4.18	-0.35
$\emptyset$	4.39	3.71	-0.68	3.66	-0.05	3.63	-0.03	-0.73	-0.76	3.64	-0.75

### Calculations in a 4 Å water environment

**Table A.7.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVDZ quality and a 4 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689-5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{FDE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.49	-1.18	3.43	-0.07	-1.25	-1.25	3.38	-1.30
Snap 2	4.51	3.22	-1.29	3.18	-0.04	-1.33	-1.33	3.15	-1.36
Snap 3	4.19	3.57	-0.61	3.50	-0.07	-0.68	-0.68	3.43	-0.76
Snap 4	4.01	3.58	-0.43	3.51	-0.08	-0.50	-0.50	3.45	-0.56
Snap 5	4.25	4.01	-0.24	3.96	-0.05	-0.29	-0.29	3.93	-0.32
Snap 6	4.54	3.68	-0.86	3.61	-0.07	-0.92	-0.92	3.56	-0.98
Snap 7	4.53	4.22	-0.31	4.17	-0.05	-0.36	-0.36	4.10	-0.43
$\emptyset$	4.39	3.68	-0.70	3.62	-0.06	-0.76	-0.76	3.57	-0.81

**Table A.8.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVDZ quality and a 4 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689-5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.49	-1.19	3.41	-0.08	3.38	-0.03	-1.26	-1.30	3.38	-1.30
Snap 2	4.51	3.20	-1.31	3.16	-0.04	3.14	-0.02	-1.35	-1.38	3.15	-1.36
Snap 3	4.19	3.57	-0.62	3.47	-0.09	3.41	-0.06	-0.71	-0.77	3.43	-0.76
Snap 4	4.01	3.55	-0.46	3.47	-0.08	3.42	-0.05	-0.54	-0.59	3.45	-0.56
Snap 5	4.25	4.00	-0.25	3.96	-0.04	3.92	-0.04	-0.30	-0.33	3.93	-0.32
Snap 6	4.54	3.67	-0.87	3.58	-0.09	3.54	-0.04	-0.96	-1.00	3.56	-0.98
Snap 7	4.53	4.22	-0.31	4.16	-0.06	4.12	-0.04	-0.37	-0.41	4.10	-0.43
$\emptyset$	4.39	3.67	-0.71	3.60	-0.07	3.56	-0.04	-0.78	-0.82	3.57	-0.81

**Table A.9.:** Excitation energies and contributions from the different polarization models to the  $\mathcal{S}$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVTZ quality and a 4 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{FDE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.51	-1.18	3.44	-0.07	-1.25	-1.25	3.40	-1.29
Snap 2	4.52	3.23	-1.29	3.19	-0.05	-1.34	-1.34	3.16	-1.36
Snap 3	4.19	3.58	-0.62	3.50	-0.07	-0.69	-0.69	3.44	-0.75
Snap 4	4.02	3.59	-0.43	3.51	-0.08	-0.51	-0.51	3.46	-0.56
Snap 5	4.26	4.02	-0.24	3.97	-0.05	-0.29	-0.29	3.94	-0.32
Snap 6	4.54	3.69	-0.85	3.61	-0.07	-0.93	-0.93	3.57	-0.97
Snap 7	4.53	4.23	-0.31	4.17	-0.05	-0.36	-0.36	4.11	-0.42
$\emptyset$	4.39	3.69	-0.70	3.63	-0.06	-0.77	-0.77	3.58	-0.81

**Table A.10.:** Excitation energies and contributions from the different polarization models to the  $\mathcal{S}$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVTZ quality and a 4 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	4.68	3.50	-1.18	3.42	-0.08	3.39	-0.03	-1.26	-1.30	3.40	-1.29
Snap 2	4.52	3.21	-1.32	3.16	-0.05	3.13	-0.03	-1.36	-1.39	3.16	-1.36
Snap 3	4.19	3.57	-0.62	3.47	-0.10	3.41	-0.06	-0.72	-0.78	3.44	-0.75
Snap 4	4.02	3.55	-0.46	3.47	-0.09	3.42	-0.05	-0.55	-0.60	3.46	-0.56
Snap 5	4.26	4.01	-0.24	3.96	-0.05	3.93	-0.04	-0.29	-0.33	3.94	-0.32
Snap 6	4.54	3.67	-0.87	3.58	-0.09	3.54	-0.04	-0.96	-1.00	3.57	-0.97
Snap 7	4.53	4.23	-0.31	4.17	-0.06	4.13	-0.04	-0.37	-0.41	4.11	-0.42
$\emptyset$	4.39	3.68	-0.71	3.60	-0.07	3.56	-0.04	-0.79	-0.83	3.58	-0.81

### PE calculations in a 5 Å water environment

**Table A.11.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVDZ quality and a 5 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	4.68	3.44	-1.23	3.35	-0.09	3.31	-0.04	-1.33	-1.36
Snap 2	4.51	3.12	-1.39	3.05	-0.07	3.02	-0.03	-1.46	-1.49
Snap 3	4.19	3.53	-0.66	3.41	-0.11	3.35	-0.06	-0.77	-0.84
Snap 4	4.01	3.50	-0.51	3.40	-0.10	3.35	-0.05	-0.61	-0.66
Snap 5	4.25	3.99	-0.27	3.93	-0.05	3.90	-0.03	-0.32	-0.36
Snap 6	4.54	3.57	-0.97	3.46	-0.11	3.41	-0.05	-1.07	-1.13
Snap 7	4.53	4.22	-0.30	4.16	-0.06	4.12	-0.04	-0.36	-0.40
$\emptyset$	4.39	3.62	-0.76	3.54	-0.09	3.49	-0.04	-0.85	-0.89

**Table A.12.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $p$ NA in aug-cc-pVTZ quality and a 5 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	4.68	3.44	-1.24	3.35	-0.09	3.31	-0.04	-1.34	-1.37
Snap 2	4.52	3.12	-1.40	3.05	-0.07	3.02	-0.03	-1.47	-1.50
Snap 3	4.19	3.53	-0.67	3.41	-0.11	3.35	-0.06	-0.78	-0.84
Snap 4	4.02	3.50	-0.51	3.40	-0.10	3.35	-0.05	-0.61	-0.67
Snap 5	4.26	3.99	-0.27	3.93	-0.05	3.90	-0.03	-0.32	-0.36
Snap 6	4.54	3.57	-0.97	3.46	-0.11	3.41	-0.05	-1.08	-1.13
Snap 7	4.53	4.22	-0.31	4.16	-0.06	4.12	-0.04	-0.37	-0.41
$\emptyset$	4.39	3.62	-0.77	3.54	-0.09	3.49	-0.04	-0.85	-0.90

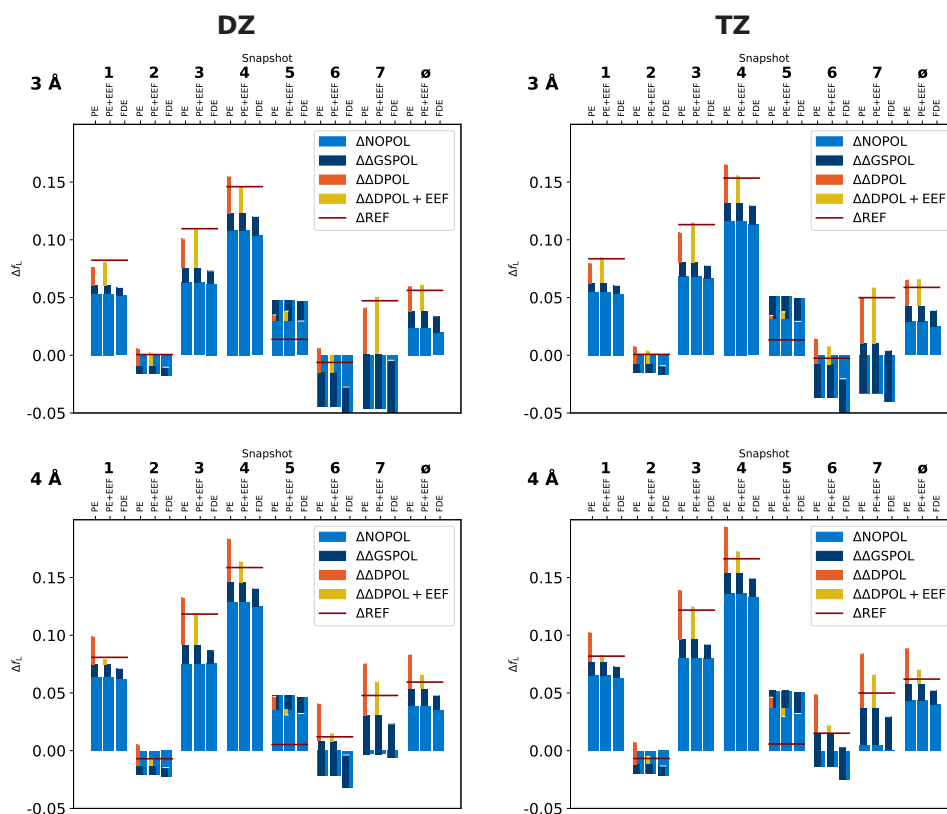
### PE calculations in a 12 Å water environment

**Table A.13.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $pNA$  in aug-cc-pVDZ quality and a 12 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	4.68	3.47	-1.21	3.37	-0.10	3.33	-0.04	-1.31	-1.35
Snap 2	4.51	3.08	-1.44	2.94	-0.13	2.90	-0.04	-1.57	-1.61
Snap 3	4.19	3.51	-0.67	3.39	-0.12	3.31	-0.08	-0.79	-0.87
Snap 4	4.01	3.33	-0.69	3.17	-0.15	3.10	-0.07	-0.84	-0.91
Snap 5	4.25	3.91	-0.34	3.70	-0.22	3.65	-0.05	-0.55	-0.60
Snap 6	4.54	3.54	-1.00	3.40	-0.13	3.34	-0.06	-1.13	-1.19
Snap 7	4.53	4.01	-0.52	3.91	-0.10	3.86	-0.06	-0.61	-0.67
$\emptyset$	4.39	3.55	-0.84	3.41	-0.14	3.36	-0.06	-0.97	-1.03

**Table A.14.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of  $pNA$  in aug-cc-pVTZ quality and a 12 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	4.68	3.49	-1.20	3.38	-0.11	3.33	-0.04	-1.31	-1.35
Snap 2	4.52	3.08	-1.44	2.94	-0.14	2.89	-0.04	-1.58	-1.63
Snap 3	4.19	3.52	-0.67	3.39	-0.13	3.31	-0.08	-0.80	-0.88
Snap 4	4.02	3.33	-0.69	3.16	-0.16	3.09	-0.08	-0.85	-0.93
Snap 5	4.26	3.93	-0.33	3.71	-0.22	3.66	-0.05	-0.55	-0.60
Snap 6	4.54	3.54	-1.00	3.40	-0.14	3.34	-0.06	-1.14	-1.20
Snap 7	4.53	4.02	-0.51	3.92	-0.10	3.86	-0.06	-0.61	-0.67
$\emptyset$	4.39	3.56	-0.84	3.41	-0.14	3.35	-0.06	-0.98	-1.04

A.2.2.2. Oscillator Strengths and  $\mathcal{F}$ -shifts for *para*-nitroaniline

**Figure A.4.:** Contributions to the  $\mathcal{F}$ -shifts and their average for different configurations of *pNA* in 3 and 4 Å water environments of water obtained from a MD simulation and subsequently calculated in a PE and FDE framework, different orders of polarization contributions and effective external field (EEF) effects obtained in calculations with aug-cc-pVDZ and aug-cc-pVTZ basis sets. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

### Calculations in a 3 Å water environment

**Table A.15.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pNA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (FDE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	0.129	0.180	0.052	0.187	0.007	0.059	0.059	0.211	0.082
Snap 2	0.159	0.142	-0.017	0.149	0.007	-0.010	-0.010	0.160	0.001
Snap 3	0.301	0.362	0.062	0.374	0.011	0.073	0.073	0.410	0.110
Snap 4	0.127	0.230	0.104	0.247	0.017	0.120	0.120	0.273	0.146
Snap 5	0.261	0.307	0.046	0.291	-0.017	0.030	0.030	0.275	0.014
Snap 6	0.273	0.215	-0.058	0.245	0.030	-0.027	-0.027	0.267	-0.006
Snap 7	0.306	0.255	-0.051	0.301	0.046	-0.004	-0.004	0.353	0.047
$\emptyset$	0.222	0.242	0.020	0.256	0.015	0.034	0.034	0.278	0.056

**Table A.16.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pNA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	DPOL+EEF	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ DPOL+EEF	REF	$\Delta$ REF
Snap 1	0.129	0.182	0.053	0.189	0.007	0.205	0.016	0.210	0.005	0.060	0.077	0.081	0.211	0.082
Snap 2	0.159	0.143	-0.016	0.150	0.007	0.165	0.015	0.162	-0.004	-0.009	0.006	0.003	0.160	0.001
Snap 3	0.301	0.363	0.063	0.376	0.013	0.401	0.025	0.410	0.009	0.076	0.101	0.109	0.410	0.110
Snap 4	0.127	0.235	0.108	0.250	0.015	0.282	0.032	0.273	-0.009	0.123	0.155	0.146	0.273	0.146
Snap 5	0.261	0.309	0.048	0.291	-0.017	0.296	0.005	0.299	0.003	0.030	0.035	0.038	0.275	0.014
Snap 6	0.273	0.228	-0.045	0.258	0.030	0.279	0.022	0.273	-0.006	-0.015	0.007	0.000	0.267	-0.006
Snap 7	0.306	0.259	-0.046	0.307	0.047	0.347	0.040	0.357	0.010	0.001	0.041	0.051	0.353	0.047
$\emptyset$	0.222	0.246	0.024	0.260	0.014	0.282	0.022	0.283	0.001	0.038	0.060	0.061	0.278	0.056

**Table A.17.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of *p*NA in aug-cc-pVTZ quality and a 3 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (FDE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	0.129	0.182	0.053	0.189	0.008	0.061	0.061	0.213	0.084
Snap 2	0.160	0.143	-0.017	0.151	0.008	-0.009	-0.009	0.161	0.001
Snap 3	0.297	0.363	0.066	0.375	0.012	0.078	0.078	0.410	0.113
Snap 4	0.119	0.233	0.113	0.249	0.017	0.130	0.130	0.273	0.153
Snap 5	0.256	0.305	0.050	0.285	-0.020	0.029	0.029	0.269	0.013
Snap 6	0.271	0.219	-0.051	0.250	0.031	-0.020	-0.020	0.268	-0.003
Snap 7	0.304	0.263	-0.040	0.308	0.045	0.004	0.004	0.353	0.050
$\emptyset$	0.219	0.244	0.025	0.258	0.014	0.039	0.039	0.278	0.059

**Table A.18.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of *p*NA in aug-cc-pVTZ quality and a 3 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	DPOL+EEF	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ DPOL+EEF	REF	$\Delta$ REF
Snap 1	0.129	0.184	0.055	0.191	0.008	0.209	0.017	0.213	0.004	0.063	0.080	0.084	0.213	0.084
Snap 2	0.160	0.145	-0.015	0.152	0.007	0.168	0.016	0.164	-0.004	-0.008	0.008	0.004	0.161	0.001
Snap 3	0.297	0.365	0.068	0.377	0.012	0.404	0.026	0.412	0.008	0.080	0.106	0.115	0.410	0.113
Snap 4	0.119	0.236	0.116	0.251	0.015	0.285	0.034	0.275	-0.010	0.132	0.166	0.156	0.273	0.153
Snap 5	0.256	0.307	0.051	0.287	-0.020	0.290	0.003	0.293	0.003	0.031	0.034	0.037	0.269	0.013
Snap 6	0.271	0.234	-0.037	0.263	0.029	0.285	0.022	0.279	-0.006	-0.008	0.015	0.008	0.268	-0.003
Snap 7	0.304	0.271	-0.033	0.314	0.043	0.354	0.040	0.363	0.009	0.010	0.050	0.059	0.353	0.050
$\emptyset$	0.219	0.249	0.029	0.262	0.014	0.285	0.023	0.286	0.001	0.043	0.066	0.066	0.278	0.059



### Calculations in a 4 Å water environment

**Table A.19.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pNA in aug-cc-pVDZ quality and a 4 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (FDE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	0.129	0.190	0.062	0.200	0.010	0.071	0.071	0.209	0.081
Snap 2	0.159	0.137	-0.023	0.145	0.008	-0.015	-0.015	0.152	-0.007
Snap 3	0.301	0.376	0.076	0.388	0.012	0.088	0.088	0.419	0.118
Snap 4	0.127	0.252	0.125	0.268	0.016	0.141	0.141	0.286	0.159
Snap 5	0.261	0.308	0.047	0.293	-0.014	0.032	0.032	0.266	0.005
Snap 6	0.273	0.241	-0.032	0.269	0.028	-0.004	-0.004	0.285	0.012
Snap 7	0.306	0.299	-0.006	0.329	0.030	0.023	0.023	0.354	0.048
$\emptyset$	0.222	0.258	0.035	0.270	0.013	0.048	0.048	0.282	0.059

**Table A.20.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pNA in aug-cc-pVDZ quality and a 4 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	DPOL+EEF	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ DPOL+EEF	REF	$\Delta$ REF
Snap 1	0.129	0.192	0.064	0.203	0.011	0.228	0.024	0.208	-0.020	0.075	0.099	0.079	0.209	0.081
Snap 2	0.159	0.139	-0.021	0.146	0.007	0.164	0.018	0.152	-0.012	-0.013	0.005	-0.007	0.152	-0.007
Snap 3	0.301	0.376	0.075	0.392	0.016	0.433	0.041	0.420	-0.013	0.092	0.133	0.119	0.419	0.118
Snap 4	0.127	0.256	0.129	0.273	0.017	0.311	0.038	0.291	-0.020	0.146	0.184	0.164	0.286	0.159
Snap 5	0.261	0.309	0.048	0.297	-0.013	0.308	0.011	0.292	-0.016	0.036	0.047	0.031	0.266	0.005
Snap 6	0.273	0.251	-0.022	0.281	0.030	0.314	0.033	0.288	-0.026	0.008	0.041	0.015	0.285	0.012
Snap 7	0.306	0.303	-0.003	0.336	0.034	0.382	0.045	0.366	-0.016	0.031	0.076	0.060	0.354	0.048
$\emptyset$	0.222	0.261	0.039	0.275	0.015	0.306	0.030	0.288	-0.018	0.053	0.083	0.066	0.282	0.059

**Table A.21.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pNA in aug-cc-pVTZ quality and a 4 Å water environment obtained with a FDE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (FDE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	0.129	0.192	0.063	0.202	0.010	0.073	0.073	0.211	0.082
Snap 2	0.160	0.138	-0.022	0.147	0.009	-0.013	-0.013	0.153	-0.007
Snap 3	0.297	0.377	0.080	0.389	0.013	0.092	0.092	0.419	0.122
Snap 4	0.119	0.253	0.133	0.269	0.016	0.150	0.150	0.286	0.166
Snap 5	0.256	0.306	0.050	0.288	-0.018	0.032	0.032	0.261	0.006
Snap 6	0.271	0.246	-0.025	0.274	0.028	0.003	0.003	0.286	0.015
Snap 7	0.304	0.304	0.001	0.333	0.029	0.030	0.030	0.353	0.050
$\emptyset$	0.219	0.259	0.040	0.272	0.012	0.052	0.052	0.281	0.062

**Table A.22.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pNA in aug-cc-pVTZ quality and a 4 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	DPOL+EEF	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ DPOL+EEF	REF	$\Delta$ REF
Snap 1	0.129	0.194	0.065	0.206	0.011	0.232	0.026	0.211	-0.021	0.077	0.103	0.082	0.211	0.082
Snap 2	0.160	0.140	-0.020	0.148	0.008	0.167	0.019	0.155	-0.012	-0.012	0.008	-0.005	0.153	-0.007
Snap 3	0.297	0.377	0.080	0.394	0.016	0.436	0.043	0.422	-0.014	0.096	0.139	0.125	0.419	0.122
Snap 4	0.119	0.256	0.136	0.274	0.018	0.314	0.040	0.292	-0.022	0.154	0.194	0.173	0.286	0.166
Snap 5	0.256	0.307	0.052	0.293	-0.015	0.302	0.009	0.285	-0.017	0.037	0.046	0.029	0.261	0.006
Snap 6	0.271	0.257	-0.014	0.286	0.029	0.320	0.034	0.293	-0.027	0.015	0.049	0.022	0.286	0.015
Snap 7	0.304	0.309	0.005	0.341	0.032	0.388	0.047	0.370	-0.018	0.037	0.084	0.066	0.353	0.050
$\emptyset$	0.219	0.263	0.044	0.277	0.014	0.308	0.031	0.290	-0.019	0.058	0.089	0.070	0.281	0.062

**PE calculations in aug-cc-pVDZ quality in a 5 Å water environment****Table A.23.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of *p*NA in aug-cc-pVDZ quality and a 5 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	0.129	0.199	0.070	0.211	0.012	0.239	0.028	0.082	0.110
Snap 2	0.159	0.152	-0.007	0.163	0.011	0.185	0.022	0.004	0.026
Snap 3	0.301	0.381	0.081	0.401	0.019	0.446	0.045	0.100	0.146
Snap 4	0.127	0.263	0.136	0.281	0.018	0.322	0.042	0.154	0.195
Snap 5	0.261	0.303	0.041	0.268	-0.035	0.257	-0.011	0.007	-0.004
Snap 6	0.273	0.278	0.005	0.303	0.025	0.341	0.038	0.030	0.069
Snap 7	0.306	0.297	-0.009	0.335	0.038	0.383	0.049	0.029	0.078
$\emptyset$	0.222	0.268	0.045	0.280	0.013	0.311	0.030	0.058	0.088

**Table A.24.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of *p*NA in aug-cc-pVTZ quality and a 5 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	0.129	0.199	0.070	0.211	0.012	0.239	0.028	0.082	0.110
Snap 2	0.160	0.152	-0.008	0.163	0.011	0.185	0.022	0.003	0.025
Snap 3	0.297	0.381	0.084	0.401	0.019	0.446	0.045	0.104	0.149
Snap 4	0.119	0.263	0.144	0.281	0.018	0.322	0.042	0.161	0.203
Snap 5	0.256	0.303	0.047	0.268	-0.035	0.257	-0.011	0.012	0.002
Snap 6	0.271	0.278	0.008	0.303	0.025	0.341	0.038	0.033	0.071
Snap 7	0.304	0.297	-0.007	0.335	0.038	0.383	0.049	0.031	0.080
$\emptyset$	0.219	0.268	0.048	0.280	0.013	0.311	0.030	0.061	0.091

**PE calculations in aug-cc-pVDZ quality in a 12 Å water environment****Table A.25.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of *p*NA in aug-cc-pVDZ quality and a 12 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	0.129	0.194	0.065	0.207	0.013	0.239	0.032	0.079	0.111
Snap 2	0.159	0.158	-0.001	0.177	0.019	0.205	0.027	0.018	0.045
Snap 3	0.301	0.382	0.082	0.403	0.021	0.457	0.054	0.102	0.156
Snap 4	0.127	0.289	0.162	0.311	0.022	0.362	0.050	0.184	0.235
Snap 5	0.261	0.236	-0.026	0.230	-0.005	0.314	0.084	-0.031	0.053
Snap 6	0.273	0.286	0.013	0.313	0.027	0.358	0.045	0.040	0.085
Snap 7	0.306	0.355	0.049	0.365	0.011	0.419	0.053	0.060	0.113
$\emptyset$	0.222	0.271	0.049	0.287	0.015	0.336	0.049	0.065	0.114

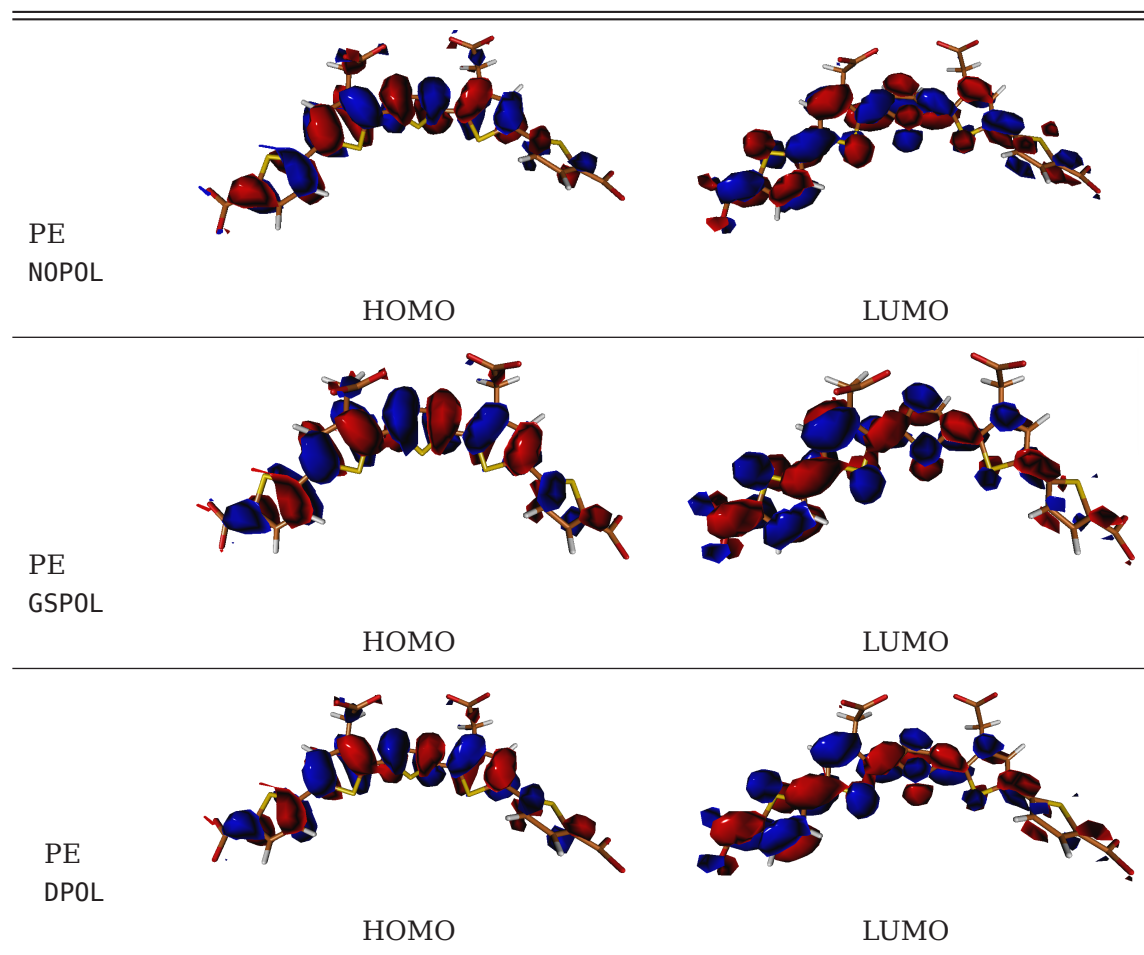
**Table A.26.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of *p*NA in aug-cc-pVTZ quality and a 12 Å water environment obtained with a PE embedding model. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL
Snap 1	0.129	0.196	0.067	0.210	0.014	0.244	0.034	0.081	0.115
Snap 2	0.160	0.160	0.000	0.179	0.020	0.208	0.029	0.020	0.048
Snap 3	0.297	0.383	0.086	0.404	0.021	0.460	0.056	0.107	0.163
Snap 4	0.119	0.289	0.170	0.313	0.023	0.366	0.053	0.193	0.246
Snap 5	0.256	0.232	-0.024	0.237	0.005	0.324	0.087	-0.018	0.068
Snap 6	0.271	0.289	0.018	0.315	0.026	0.362	0.047	0.045	0.091
Snap 7	0.304	0.357	0.053	0.368	0.011	0.424	0.056	0.064	0.120
$\emptyset$	0.219	0.272	0.053	0.289	0.017	0.341	0.052	0.070	0.122

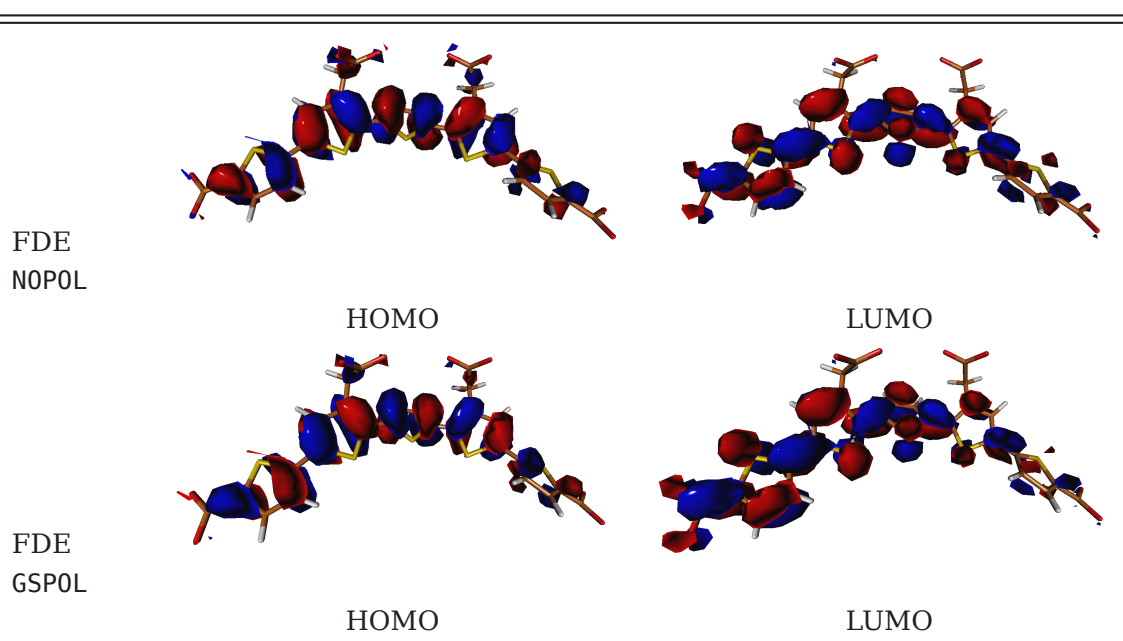
### A.2.3. Pentameric Formyl Thiophene Acetic Acid

**Table A.27.:** Excitation energies, oscillator strengths, and contributions of the strongest transitions of snapshot 1 calculated for different polarization models. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

Model	Trans.	$E^{\text{ex}} / \text{eV}$	$f_{\text{L}}$	Contrib.	Orbitals transition
PE NOPOL	1	3.08	1.221	0.66	$\pi \rightarrow \pi^*$
PE GSPOL	1	3.09	1.280	0.65	$\pi \rightarrow \pi^*$
PE DPOL	1	3.06	1.343	0.65	$\pi \rightarrow \pi^*$
FDE NOPOL	1	3.07	1.195	0.66	$\pi \rightarrow \pi^*$
FDE GSPOL	1	3.09	1.248	0.65	$\pi \rightarrow \pi^*$
Reference	1	3.05	1.2882	0.65	

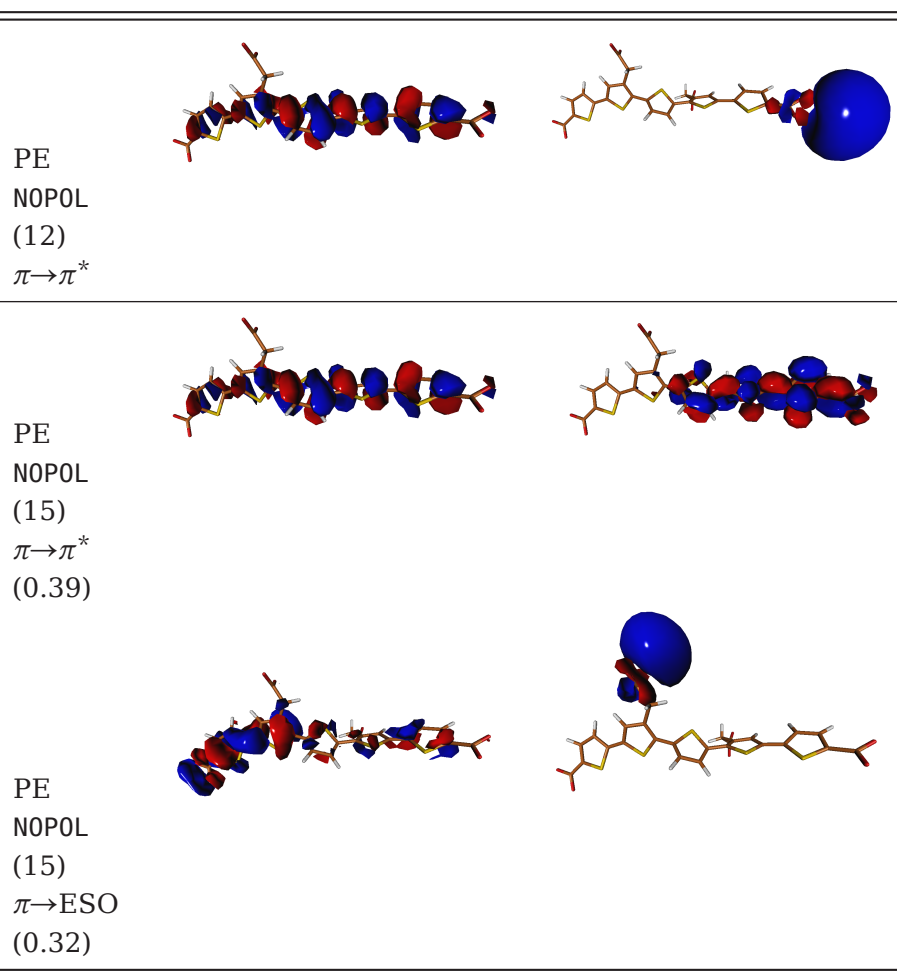


**Figure A.5.:** Orbitals involved in the strongest transition of pFTAA in the different models for snapshot 1 for the PE scheme. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.



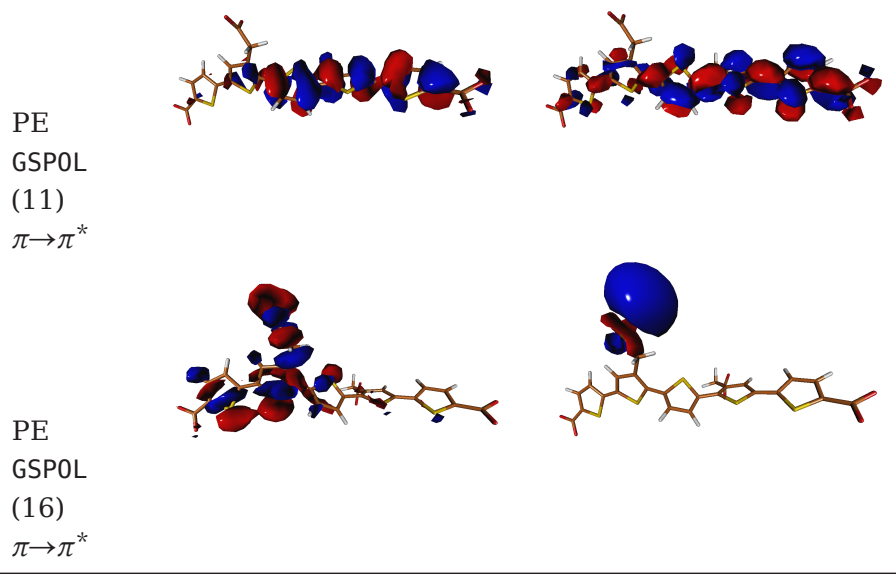
**Figure A.6.:** Orbitals involved in the strongest transition of pFTAA in the different models for snapshot 1 for the FDE scheme. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

### A.2.3.1. Electron-Spill-Out Effect

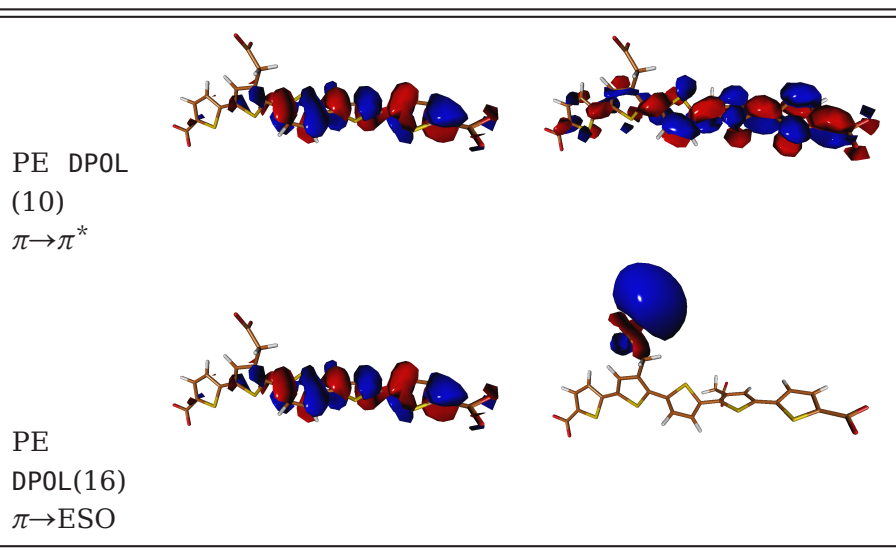


**Figure A.7.:** Orbitals involved in the strongest transitions of snapshot 2 of pFTAA for the PE NOPOL model. Numbers in brackets denote the calculated transition number.





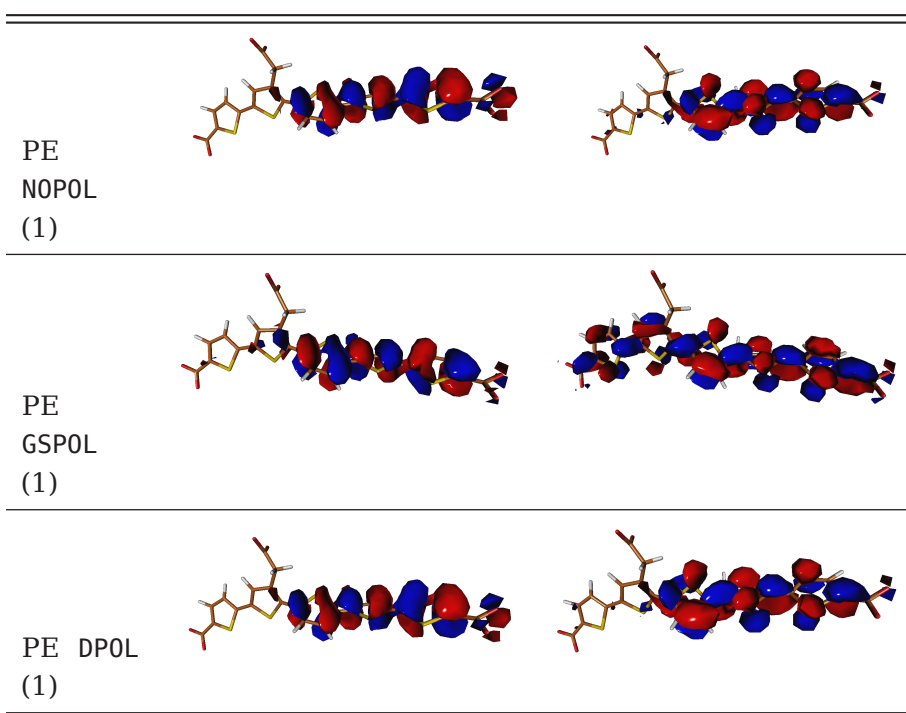
**Figure A.8.:** Orbitals involved in the strongest transitions of snapshot 2 of pFTAA for the PE GSPOL model



**Figure A.9.:** Orbitals involved in the strongest transitions of snapshot 2 of pFTAA for the PE DPOL model

**Table A.28.:** Strongest transitions and identification of the contributing orbitals of the PE calculations for snapshots 2 of pFTAA in a 3 Å water environment and the different polarization models. If more than one high contribution was found, those were given in brackets.

Model	Transition	$E^{\text{ex}}$	$f_{\text{L}}$	Orbitals
PE NOPOL	12	3.02	0.3080	$\pi \rightarrow \text{ESO}$
				$\pi \rightarrow \pi^*$ (0.39)
	15	3.22	0.4151	$\pi \rightarrow \text{ESO}$ (0.32)
				$\pi \rightarrow \text{ESO}$ (0.30)
	16	3.24	0.6427	$\pi \rightarrow \pi^*$
	20	3.72	0.2196	$\pi \rightarrow \pi^*$
PE GSPOL	11	3.22	1.3817	$\pi \rightarrow \pi^*$
	16	3.65	0.1824	$\pi \rightarrow \text{ESO}$ (0.48)
				$\pi \rightarrow \pi^*$ (0.16)
	22	4.01	0.2000	$\pi \rightarrow \pi^*$
				$\pi \rightarrow \pi^*$
PE DPOL	10	3.19	1.4363	$\pi \rightarrow \pi^*$
	16	3.64	0.2248	$\pi \rightarrow \text{ESO}$ (0.44)
				$n \rightarrow \text{ESO}$ (0.21)
				$\pi \rightarrow \pi^*$ (0.46)
	22	3.99	0.2013	$\pi \rightarrow \pi^*$ (0.33)
				$\pi \rightarrow \pi^*$ (0.24)
FDE NOPOL	1	3.21	1.3042	$\pi \rightarrow \pi^*$
	3	3.73	0.2594	$\pi \rightarrow \pi^*$ (0.46)
				$\pi \rightarrow \pi^*$ (0.38)
	5	4.02	0.1711	$\pi \rightarrow \pi^*$
FDE GSPOL	1	3.24	1.3477	$\pi \rightarrow \pi^*$
	3	3.74	0.2435	$\pi \rightarrow \pi^*$
				$\pi \rightarrow \pi^*$
				$\pi \rightarrow \text{ESO}$ (0.36)
	4	4.03	0.1459	$\pi \rightarrow \pi^*$ (0.30)
				$\pi \rightarrow \pi^*$ (0.27)
Reference	1	3.20	1.1573	
	2	3.27	0.1764	
	3	3.66	0.1564	
	5	4.01	0.1831	

**A.2.3.2. Introduction of Atomic Pseudopotentials to PE**

**Figure A.10.:** Orbitals involved in the strongest transition of pFTAA in the different models for snapshot 2 calculated with atomic pseudo potentials. Numbers in brackets after the model denote the calculated transition shown. Numbers in brackets after the orbital type transition denote the contribution of the orbitals to the calculated transition.

**Table A.29.:** Strongest transitions and identification of the contributing orbitals of the PE calculations for snapshots 2 of pFTAA in a 3 Å water environment and the different polarization models including pseudo potentials (PP). If more than one high contribution was found, those were given in brackets.

Model	Transition	$E^{\text{ex}}$	$f_{\text{L}}$	Orbitals
PE NOPOL + PP	1	3.18	1.3077	$\pi \rightarrow \pi^*$
	2	3.59	0.1480	$\pi \rightarrow \pi^*$
	3	3.90	0.1999	$\pi \rightarrow \pi^*$
PE GSPOL + PP	1	3.24	1.3320	$\pi \rightarrow \pi^*$ $\pi \rightarrow \pi^*$ (0.42)
	2	3.71	0.1941	$\pi \rightarrow \pi^*$ (0.39) $\pi \rightarrow \pi^*$ (0.36)
	3	3.99	0.2271	$\pi \rightarrow \pi^*$ (0.41)
PE DPOL + PP	1	3.16	1.3839	$\pi \rightarrow \pi^*$ (0.57) $\pi \rightarrow \pi^*$ (0.51)
	2	3.53	0.1208	$\pi \rightarrow \pi^*$ (0.31) $\pi \rightarrow \pi^*$ (0.24)
	3	3.83	0.2510	$\pi \rightarrow \pi^*$ (0.47) $n \rightarrow \pi^*$ (0.27) $\pi \rightarrow \pi^*$ (0.26)

### A.2.3.3. Shifts of Pentameric Formyl Thiophene Acetic Acid

The calculated excitation energies (in eV) and oscillator strengths for the models NOPOL, GSPOL and PE DPOL of the strongest  $\pi \rightarrow \pi^*$  transition are given. Calculations were done with PE and FDE embedding models for different configurations of pNA and pFTAA. The definition of the different shifts ( $\Delta$ NOPOL,  $\Delta$ GSPOL,  $\Delta$ DPOL,  $\Delta$  DPOL+EEF, and  $\Delta$ REF) as well as contributions ( $\Delta\Delta$ NOPOL,  $\Delta\Delta$ GSPOL,  $\Delta\Delta$ DPOL,  $\Delta\Delta$ DPOL+EEF, and  $\Delta\Delta$  EEF) are described in Fig. 4.2.4.

### Excitation energies and $S$ -shifts of pentameric formyl thiophene acetic acid in a 3 Å water environment in aug-cc-pVDZ quality

**Table A.30.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of pFTAA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a FDE embedding model. All snapshots containing two sodium ions in close proximity are marked with \*\*. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{FDE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	2.92	3.07	0.15	3.09	0.01	0.17	0.17	3.05	0.14
Snap 2**	3.17	3.21	0.04	3.24	0.02	0.07	0.07	3.20	0.03
$\emptyset$	3.05	3.14	0.10	3.16	0.02	0.12	0.12	3.13	0.09

**Table A.31.:** Excitation energies and contributions from the different polarization models to the  $S$ -shifts and their average for different configurations of pFTAA in aug-cc-pVDZ quality and a 3 Å water environment obtained with a PE embedding model (and optional effective external potential (EEF)). All snapshots containing one or two sodium ions in close proximity are marked with \* and \*\*, respectively, and incorporate pseudopotentials in the PE calculations. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$\Delta E^{\text{ex}}(\text{PE})$ :	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	2.92	3.08	0.16	3.09	0.01	3.06	-0.03	0.17	0.14	3.05	0.14
Snap 2**	3.17	3.18	0.01	3.24	0.05	3.16	-0.08	0.07	-0.02	3.20	0.03
Snap 3*	2.78	2.68	-0.09	2.85	0.17	2.82	-0.03	0.07	0.04	2.64	-0.13
Snap 4*	2.98	2.90	-0.08	2.97	0.07	2.95	-0.02	-0.01	-0.04	2.90	-0.09
Snap 5	3.20	3.31	0.11	3.32	0.02	3.29	-0.04	0.13	0.09	3.28	0.08
Snap 6	3.36	3.46	0.10	3.46	0.00	3.43	-0.03	0.10	0.07	3.42	0.06
Snap 7	2.74	2.80	0.06	2.83	0.03	2.79	-0.04	0.09	0.05	2.80	0.06
Snap 8**	3.05	2.91	-0.14	2.95	0.03	2.91	-0.03	-0.10	-0.14	2.95	-0.10
$\emptyset$	3.02	3.04	0.02	3.09	0.05	3.05	-0.04	0.07	0.03	3.03	0.01

### Oscillator strengths and $\mathcal{F}$ -shifts of pFTAA in a 3 Å water environment in aug-cc-pVDZ quality

**Table A.32.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pFTAA in aug-cc-pVDZ quality in a 3 Å water environment obtained with a FDE embedding model. All snapshots containing two sodium ions in close proximity are marked with \*\*. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.

$f_L$ (FDE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	$\Delta$ GSPOL	$\Delta$ DPOL	REF	$\Delta$ REF
Snap 1	0.900	1.195	0.294	1.248	0.053	0.348	0.348	1.288	0.388
Snap 2**	1.165	1.304	0.140	1.348	0.044	0.183	0.183	1.157	-0.007
$\emptyset$	1.033	1.249	0.217	1.298	0.049	0.265	0.265	1.223	0.191

**Table A.33.:** Oscillator strengths and contributions from the different polarization models to the  $\mathcal{F}$ -shifts and their average for different configurations of pFTAA in aug-cc-pVDZ quality in a 3 Å water environment obtained with a PE embedding model (and optional effective external potential (EEF)). All snapshots containing one or two sodium ions in close proximity are marked with \* and \*\*, respectively, and incorporate pseudopotentials in the PE calculations. Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. *J. Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society..

$\Delta f_L$ (PE):	VACUUM	NOPOL	$\Delta$ NOPOL	GSPOL	$\Delta\Delta$ GSPOL	DPOL	$\Delta\Delta$ DPOL	DPOL+EEF	$\Delta\Delta$ EEF	$\Delta$ GSPOL	$\Delta$ DPOL	$\Delta$ DPOL+EEF	REF	$\Delta$ REF
Snap 1	0.900	1.221	0.320	1.280	0.060	1.343	0.063	1.314	-0.029	0.380	0.443	0.414	1.288	0.388
Snap 2**	1.165	1.308	0.143	1.332	0.024	1.384	0.052	1.308	-0.076	0.167	0.219	0.143	1.157	-0.007
Snap 3*	1.150	1.076	-0.074	1.226	0.150	1.284	0.058	1.212	-0.072	0.076	0.134	0.0617	1.086	-0.064
Snap 4*	1.214	1.089	-0.125	1.197	0.108	1.274	0.078	1.214	-0.060	-0.017	0.061	-0.0001	1.065	-0.148
Snap 5	1.242	1.439	0.197	1.472	0.034	1.556	0.084	1.518	-0.038	0.231	0.315	0.276	1.504	0.263
Snap 6	0.802	1.299	0.497	1.350	0.000	1.422	0.071	1.348	-0.074	0.548	0.620	0.546	1.274	0.472
Snap 7	1.094	1.282	0.188	1.358	0.077	1.420	0.061	1.352	-0.068	0.265	0.326	0.258	1.340	0.246
Snap 8**	1.359	1.470	0.111	1.500	0.030	1.534	0.034	1.440	-0.094	0.141	0.175	0.081	1.416	0.057
$\emptyset$	1.116	1.273	0.157	1.339	0.060	1.402	0.063	1.338	-0.064	0.224	0.287	0.222	1.266	0.151

## A.2.4. Publications

### A.2.4.1. Jansen, M.; Thi Minh, N. N.; Hedegård, E. D.; König, C. In *Chemical Modelling: Volume 17, 2022*

Reproduced from Ref. Jansen, M.; Thi Minh, N. N.; Hedegård, E. D.; König, C. In *Chemical Modelling: Volume 17, 2022* with permission from the Royal Society of Chemistry.

## Quantum-derived embedding schemes for local excitations

Marina Jansen,<sup>a</sup> Nghia Nguyen Thi Minh,<sup>a</sup>  
Erik D. Hedegård<sup>b</sup> and Carolin König<sup>\*a</sup>

DOI: 10.1039/9781839169342-00024

Quantum-mechanical and classical embedding schemes both employ effective operators that incorporate the environmental effects for local properties. Despite their common goal, the two methods follow rather different philosophies: quantum-mechanical density-based embedding schemes aim at a reformulation of density-functional theory methods through a subsystem formalism, while in classical polarizable embedding models, perturbation theory is employed to derive and parametrize the effective operators. In this book chapter, we derive these two methodologies in a common framework, highlighting the common ideas as well as differences between them. We further review the latest extensions to these methodologies as well as recent applications to local absorption properties of small molecules as well as the green fluorescent protein.

### 1 Introduction

Accurate theoretical calculations of absorption and fluorescent spectra have been a long-standing goal of theoretical chemistry. If fulfilled, such calculations can enable design of both, novel optical materials as well as new diagnostics. One of the most obvious examples of the latter are dyes used in bioimaging.<sup>1–6</sup> For instance, fluorescent dyes are good candidates for the detection of protein misfolds that are hallmarks for neurodegenerative diseases, like Alzheimer's disease and Parkinson's disease.<sup>7</sup> Another, prominent example is the green fluorescent protein (GFP) and related fluorescent proteins, which serve as protein biomarkers.<sup>8,9</sup> An understanding of which factors govern the optical properties of dyes or proteins alike can only be achieved in atomistic detail with the help of theoretical methods.<sup>10–12</sup>

For the calculation of optical properties, quantum-mechanical (QM) approaches are indispensable, but their high computational cost makes their application unfeasible for most biomolecular systems. For solvated systems, one way to tackle this situation is to treat the solvent as a structureless continuum.<sup>13</sup> Yet, this approach does not cover strong solvent-solute interactions (*e.g.* hydrogen bonds), where explicit inclusion of solvent molecules is required. Further, continuum models are not constructed for heterogeneous environments such as proteins. The most prominent approach to heterogeneous as well as hydrogen-binding environments is a quantum-mechanical/molecular-mechanical (QM/MM) model.<sup>14,15</sup> In the QM/MM models, the dye is treated by QM methods and its environment by

<sup>a</sup>*Institute of Physical Chemistry and Electrochemistry, Leibniz University Hannover, Callinstr. 3A, D 30167 Hannover, Germany. E-mail: carolin.koenig@pci.uni-hannover.de*

<sup>b</sup>*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, DK 5230 Odense M, Denmark*

a classical description. This classical description usually means that the electrostatic part of the interaction energy is treated by point charges. This treatment has been shown to be insufficient in many cases for optical properties.<sup>16</sup> Indeed, for GFP it has been shown that protein models beyond the electrostatics provided by standard MM calculations are needed for agreement with experimental absorption spectra.<sup>17–19</sup>

In this book chapter, we focus on *embedding schemes* for optical transitions. The QM/MM model can be seen as a very simple formulation of an embedding scheme, but we focus on more advanced formulations in which the environmental description is obtained from *fragments* or *sub-systems* (the environment can be solvent, protein or alike). We only discuss *local* absorption and emission phenomena, where local refers to transitions that can be described by orbitals within the active region (the dye). Thus, we assume that no (or little) excitonic coupling or charge transfer occurs with the environment.

Historically, two main methodologies have been employed for describing electronic transitions: state-specific or linear-response-based methodologies, where the former directly optimizes the individual excited state wave functions. Linear-response methodologies<sup>20–24</sup> are time-dependent perturbation-based models that avoid explicit optimization of the excited state wave function. Response methods have the further advantage that they give rise to a full set of possible optical transitions and intensities within one matrix eigenvalue equation. A large part of this chapter is devoted to the theoretical foundation of embedding schemes when combined with linear-response theory. We also briefly comment on the state-specific methodology.

We distinguish between two different classes of embedding approaches: (i) QM/classical polarizable embedding (PE) approaches including multipole description of the environment with additional classical polarization and (ii) density-based QM/QM embedding approaches in which the environment is also treated quantum-mechanically. Both methods have been combined with a large number of electronic-structure methods,<sup>25–59</sup> but we will focus on embedding with density-functional theory (DFT) in this book chapter. Recent reviews of PE approaches<sup>16,60–62</sup> as well as QM/QM embedding schemes<sup>48,63–65</sup> can be found in the literature. They have also been collectively reviewed about a decade ago by Gomes and Jacob.<sup>66</sup> Despite QM/QM and QM/classical approaches having the same goal, their origins are rather different and comparisons are sparse. The goal of this chapter is to present a common framework for these two types of embedding approaches and use this framework to derive the effective operators employed in the respective environment models. We next compare how the effective potential manifests in the linear-response equations. Based on this common framework, the differences as well as advantages and disadvantages of the methodologies are discussed. We further review recent extensions to both density-based QM/QM and QM/classical approaches and show that some of these extensions bring these two classes of embedding methods closer together. Finally, we review studies where the methods have been applied to small model systems and GFP, respectively.



## 2 Defining the interaction energy and Hamiltonian

In this section, we formally compare the energy expressions from density-based QM/QM and classical PE methods. We will investigate two different formulations; one is denoted frozen-density embedding (FDE)<sup>48,63,64,67</sup> and the other relies on a multipole expansion of the densities/potentials.<sup>16,68–70</sup> As a toy model we consider a total system comprised of two subsystems A and B. We define one of these subsystems as the active subsystem, while we define the other as the environment; we arbitrarily choose subsystem A as the active one. We first focus on the interaction energy and Hamiltonian for the full system. Next, we define an effective Hamiltonian for the active subsystem only. Finally, we employ this effective Hamiltonian in a linear-response time-dependent density-functional theory (LR-TD-DFT) framework to obtain local excitations.

### 2.1 Density- and multipole-based interaction energies

We decompose the total energy of the combined A and B subsystems as a sum of the energy functionals

$$E_{\text{tot}}[\rho_{\text{tot}}] = E_{\text{A}}[\rho_{\text{A}}] + E_{\text{B}}[\rho_{\text{B}}] + E_{\text{int}}[\rho_{\text{A}}, \rho_{\text{B}}], \quad (1)$$

where the first two terms denote energy for subsystem A and B, respectively, and  $E_{\text{int}}[\rho_{\text{A}}, \rho_{\text{B}}]$  defines the interaction energy, *i.e.*  $E_{\text{int}}[\rho_{\text{A}}, \rho_{\text{B}}] = E_{\text{tot}}[\rho_{\text{tot}}] - E_{\text{A}}[\rho_{\text{A}}] - E_{\text{B}}[\rho_{\text{B}}]$ . Thus, the interaction energy in eqn (1) is defined with reference to  $E_{\text{A}}[\rho_{\text{A}}]$  and  $E_{\text{B}}[\rho_{\text{B}}]$ , but we have so far left the form of these energies unspecified. Note that the interaction energy  $E_{\text{int}}$  does not only depend on  $\rho_{\text{A}}$  and  $\rho_{\text{B}}$ , but also on  $\rho_{\text{tot}}$ . We, however, skip the  $\rho_{\text{tot}}$ -dependence here as we below choose to define  $\rho_{\text{tot}} = \rho_{\text{A}} + \rho_{\text{B}}$ .

We can further decompose the interaction energy into Coulomb (classical),  $E_{\text{int}}^{\text{C}}[\rho_{\text{A}}, \rho_{\text{B}}]$ , and QM contributions,  $E_{\text{int}}^{\text{QM}}[\rho_{\text{A}}, \rho_{\text{B}}]$

$$E_{\text{int}}[\rho_{\text{A}}, \rho_{\text{B}}] = E_{\text{int}}^{\text{C}}[\rho_{\text{A}}, \rho_{\text{B}}] + E_{\text{int}}^{\text{QM}}[\rho_{\text{A}}, \rho_{\text{B}}]. \quad (2)$$

Both contributions can be obtained analogously to eqn (1)

$$E_{\text{int}}^{\text{C}}[\rho_{\text{A}}, \rho_{\text{B}}] = E_{\text{tot}}^{\text{C}}[\rho_{\text{tot}}] - E_{\text{A}}^{\text{C}}[\rho_{\text{A}}] - E_{\text{B}}^{\text{C}}[\rho_{\text{B}}] \quad (3)$$

$$E_{\text{int}}^{\text{QM}}[\rho_{\text{A}}, \rho_{\text{B}}] = E_{\text{tot}}^{\text{QM}}[\rho_{\text{tot}}] - E_{\text{A}}^{\text{QM}}[\rho_{\text{A}}] - E_{\text{B}}^{\text{QM}}[\rho_{\text{B}}]. \quad (4)$$

We focus on the Coulomb part as this is included in both density-embedding and classical embedding schemes. In practice,  $E^{\text{QM}}$  will of course be present and we describe how to deal with this part in Section 2.2. The Coulomb energy for a molecular multi-electron system with density  $\rho$  in atomic units is defined by

$$E^{\text{C}}[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \text{d}\mathbf{r}\text{d}\mathbf{r}' - \sum_{\text{N}} \int \frac{\rho(\mathbf{r})Z_{\text{N}}}{|\mathbf{r} - \mathbf{R}_{\text{N}}|} \text{d}\mathbf{r} + \sum_{\text{N}} \sum_{\text{M} \neq \text{N}} \frac{Z_{\text{N}}Z_{\text{M}}}{|\mathbf{R}_{\text{N}} - \mathbf{R}_{\text{M}}|}, \quad (5)$$

where  $Z_{\text{N}}$  defines the charge of nucleus  $N$  and the electronic and nuclear coordinates are denoted by small  $\mathbf{r}$  and capital  $\mathbf{R}$ , respectively. We will in the following use eqn (5) to define the classical interaction energy

between subsystems A and B in our toy system. The total density will be expressed as  $\rho_{\text{tot}}(\mathbf{r}) = \rho_A(\mathbf{r}_a) + \rho_B(\mathbf{r}_b)$ . Inserting this into eqn (5) and using the definition in eqn (3) for the Coulombic interaction energy between subsystems A and B, one obtains

$$E_{\text{int}}^{\text{C}}[\rho_A, \rho_B] = \iint \frac{\rho_A(\mathbf{r}_a)\rho_B(\mathbf{r}_b)}{|\mathbf{r}_a - \mathbf{r}_b|} d\mathbf{r}_a d\mathbf{r}_b - \sum_{I \in A} \int \frac{Z_I \rho_B(\mathbf{r}_b)}{|\mathbf{R}_I - \mathbf{r}_b|} d\mathbf{r}_b - \sum_{J \in B} \int \frac{Z_J \rho_A(\mathbf{r}_a)}{|\mathbf{R}_J - \mathbf{r}_a|} d\mathbf{r}_a + \sum_{I \in A} \sum_{J \in B} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (6)$$

where sums run over all nuclei with charges  $Z_I$  and  $Z_J$  in subsystem A and B, respectively. We stress that we have to this point not made any assumption regarding possible overlap of the densities and underlying subsystem wave functions.

Since the classical embedding schemes often are derived in terms of the underlying Hamilton operator, rather than the densities,<sup>71</sup> we also define a Hamiltonian for the total system. This Hamiltonian is defined as

$$\hat{H}_{\text{tot}} = \hat{H}_A + \hat{H}_B + \hat{H}_{\text{int}}, \quad (7)$$

where  $\hat{H}_X$  are Hamiltonians for the isolated subsystems ( $X=A, B$ ). The interaction Hamiltonian,  $\hat{H}_{\text{int}}$ , contains interactions between electrons and nuclei in the two subsystems,

$$\hat{H}_{\text{int}} = \sum_{i \in A} \sum_{j \in B} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i \in A} \sum_{J \in B} \frac{Z_J}{|\mathbf{r}_i - \mathbf{R}_J|} - \sum_{I \in A} \sum_{j \in B} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_j|} + \sum_{I \in A} \sum_{J \in B} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \quad (8)$$

The interaction energy can be obtained by using  $\hat{H}_{\text{int}}$  as the perturbation in Rayleigh–Schrödinger perturbation theory; classical embedding theory including polarization is formally derived in this framework as will be elaborated later in this section.

We can now define  $|\Psi_X^{(0)}\rangle$  as the optimized wave function for an isolated subsystem and  $\hat{H}_{\text{int}}$  (eqn (8)) as perturbation. However, using the unperturbed wave functions requires specific care as we then define  $|\Psi_{\text{tot}}^{(0)}\rangle = |\Psi_A^{(0)}\rangle |\Psi_B^{(0)}\rangle$  as the total unperturbed wave function. This wave function is an eigenfunction of the unperturbed Hamiltonian  $\hat{H}_0 = \hat{H}_A + \hat{H}_B$ , ensuring that the monomer energies for the isolated subsystems are obtained at zeroth order. However,  $|\Psi_{\text{tot}}^{(0)}\rangle$  violates the anti-symmetry requirement with respect to interchanging electrons in subsystems A and B. We could define an anti-symmetric wave function  $\mathcal{A}|\Psi_{\text{tot}}^{(0)}\rangle$ , but this wave function is not an eigenfunction of  $\hat{H}_0$ , removing the theoretical ground for using perturbation theory. Theories to circumvent the issue of anti-symmetrization while treating  $\hat{H}_{\text{int}}$  in the basis of unperturbed wave functions have been developed within the framework of symmetry-adapted perturbation theory.<sup>72</sup> The classical

embedding models discussed in this book chapter take a different (and less rigorous) approach assuming that the subsystems A and B do not overlap. We also assume that the wave functions for A and B are always individually normalized, *i.e.*  $\langle \Psi_X | \Psi_Y \rangle = \delta_{XY}$ .

In case A and B do not overlap, the  $E_{\text{int}}^{\text{QM}}[\rho_A, \rho_B]$  in eqn (2) vanishes, and we can focus on the Coulombic interaction energy in eqn (3). Moreover, we can ascribe certain electron densities to subsystem A and B, respectively. The Hamiltonian can be connected to the density-based expression through the density operator

$$\hat{\rho}_X(\mathbf{r}_x) = \sum_{i \in X} \delta(\mathbf{r}_i - \mathbf{r}_x), \quad (9)$$

where the italic index  $i$  refers to a single-electron coordinate and the index  $x$  refers to a real-space coordinate of the electron density of subsystem X. From the density operator we can obtain the density as  $\rho_X(\mathbf{r}_x) = \langle \Psi_X | \hat{\rho}_X(\mathbf{r}_x) | \Psi_X \rangle$ .

The density operator is a central operator in the present comparison as density-based embedding schemes are often formulated in terms of real-space electron densities, whereas QM/classical approaches are more frequently expressed in terms of operators acting on electron coordinates. Using the density operator introduced in eqn (9), we obtain the equivalent definition of  $\hat{H}_{\text{int}}$

$$\begin{aligned} \hat{H}_{\text{int}} = & \iint \frac{\hat{\rho}_A(\mathbf{r}_a) \hat{\rho}_B(\mathbf{r}_b)}{|\mathbf{r}_a - \mathbf{r}_b|} d\mathbf{r}_a d\mathbf{r}_b - \sum_{I \in A} \int \frac{\hat{\rho}_B(\mathbf{r}_b) Z_I}{|\mathbf{R}_I - \mathbf{r}_b|} d\mathbf{r}_b \\ & - \sum_{J \in B} \int \frac{\hat{\rho}_A(\mathbf{r}_a) Z_J}{|\mathbf{R}_J - \mathbf{r}_a|} d\mathbf{r}_a + \sum_{I \in A} \sum_{J \in B} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \end{aligned} \quad (10)$$

In the following, we will use eqn (6)-(10) to analyze different contributions to the Coulombic interaction energy. We first investigate a scenario where all densities are kept frozen, and next a scenario in which one of the subsystems is allowed to relax. Finally, we investigate the case where polarization effects of both densities with respect to each other are considered.

**2.1.1 Interaction energy for frozen densities.** In the case where there is no relaxation of the two subsystems, the isolated (non-polarized) density for the total system is  $\rho_{\text{tot}}^{(0)}(\mathbf{r}) = \rho_A^{(0)}(\mathbf{r}_a) + \rho_B^{(0)}(\mathbf{r}_b)$  and we obtain from eqn (3) and (5)

$$\begin{aligned} E_{\text{int}}^{\text{C}}[\rho_A^{(0)}, \rho_B^{(0)}] = & \iint \frac{\rho_A^{(0)}(\mathbf{r}_a) \rho_B^{(0)}(\mathbf{r}_b)}{|\mathbf{r}_a - \mathbf{r}_b|} d\mathbf{r}_a d\mathbf{r}_b - \sum_{I \in A} \int \frac{Z_I \rho_B^{(0)}(\mathbf{r}_b)}{|\mathbf{R}_I - \mathbf{r}_b|} d\mathbf{r}_b \\ & - \sum_{J \in B} \int \frac{Z_J \rho_A^{(0)}(\mathbf{r}_a)}{|\mathbf{R}_J - \mathbf{r}_a|} d\mathbf{r}_a + \sum_{I \in A} \sum_{J \in B} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \end{aligned} \quad (11)$$

similarly to how we obtained eqn (6).

In density-based embedding methods, all involved subsystems within eqn (11) are treated by means of real-space densities. In the PE embedding methods, the density of the environmental subsystem is expressed by discretized multipoles. We show the procedure to obtain the multipoles below: using  $\hat{H}_{\text{int}}$  in the form of eqn (10), we obtain the expression

$$\begin{aligned}
 E_{\text{int}}^{\text{C}}[\rho_{\text{A}}^{(0)}, \rho_{\text{B}}^{(0)}] &= \langle \Psi_{\text{tot}}^{(0)} | \hat{H}_{\text{int}} | \Psi_{\text{tot}}^{(0)} \rangle = \langle \Psi_{\text{A}}^{(0)} | \left( \langle \Psi_{\text{B}}^{(0)} | \hat{H}_{\text{int}} | \Psi_{\text{B}}^{(0)} \rangle \right) | \Psi_{\text{A}}^{(0)} \rangle \\
 &= \langle \Psi_{\text{A}}^{(0)} | \left( \langle \Psi_{\text{B}}^{(0)} | \left( \iint \frac{\hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}})\hat{\rho}_{\text{B}}(\mathbf{r}_{\text{b}})}{|\mathbf{r}_{\text{a}} - \mathbf{r}_{\text{b}}|} d\mathbf{r}_{\text{a}} d\mathbf{r}_{\text{b}} - \sum_{J \in \text{B}} \int \frac{Z_J \hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}})}{|\mathbf{R}_J - \mathbf{r}_{\text{a}}|} d\mathbf{r}_{\text{a}} \right. \right. \\
 &\quad \left. \left. - \sum_{I \in \text{A}} \int \frac{Z_I \hat{\rho}_{\text{B}}(\mathbf{r}_{\text{b}})}{|\mathbf{R}_I - \mathbf{r}_{\text{b}}|} d\mathbf{r}_{\text{b}} + \sum_{I \in \text{A}} \sum_{J \in \text{B}} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \right) | \Psi_{\text{B}}^{(0)} \rangle \right) | \Psi_{\text{A}}^{(0)} \rangle.
 \end{aligned}
 \tag{12}$$

Note that taking the expectation values over  $\hat{\rho}_{\text{A}}$  and  $\hat{\rho}_{\text{B}}$  leads directly to the expression in eqn (11). We can now perform a Taylor expansion of the inverse distances  $|\mathbf{r}_{\text{a}} - \mathbf{r}_{\text{b}}|^{-1}$ ,  $|\mathbf{R}_I - \mathbf{r}_{\text{b}}|^{-1}$  and  $|\mathbf{R}_J - \mathbf{r}_{\text{a}}|^{-1}$ ,  $|\mathbf{R}_I - \mathbf{R}_J|^{-1}$  using eqn (76) and (77) in the Appendix. To shorten the resulting expression, we define the interaction operators  $T_{ji}^{(0)} = |\mathbf{R}_i - \mathbf{R}_j|^{-1}$ ,  $T_{ji,\alpha}^{(1)} = \frac{\partial}{\partial R_{i,\alpha}} |\mathbf{R}_i - \mathbf{R}_j|^{-1}$  and so on, according to the book by Stone.<sup>71</sup> Here,  $\alpha$  denotes a Cartesian coordinate,  $x$ ,  $y$ , or  $z$ . We choose to describe the electron density of subsystem B,  $\rho_{\text{B}}$ , in terms of multipoles and define the expansion point of the Taylor series as the centre of mass of system B ( $\mathbf{R}_{\text{B}}$ ). This leads to

$$\begin{aligned}
 E_{\text{int}}^{\text{C}}[\rho_{\text{A}}^{(0)}, \rho_{\text{B}}^{(0)}] &\approx E_{\text{int}}^{\text{mult}}[\rho_{\text{A}}^{(0)}, \rho_{\text{B}}^{(0)}] \\
 &= \left( -\langle \Psi_{\text{A}}^{(0)} | \int \hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}}) T_{\text{Ba}}^{(0)} d\mathbf{r}_{\text{a}} | \Psi_{\text{A}}^{(0)} \rangle + \sum_{I \in \text{A}} Z_I T_{\text{BI}}^{(0)} \right) \langle \Psi_{\text{B}}^{(0)} | \hat{q}_{\text{B}} | \Psi_{\text{B}}^{(0)} \rangle \\
 &\quad - \sum_{\alpha} \left( -\langle \Psi_{\text{A}}^{(0)} | \int \hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}}) T_{\text{Ba},\alpha}^{(1)} d\mathbf{r}_{\text{a}} | \Psi_{\text{A}}^{(0)} \rangle + \sum_{I \in \text{A}} Z_I T_{\text{BI},\alpha}^{(1)} \right) \\
 &\quad \times \langle \Psi_{\text{B}}^{(0)} | \hat{\mu}_{\text{B},\alpha} | \Psi_{\text{B}}^{(0)} \rangle + \dots
 \end{aligned}
 \tag{14}$$

Note that  $\mathbf{T}_{\text{Ba}}^{(k)}$  is also dependent on  $\mathbf{r}_{\text{a}}$  and although this dependence is not parametric, we let this dependence be implicit as in the definition in the book by Stone.<sup>71</sup> In eqn (14), we have defined the multipole moment operators with respect to the environmental charge density

$$\hat{q}_{\text{B}} = - \int \hat{\rho}_{\text{B}}(\mathbf{r}_{\text{b}}) d\mathbf{r}_{\text{b}} + \sum_{J \in \text{B}} Z_J
 \tag{15}$$

$$\hat{\mu}_{\text{B},\alpha} = - \int \hat{\rho}_{\text{B}}(\mathbf{r}_{\text{b}}) (r_{\text{b},\alpha} - R_{\text{B},\alpha}) d\mathbf{r}_{\text{b}} + \sum_{J \in \text{B}} Z_J (R_{J,\alpha} - R_{\text{B},\alpha}).
 \tag{16}$$

Note that the multipole moment operators combine nuclei and electron contributions from the environment. When the multipoles are obtained from  $|\Psi_B^{(0)}\rangle$  we use the short-hand notation  $q_B[\rho_B^{(0)}] \equiv \langle \Psi_B^{(0)} | \hat{q}_B | \Psi_B^{(0)} \rangle$ ,  $\mu_{B,\alpha}[\rho_B^{(0)}] \equiv \langle \Psi_B^{(0)} | \hat{\mu}_{B,\alpha} | \Psi_B^{(0)} \rangle$  and so on. In later expressions we will even further shorten the notation as  $q_B^0$ ,  $\mu_{B,\alpha}^0$  and so forth. Taking the expectation values in eqn (14), we obtain the Coulombic interaction energy of the unperturbed electron density  $\rho_A^{(0)}$  in subsystem A with the multipole description of subsystem B,

$$E_{\text{int}}^{\text{mult}}[\rho_A^{(0)}, \rho_B^{(0)}] = \left( -\int \rho_A^{(0)}(\mathbf{r}_a) T_{\text{Ba}}^{(0)} d\mathbf{r}_a + \sum_{I \in A} Z_I T_{\text{BI}}^{(0)} \right) q_B[\rho_B^{(0)}] - \sum_{\alpha} \left( -\int \rho_A^{(0)}(\mathbf{r}_a) T_{\text{Ba},\alpha}^{(1)} d\mathbf{r}_a + \sum_{I \in A} Z_I T_{\text{BI},\alpha}^{(1)} \right) \mu_{B,\alpha}[\rho_B^{(0)}] + \dots \quad (17)$$

The equation with all densities frozen in eqn (17) is rarely used with one notable exception: in the so-called mechanical embedding strategy to describe electrostatic interactions in QM/MM methods the first term (point charges) in eqn (17) is employed to describe the electrostatics between QM and MM regions.<sup>15</sup>

**2.1.2 Polarization of one subsystem density.** In a next step, we allow the active subsystem to relax in the potential from the (still frozen) environment density  $\rho_B^{(0)}$ . This will lead to a polarization of the electron density of subsystem A, and we denote the *polarized* density  $\rho_A^{(1)}$ . Within the density-based embedding schemes, this approach refers to FDE.<sup>25</sup> The Coulombic interaction energy can in this case be obtained by inserting  $\rho_{\text{tot}}^{(1)} = \rho_A^{(1)} + \rho_B^{(0)}$  in eqn (5)

$$E_{\text{tot}}^{\text{C}}[\rho_{\text{tot}}^{(1)}] = E_{\text{int}}^{\text{C}}[\rho_A^{(1)}, \rho_B^{(0)}] + E_{\text{A}}^{\text{C}}[\rho_A^{(1)}] + E_{\text{B}}^{\text{C}}[\rho_B^{(0)}]. \quad (18)$$

The  $E_{\text{int}}^{\text{C}}[\rho_A^{(1)}, \rho_B^{(0)}]$  term in eqn (18) is defined in eqn (11) (with  $\rho_A^{(0)}$  replaced by  $\rho_A^{(1)}$ ), but can naturally also be defined as a multipole expansion. The procedure is identical to the one used to obtain  $E_{\text{int}}^{\text{C}}[\rho_A^{(0)}, \rho_B^{(0)}]$  in the unrelaxed case. The sole exception is that we now allow the active subsystem to relax, and hence wave function  $|\Psi_A^{(0)}\rangle$  in eqn (17) changes to  $|\Psi_A^{(1)}\rangle$ . We thus perform the same manipulations for the multipole expansion as in eqn (12)–(14), but we use the opportunity to collect the orders of the Taylor expansion into a multi-index  $k$  (see eqn (78) and (79) in the Appendix)

$$E_{\text{int}}^{\text{mult}}[\rho_A^{(1)}, \rho_B^{(0)}] = \sum_{k=0} \frac{(-1)^{|k|}}{k!} \left( -\langle \Psi_A^{(1)} | \int \hat{\rho}_A(\mathbf{r}_a) \mathbf{T}_{\text{Ba}}^{(k)} d\mathbf{r}_a | \Psi_A^{(1)} \rangle + \sum_{I \in A} Z_I \mathbf{T}_{\text{BI}}^{(k)} \right) \times \langle \Psi_B^{(0)} | \hat{\mathbf{Q}}_B^{(k)} | \Psi_B^{(0)} \rangle. \quad (19)$$

In eqn (19), we have further defined a general multipole-moment vector operator,  $\hat{\mathbf{Q}}_B^{(k)}$  where the orders correspond to point charge ( $k=0$ ), dipole ( $k=1$ ), quadrupole ( $k=2$ ), *etc.* Taking the expectation values in eqn (19) yields the expression

$$E_{\text{int}}^{\text{mult}}[\rho_A^{(1)}, \rho_B^{(0)}] = \sum_{k=0} \frac{(-1)^{|k|}}{k!} \left( - \int \rho_A^{(1)}(\mathbf{r}_a) \mathbf{T}_{Ba}^{(k)} d\mathbf{r}_a + \sum_{I \in A} Z_I \mathbf{T}_{BI}^{(k)} \right) \mathbf{Q}_B^{(k)}[\rho_B^{(0)}], \quad (20)$$

where  $\mathbf{Q}_B^{(k)}[\rho_B^{(0)}] \equiv \langle \Psi_B^{(0)} | \hat{\mathbf{Q}}_B^{(k)} | \Psi_B^{(0)} \rangle$ . The first term in the expansion ( $k=0$ ) in eqn (20) corresponds to the parametrization employed in electrostatically embedding QM/MM models.<sup>15</sup> At this point, we also note that electrostatically embedded QM/MM (and also polarizable extensions) usually employ a distributed origin and we return to this below (*cf.* Section 2.2).

**2.1.3 Mutual polarization of densities.** The above-described change from  $\rho_A^{(0)}$  to  $\rho_A^{(1)}$  will naturally also induce a change on subsystem B, which will then again change  $\rho_A$ . One method to capture this *mutual polarization* effect is a step-wise optimization of  $\rho_A$  and  $\rho_B$  in a number of steps until self-consistency. This process of alternating which subsystem is optimized, also known as freeze-and-thaw<sup>26</sup> procedure, is often applied in density-based embedding methods.<sup>26,40,43,73</sup> We here denote relaxed densities obtained from such freeze-and-thaw cycles  $\rho_A^{(2)}$  and  $\rho_B^{(2)}$ . Within the context of density-based embedding schemes, these freeze-and-thaw cycles over many FDE calculations formally recover the super-system result.<sup>26,63</sup>

Similarly, also classical multipole expansions of the densities may be updated in a self-consistent scheme to describe mutual polarization; this corresponds to employing a  $\mathbf{Q}_B^{(k)}[\rho_B]$  with a mutually polarized  $\rho_B$  in eqn (20). This is done in the embedded X-Pol model.<sup>69,74,75</sup> A more common method in classical embedding schemes is to construct an energy expression based on perturbation theory<sup>71</sup> and we spend the rest of this section to show this procedure.

If we assume that a full set of solutions is available for the isolated subsystems,

$$\hat{H}_X |\Psi_{X,n}^{(0)}\rangle = E_{X,n}^{(0)} |\Psi_{X,n}^{(0)}\rangle, \quad (21)$$

and choose  $\hat{H}_{\text{int}}$  from eqn (7) as the perturbation operator, we obtain from Rayleigh–Schrödinger perturbation theory,

$$E_{\text{tot}} = E^{(0)} + E^{(1)} + E^{(2)}. \quad (22)$$

We can from eqn (1) identify  $E^{(0)} = E_A[\rho_A^{(0)}] + E_B[\rho_B^{(0)}]$  and the interaction energy,  $E_{\text{int}}[\rho_A^{(0)}, \rho_B^{(0)}]$ , comes through first and second-order energy corrections. We recognize the former as

$$E^{(1)} = \langle \Psi_{\text{tot}}^{(0)} | \hat{H}_{\text{int}} | \Psi_{\text{tot}}^{(0)} \rangle = E_{\text{int}}^{\text{C}}[\rho_A^{(0)}, \rho_B^{(0)}] \approx E_{\text{int}}^{\text{mult}}[\rho_A^{(0)}, \rho_B^{(0)}], \quad (23)$$

where the term  $E_{\text{int}}^{\text{C}}[\rho_{\text{A}}^{(0)}, \rho_{\text{B}}^{(0)}]$  is given in eqn (12). As implied by the above equation, we can follow eqn (12)–(17) to replace the densities by a multipole expansion to obtain the first-order energy correction  $E^{(1)}$ . The second-order corrections are more complicated, as they involve excited-state wave functions. We can decompose this part into<sup>71</sup>

$$E^{(2)} = E_{\text{A}}^{\text{pol}} + E_{\text{B}}^{\text{pol}} + E^{\text{disp}}. \quad (24)$$

The terms are sorted such that the expression for  $E_{\text{B}}^{\text{pol}}$  includes excited states for subsystem B, whereas the wave function for A remains in the ground state (and *vice versa* for  $E_{\text{A}}^{\text{pol}}$ ). The dispersion part involves excited states from both subsystems simultaneously,<sup>71</sup> but we will disregard this part. The perturbation expression for  $E_{\text{B}}^{\text{pol}}$  becomes

$$E_{\text{B}}^{\text{pol}} = \sum_{n \neq 0} \frac{\langle \Psi_{\text{B},0}^{(0)} | \langle \Psi_{\text{A},0}^{(0)} | \hat{H}_{\text{int}} | \Psi_{\text{A},0}^{(0)} \rangle | \Psi_{\text{B},n}^{(0)} \rangle \langle \Psi_{\text{B},n}^{(0)} | \langle \Psi_{\text{A},0}^{(0)} | \hat{H}_{\text{int}} | \Psi_{\text{A},0}^{(0)} \rangle | \Psi_{\text{B},0}^{(0)} \rangle}{E_{\text{B},0}^{(0)} - E_{\text{B},n}^{(0)}}, \quad (25)$$

and  $E_{\text{A}}^{\text{pol}}$  can be obtained analogously. Inserting  $\hat{H}_{\text{int}}$  and performing the Taylor expansion as done above, leads to

$$E_{\text{B}}^{\text{pol}} = \sum_{\alpha\beta} \left\{ \left( - \langle \Psi_{\text{A},0}^{(0)} | \int \hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}}) T_{\text{Ba},\alpha}^{(1)} d\mathbf{r}_{\text{a}} | \Psi_{\text{A},0}^{(0)} \rangle + \sum_{I \in \text{A}} Z_I T_{\text{BI},\alpha}^{(1)} \right) \right. \\ \times \sum_{n \neq 0} \frac{\langle \Psi_{\text{B},0}^{(0)} | \hat{\mu}_{\text{B},\alpha} | \Psi_{\text{B},n}^{(0)} \rangle \langle \Psi_{\text{B},n}^{(0)} | \hat{\mu}_{\text{B},\beta} | \Psi_{\text{B},0}^{(0)} \rangle}{E_{\text{B},0}^{(0)} - E_{\text{B},n}^{(0)}} \\ \left. \times \left( - \langle \Psi_{\text{A},0}^{(0)} | \int \hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}}) T_{\text{Ba},\beta}^{(1)} d\mathbf{r}_{\text{a}} | \Psi_{\text{A},0}^{(0)} \rangle + \sum_{I \in \text{A}} Z_I T_{\text{BI},\beta}^{(1)} \right) \right\}, \quad (26)$$

where we have truncated the multipole expansion after the dipole contribution. Note that contrary to eqn (17), the zeroth-order multipoles (point charges) do not contribute in eqn (26) since  $\langle \Psi_{\text{B},0}^{(0)} | \hat{q}_{\text{B}} | \Psi_{\text{B},n}^{(0)} \rangle = q_{\text{B}} \langle \Psi_{\text{B},0}^{(0)} | \Psi_{\text{B},n}^{(0)} \rangle = 0$ , and hence

$$\langle \Psi_{\text{B},0}^{(0)} | \left\{ \left( - \langle \Psi_{\text{A},0}^{(0)} | \int \hat{\rho}_{\text{A}}(\mathbf{r}_{\text{a}}) T_{\text{Ba}}^{(0)} d\mathbf{r}_{\text{a}} | \Psi_{\text{A},0}^{(0)} \rangle + \sum_{I \in \text{A}} Z_I T_{\text{BI}}^{(0)} \right) \hat{q}_{\text{B}} \right\} | \Psi_{\text{B},n}^{(0)} \rangle = 0. \quad (27)$$

The expression in eqn (26) can be further rewritten: first, we recognize the static polarizability for subsystem B with the elements

$$\alpha_{\text{B},\alpha\beta}^0 = -2 \sum_{n \neq 0} \frac{\langle \Psi_{\text{B},0}^{(0)} | \hat{\mu}_{\text{B},\alpha} | \Psi_{\text{B},n}^{(0)} \rangle \langle \Psi_{\text{B},n}^{(0)} | \hat{\mu}_{\text{B},\beta} | \Psi_{\text{B},0}^{(0)} \rangle}{E_{\text{B},0}^{(0)} - E_{\text{B},n}^{(0)}}. \quad (28)$$

Next, we use that the potential at  $\mathbf{R}_{\text{B}}$  due to presence of an electron in subsystem A at  $\mathbf{r}_{\text{a}}$  is  $V(\mathbf{r}_{\text{a}}) = -|\mathbf{r}_{\text{a}} - \mathbf{R}_{\text{B}}|^{-1}$  and the potential due to a

nucleus in subsystem A is  $V(\mathbf{R}_I) = Z_I |\mathbf{R}_I - \mathbf{R}_B|^{-1}$ . With these definitions along with the usual definition of an electric field  $\mathcal{E}$  as the negative derivative of the potential, we see that the terms involving the interactions operators  $\mathbf{T}_{Ba}^{(1)}$  and  $\mathbf{T}_B^{(1)}$  in eqn (26) are

$$\sum_{I \in A} Z_I T_{Bf,\alpha}^{(1)} = \sum_{I \in A} \frac{\partial V(\mathbf{R}_I)}{\partial R_{I,\alpha}} = -\mathcal{E}_\alpha^n \quad (29)$$

$$\left\langle \Psi_A^{(0)} \left| \int \hat{\rho}_A(\mathbf{r}_a) T_{Ba,\alpha}^{(1)} d\mathbf{r}_a \right| \Psi_A^{(0)} \right\rangle = - \int \rho_A^{(0)}(\mathbf{r}_a) \left( \frac{\partial V(\mathbf{r}_a)}{\partial r_{a,\alpha}} \right) d\mathbf{r}_a = \mathcal{E}_\alpha^e [\rho_A^{(0)}]. \quad (30)$$

We denote the  $\alpha$  component of field from nuclei in system A as  $\mathcal{E}_\alpha^n$  and the corresponding  $\alpha$  component from the electrons  $\mathcal{E}_\alpha^e [\rho_A^{(0)}]$ . Inserting eqn (28)–(30) in eqn (26), we arrive at

$$E_B^{\text{pol}} [\rho_A^{(0)}] = -\frac{1}{2} \left( \mathcal{E}_A^e [\rho_A^{(0)}] + \mathcal{E}_A^n \right)^T \alpha_B^0 \left( \mathcal{E}_A^e [\rho_A^{(0)}] + \mathcal{E}_A^n \right), \quad (31)$$

where we note that we have encapsulated all dependence on  $\Psi_{B,n}^{(0)}$  into a parametric dependence through the polarizabilities,  $\alpha_B^0$ , similar to what is done with the multipoles in eqn (17). The total energy (in eqn (1)) is now defined by combining eqn (22)–(24) while neglecting  $E^{\text{disp}}$  as

$$E_{\text{tot}}^{\text{PE}} [\rho_A^{(0)}, \rho_B^{(0)}] = E_A [\rho_A^{(0)}] + E_B [\rho_B^{(0)}] + E_{\text{int}}^{\text{mult}} [\rho_A^{(0)}, \rho_B^{(0)}] + E_B^{\text{pol}} [\rho_A^{(0)}] + E_A^{\text{pol}} [\rho_B^{(0)}]. \quad (32)$$

The expressions for the last three terms are given in eqn (20) and (31), respectively. For consistency with the definition in eqn (1), we write the total energy as a sum of the energies of the two subsystems A and B plus an interaction energy,

$$E_{\text{tot}}^{\text{PE}} [\rho_A^{(0)}, \rho_B^{(0)}] = E_A^{\text{PE}} [\rho_A^{(0)}, \rho_B^{(0)}] + E_B^{\text{PE}} [\rho_B^{(0)}, \rho_A^{(0)}] + E_{\text{int}}^{\text{mult}} [\rho_A^{(0)}, \rho_B^{(0)}], \quad (33)$$

where we have combined the energy of the isolated subsystem and polarization in the term

$$E_A^{\text{PE}} [\rho_A^{(0)}, \rho_B^{(0)}] = E_A [\rho_A^{(0)}] + E_A^{\text{pol}} [\rho_B^{(0)}]. \quad (34)$$

The term  $E_B^{\text{PE}} [\rho_B^{(0)}, \rho_A^{(0)}]$  is defined in a similar fashion. In next section, we will allow  $\rho_A$  to change (*e.g.*, through a self-consistent procedure), while  $\rho_B^{(0)}$  will remain the isolated density of subsystem B (represented by a multipole expansion). We denote the density that is freely optimized without superscript as  $\rho_A$ , thus writing the total energy for the PE model becomes

$$E_{\text{tot}}^{\text{PE}} [\rho_A, \rho_B^{(0)}] = E_A^{\text{PE}} [\rho_A, \rho_B^{(0)}] + E_B^{\text{PE}} [\rho_B^{(0)}, \rho_A] + E_{\text{int}}^{\text{mult}} [\rho_A, \rho_B^{(0)}]. \quad (35)$$

## 2.2 Effective embedding potentials

In the previous subsection, we considered the total system. We now focus on a selected active subsystem, while including the environment through



an effective operator. We again select subsystem A as the active subsystem and the additional embedding terms are added to the subsystem Hamiltonian,

$$\hat{H}_A^{\text{eff}} = \hat{H}_A + \hat{v}_A^{\text{emb}}. \quad (36)$$

We assume  $\hat{v}_A^{\text{emb}}$  to be a one-electron operator that can be expressed *via* a real-space potential  $v_A^{\text{emb}}(\mathbf{r}_a)$  as

$$\hat{v}_A^{\text{emb}} = \int \hat{\rho}_A v_A^{\text{emb}}(\mathbf{r}_a) d\mathbf{r}_a. \quad (37)$$

The embedding potential  $v_A^{\text{emb}}(\mathbf{r}_a)$  may then be obtained as<sup>25,67</sup>

$$v_A^{\text{emb}}(\mathbf{r}_a) = \frac{\delta}{\delta \rho_A} (E_{\text{tot}}[\rho_{\text{tot}}] - E_A[\rho_A]) = \frac{\delta E_{\text{tot}}[\rho_{\text{tot}}]}{\delta \rho_A} - \frac{\delta E_A[\rho_A]}{\delta \rho_A}, \quad (38)$$

where  $E_{\text{tot}}[\rho_{\text{tot}}]$  is the total energy as introduced in eqn (1). Directly employing the general forms in eqn (1) and (2) yields

$$v_A^{\text{emb}}(\mathbf{r}_a) = \frac{\delta E_{\text{tot}}^{\text{C}}[\rho_{\text{tot}}]}{\delta \rho_A} - \frac{\delta E_A^{\text{C}}[\rho_A]}{\delta \rho_A} + \frac{\delta E_{\text{tot}}^{\text{QM}}[\rho_{\text{tot}}]}{\delta \rho_A} - \frac{\delta E_A^{\text{QM}}[\rho_A]}{\delta \rho_A}. \quad (39)$$

We now examine, how the FDE and PE methods obtain different forms of  $v_A^{\text{emb}}(\mathbf{r}_a)$  by employing different expressions for  $E_{\text{tot}}[\rho_{\text{tot}}]$  and  $E_A[\rho_A]$ . In FDE, we assume the environmental electron density  $\rho_B$  to be constant (frozen). In this case

$$\left. \frac{\delta E_B^{\text{C}}[\rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} = \left. \frac{\delta E_B^{\text{QM}}[\rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} = 0 \quad (40)$$

so that

$$v_A^{\text{FDE}}(\mathbf{r}_a) = \left. \frac{\delta E_{\text{int}}^{\text{C}}[\rho_A, \rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} + \left. \frac{\delta E_{\text{int}}^{\text{QM}}[\rho_A, \rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}}. \quad (41)$$

It is further realized that the Coulomb contribution to the energy is additive such that the Coulomb contribution to the embedding potential depends on the environmental subsystem density  $\rho_B$  only.<sup>25,67</sup> Thus, if we insert eqn (11) in the first term of eqn (41) (but with  $\rho_A$  free to be optimized and possibly a different  $\rho_B$  than the one for an isolated system), we obtain

$$\left. \frac{\delta E_{\text{int}}^{\text{C}}[\rho_A, \rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} = - \sum_{J \in B} \frac{Z_J}{|\mathbf{r}_a - \mathbf{R}_J|} + \int \frac{\rho_B(\mathbf{r}_b)}{|\mathbf{r}_a - \mathbf{r}_b|} d\mathbf{r}_b = v_A^{\text{C}}[\rho_B](\mathbf{r}_a). \quad (42)$$

The situation is different for the QM parts:  $E_{\text{int}}^{\text{QM}}[\rho_A, \rho_B]$  (eqn (4)), is comprised of a kinetic energy contribution  $E_{\text{int}}^{\text{kin}}[\rho_A, \rho_B]$  and an exchange–correlation

contribution  $E_{\text{int}}^{\text{xc}}[\rho_A, \rho_B]$ , both of which are non-linear (non-additive) in the electron density.<sup>25,67</sup> This means that

$$\begin{aligned} v_A^{\text{QM}}[\rho_A, \rho_B](\mathbf{r}_a) &= \left. \frac{\delta E_{\text{int}}^{\text{QM}}[\rho_A, \rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} \\ &= \left. \frac{\delta E_{\text{int}}^{\text{kin}}[\rho_A, \rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} + \left. \frac{\delta E_{\text{int}}^{\text{xc}}[\rho_A, \rho_B]}{\delta \rho_A} \right|_{\rho_B = \text{const.}} \\ &= v_A^{\text{nadd,kin}}[\rho_A, \rho_B](\mathbf{r}_a) + v_A^{\text{nadd,xc}}[\rho_A, \rho_B](\mathbf{r}_a), \end{aligned} \quad (43)$$

is dependent on the electron density of the active subsystem  $\rho_A$  and that of the environment  $\rho_B$ . The overall FDE potential, hence, reads

$$v_A^{\text{FDE}}(\mathbf{r}_a) = v_A^{\text{C}}[\rho_B](\mathbf{r}_a) + v_A^{\text{nadd,kin}}[\rho_A, \rho_B](\mathbf{r}_a) + v_A^{\text{nadd,xc}}[\rho_A, \rho_B](\mathbf{r}_a). \quad (44)$$

Even though the environmental density  $\rho_B$  is assumed to be frozen here, it is possible by subsequent optimization of the subsystem electron densities in freeze-and-thaw cycles<sup>26</sup> to obtain the mutually relaxed electron densities  $\rho_A^{(2)}$  and  $\rho_B^{(2)}$  introduced in Section 2.1. At the end of the freeze-and-thaw procedure, an embedding potential  $v_A^{\text{FDE}}$  for subsystem A is obtained that incorporates the ground-state polarization of subsystem B and *vice versa*.

For fully relaxed electron densities ( $\rho_A^{(2)}$  and  $\rho_B^{(2)}$ ), eqn (44) formally leads to the supermolecular results, but the applied approximations in the embedding potential lead to deviations from that result. The latter can, however, be obtained by reconstruction schemes (see ref. 63, 76 and 77 for further information).

For the PE method, we only include the Coulomb-type embedding potentials and turn to the multipole expansion. This means, we insert eqn (35) and (34) in eqn (38) and obtain the functional derivative in close analogy with FDE above

$$v_A^{\text{PE}} = \frac{\delta E_{\text{tot}}^{\text{PE}}[\rho_A, \rho_B^{(0)}]}{\delta \rho_A} = \frac{\delta E_A^{\text{PE}}[\rho_A]}{\delta \rho_A} = \frac{\delta E_{\text{int}}^{\text{mult}}[\rho_A, \rho_B^{(0)}]}{\delta \rho_A} + \frac{\delta E_B^{\text{pol}}[\rho_A]}{\delta \rho_A}. \quad (45)$$

For the second equality in eqn (45), we exploited that  $\rho_B^{(0)}$  is represented by parameters, which are assumed to be constant.

The functional derivative of the  $E_{\text{int}}^{\text{mult}}[\rho_A, \rho_B^{(0)}]$  (cf. eqn (20)) is straightforward, whereas the functional derivative of  $E_B^{\text{pol}}[\rho_A]$  (cf. eqn (31)) requires use of the functional derivative product rule (see the Appendix of ref. 78). The final potential becomes

$$\begin{aligned} v_A^{\text{PE}}(\mathbf{r}_a) &= v^{\text{mult}}(\mathbf{r}_a) + v^{\text{pol}}[\rho_A](\mathbf{r}_a) \\ &= - \sum_{k=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{\text{Ba}}^{(k)}(\mathbf{r}_a) \mathbf{Q}_B^{0,(k)} - (\boldsymbol{\varepsilon}_A^e[\rho_A] + \boldsymbol{\varepsilon}_A^n)^T \boldsymbol{\alpha}_B^0 \boldsymbol{\varepsilon}_A^e(\mathbf{r}_a), \end{aligned} \quad (46)$$

where we have used the short form  $\mathbf{Q}_B^{0,(k)} \equiv \mathbf{Q}_B^{(k)}[\rho_B^{(0)}]$  and made the  $\mathbf{r}_a$  dependence of  $T_{Ba}^{(k)}(\mathbf{r}_a)$  explicit. Note that we in the above equation have used eqn (30) to write

$$\frac{\delta \mathcal{E}_{A,\alpha}^e[\rho_A(\mathbf{r}_a)]}{\delta \rho_A(\mathbf{r}_a)} = \varepsilon_{A,\alpha}^e(\mathbf{r}_a), \quad (47)$$

with  $\varepsilon_{A,\alpha}^e(\mathbf{r}_a)$  being the  $\alpha$  component of the electronic part of the electric field operator, defined in real-space coordinates. Since the second term in eqn (46) is dependent on  $\rho_A$ , the PE embedding operator is an effective operator.

In classical electrostatics, the term  $\boldsymbol{\alpha} \cdot \boldsymbol{\mathcal{E}}$  in eqn (46) corresponds to the first term in the induced dipole. We can thus define

$$\nu^{\text{pol}}[\rho_A](\mathbf{r}_a) = -(\boldsymbol{\mu}^{\text{ind}}[\rho_A])^T \boldsymbol{\mathcal{E}}_A^e(\mathbf{r}_a), \quad (48)$$

with  $\boldsymbol{\mu}^{\text{ind}}[\rho_A] = \boldsymbol{\alpha}_B^0(\boldsymbol{\mathcal{E}}_A^e[\rho_A] + \boldsymbol{\mathcal{E}}_A^n)$ . Below, we use this formulation in a more general form of  $\nu^{\text{pol}}[\rho_A](\mathbf{r}_a)$  including more than one environment subsystem density.

### 2.3 Extension to multiple environment subsystems

We have so far only considered two subsystem densities, of which we denoted  $\rho_B$  the environment density. However, a much more common situation is to have many subsystem densities. For instance, when  $\rho_A$  is a solute and the environment densities represent solvent molecules. In this case, the total density can be written as

$$\rho_{\text{tot}} = \rho_A + \rho_{\text{env}} = \rho_A + \sum_{X \neq A} \rho_X. \quad (49)$$

This leads to a modification of the FDE and PE potentials in eqn (44) and (46). For the former, we define  $\rho_{\text{env}} = \sum_{X \neq A} \rho_X$  and introduce sums over the subsystems. The resulting Coulomb potential reads

$$\begin{aligned} v_A^C[\rho_{\text{env}}](\mathbf{r}_a) &= - \sum_{X \neq A} \sum_{J \in X} \frac{Z_{JX}}{|\mathbf{r}_a - \mathbf{R}_{JX}|} + \sum_{X \neq A} \int \frac{\rho_X(\mathbf{r}_b)}{|\mathbf{r}_a - \mathbf{r}_b|} d\mathbf{r}_b \\ &= - \sum_{X \neq A} \sum_{J \in X} \frac{Z_{JX}}{|\mathbf{r}_a - \mathbf{R}_{JX}|} + \int \frac{\rho_{\text{env}}(\mathbf{r}_b)}{|\mathbf{r}_a - \mathbf{r}_b|} d\mathbf{r}_b, \end{aligned} \quad (50)$$

where  $Z_{JX}$  is the charge of the  $J$ th nucleus in the subsystem  $X$  and  $\mathbf{R}_{JX}$ , its spatial coordinate. The non-additive kinetic and exchange–correlation contributions can be obtained analogously by replacing  $\rho_B$  in eqn (44) by  $\rho_{\text{env}}$ , that is  $v_A^{\text{nadd,kin}}[\rho_A, \rho_{\text{env}}](\mathbf{r}_a) + v_A^{\text{nadd,xcr}}[\rho_A, \rho_{\text{env}}](\mathbf{r}_a)$ . Within a freeze-and-thaw procedure  $\rho_{\text{env}}$  is redefined for every subsystem  $Y$  in the procedure as  $\rho_{\text{env}} = \sum_{X \neq Y} \rho_X$ .

The PE potential in eqn (46) can also straightforwardly be extended to more densities, by introducing a sum over the environment multipoles. However, even for a single environment density, it is an advantage to

choose distributed origins for the multipole expansion. A distributed origin leads to an improved radius of convergence of the  $1/R$  Taylor series.<sup>79,80</sup> In practice, that means replacing  $\mathbf{R}_B$  with a sum over several, distributed *sites* in eqn (14)–(16). The most obvious sites are the atomic positions in subsystem B, but occasionally also bond-midpoints are used. We additionally allow this sum over sites ( $s$ ) to include multiple fragments.  $v^{\text{mult}}(\mathbf{r}_a)$  in eqn (46) thus becomes

$$v^{\text{mult}}(\mathbf{r}_a) = \sum_s v_s^{\text{mult}}(\mathbf{r}_a) = - \sum_s \sum_{k=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{sa}^{(k)}(\mathbf{r}_a) \mathbf{Q}_s^{0,(k)}. \quad (51)$$

For the polarization part,  $v^{\text{pol}}(\mathbf{r}_a)$ , it should additionally be noted that with multiple environment subsystems the electric fields in eqn (31) and thus also eqn (46) need to be augmented with a field due to the multipoles ( $\mathcal{E}_s^{\text{mult}}$ ). That means, the total field on a given site  $s$  becomes

$$\mathcal{E}_{A,s}^{\text{tot}}[\rho_A] = \mathcal{E}_s^{\text{mult}} + \mathcal{E}_{A,s}^e[\rho_A] + \mathcal{E}_{A,s}^n. \quad (52)$$

The fields from the electrons and nuclei in the active subsystem A is defined as in eqn (29) and (30), but with interaction operators involving sites  $\mathbf{R}_s$  rather than  $\mathbf{R}_B$ . Further, the induced dipole in eqn (48) will be modified to<sup>81</sup>

$$\boldsymbol{\mu}_s^{\text{ind}}[\rho_A] = \boldsymbol{\alpha}_s^0 \left( \mathcal{E}_{A,s}^{\text{tot}}[\rho_A] + \sum_{t \neq s} \mathbf{T}_{st}^{(2)} \boldsymbol{\mu}_t^{\text{ind}}[\rho_A] \right), \quad (53)$$

where the total field,  $\mathcal{E}_{A,s}^{\text{tot}}$ , was given above (*cf.* eqn (52)). The last term is the field from the induced dipoles from all other sites. Applequist and co-workers showed how eqn (53) can be rewritten into a matrix equation<sup>81</sup>

$$\boldsymbol{\mu}_s^{\text{ind}}[\rho_A] = \sum_t \mathbf{R}_{ts} \mathcal{E}_{A,s}^{\text{tot}}[\rho_A], \quad (54)$$

where  $\mathbf{R}$  is the so-called response matrix

$$\mathbf{R} = \begin{pmatrix} \boldsymbol{\alpha}_1^{-1} & -\mathbf{T}_{12}^{(2)} & \dots & -\mathbf{T}_{1S}^{(2)} \\ -\mathbf{T}_{21}^{(2)} & \boldsymbol{\alpha}_2^{-1} & \dots & -\mathbf{T}_{2S}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{T}_{S1}^{(2)} & -\mathbf{T}_{S2}^{(2)} & \dots & \boldsymbol{\alpha}_S^{-1} \end{pmatrix}. \quad (55)$$

With eqn (54), we can write  $v^{\text{pol}}[\rho_A](\mathbf{r}_a)$  in eqn (48) as

$$v^{\text{pol}}[\rho_A](\mathbf{r}_a) = \sum_s v_s^{\text{pol}}[\rho_A](\mathbf{r}_a) = - \sum_s (\boldsymbol{\mu}_s^{\text{ind}}[\rho_A])^T \mathcal{E}_{A,s}^e(\mathbf{r}_a). \quad (56)$$

The total PE-type potential can, hence, be summarized as

$$\begin{aligned} v^{\text{PE}}[\rho_A](\mathbf{r}_a) &= v^{\text{mult}}(\mathbf{r}_a) + v^{\text{pol}}[\rho_A](\mathbf{r}_a) \\ &= \sum_s (v_s^{\text{mult}}(\mathbf{r}_a) + v_s^{\text{pol}}[\rho_A](\mathbf{r}_a)). \end{aligned} \quad (57)$$

**Table 1** Overview of the main differences between different PE-type embedding methods. The method abbreviations are direct reaction field (DRF), polarizable embedding (PE), molecular mechanics with polarization (MMPol), and effective fragment potential (EFP). The last method is the atomic multipole optimized energetics for biomolecular applications (AMOEBA), developed originally for pure classical simulations.

Method	Multipoles	Polarizabilities	Comments
DRF <sup>82,83</sup>	Charges	Isotropic	Localized charges and polarizabilities chosen to reproduce gas-phase solvent dipole and polarizabilities (with Thole damping).
PE <sup>84</sup>	Higher-order	Anisotropic	Local multipoles and polarizabilities obtained with QM calculation using LoProp. <sup>85</sup>
MMPol <sup>86</sup>	Charges	Isotropic	AMBER charges and polarizabilities from ref. 81 and 87.
EFP <sup>68,88</sup>	Higher-order	Anisotropic	Localized multipoles and polarizabilities from QM calculations. Multipoles from DMA and polarizabilities from localized orbitals. May contain additional terms, <i>i.e.</i> exchange repulsion, charge transfer, short-range correlation, as well as dispersion terms.
AMOEBA <sup>89</sup>	Higher-order	Isotropic	Parameters from AMOEBA, <i>i.e.</i> DMA-based local multipoles and (occasionally modified) polarizabilities from Thole.

In the literature, many different classical polarizable models based on eqn (57) have been reported. The main difference between these models is the employed parameters used to define the operators in eqn (57), *e.g.* what type of multipole expansion, their truncation, and the form of the polarization terms. Several methods employ the distributed multipole analysis (DMA) by Stone<sup>90</sup> for the multipoles. For the polarizabilities, it should be noted that calculating the point-dipole interaction model in eqn (53) can lead to divergence for short distances, which has been denoted the *polarizability catastrophe*. Thole introduced a damped form of the point-dipole interaction, which reduces the issue,<sup>91</sup> and this scheme is employed in some PE models. A very brief summary is provided in Table 1 (note that we do not aim at a complete list here).

### 3 Embedding for local optical properties

To shed light on the incorporation of embedding contributions to local response calculations, we first briefly summarize the linear-response (or LR-TD-DFT) equations for an isolated (vacuum) system. We use one of many possible derivations,<sup>21,23,92</sup> which is described in more detail in ref. 93. More general derivations also valid for non-variational correlated wave functions<sup>23,94</sup> have also been employed for PE<sup>95-97</sup> as well as density-based embedding<sup>35,36</sup> models, and we refer to the literature for these cases.

In vacuum the electronic excitation can be described as a response to an external, time-dependent (electric) field

$$V(\mathbf{r}, t) = \frac{1}{2} (V_0(\mathbf{r})e^{-i\omega t} + V_0^*(\mathbf{r})e^{i\omega t}), \quad (58)$$

where  $\omega$  is the frequency of the electric field and  $V_0$  quantifies its strength. Since we generally work in a Kohn–Sham–DFT framework for the active subsystem, we assume a single-determinant wave function, comprised of Kohn–Sham orbitals that are eigenfunction to the Kohn–Sham operator,  $\hat{f}_{\text{iso}}^{(0)}$ . Our interest here is in the time-dependent Kohn–Sham operator described by adding  $V(\mathbf{r}, t)$  to the Kohn–Sham operator for the isolated system  $\hat{f}_{\text{iso}}^{(0)}$

$$\hat{f}(t) = \hat{f}_{\text{iso}}^{(0)} + V(\mathbf{r}, t). \quad (59)$$

At this point, it should be noted that, although we until now exclusively worked in a density-based formalism, the density can be directly related to the Kohn–Sham orbitals through the *density matrix*,  $\mathbf{P}$ . This relation is

$$\rho(\mathbf{r}) = \sum_{pq} P_{pq} \phi_p(\mathbf{r}) \phi_q^*(\mathbf{r}) \quad (60)$$

where  $\phi_p$  and  $\phi_q$  are atomic or molecular orbitals.

By means of the density matrix, the time-dependent Kohn–Sham equations can be formulated as

$$i \frac{\partial \mathbf{P}}{\partial t} = [\mathbf{F}, \mathbf{P}], \quad (61)$$

where  $\mathbf{F}$  is the matrix representation of the Kohn–Sham operator. We assume the response of the density matrix elements to be linear and immediate, so that the perturbed matrix elements read

$$P_{pq} = P_{pq}^{(0)} + P_{pq}^{(1)} = P_{pq}^{(0)} + \frac{1}{2} \left( d_{pq} e^{-i\omega t} + d_{pq}^* e^{i\omega t} \right). \quad (62)$$

We will now assume the orbitals to be the canonical orbitals of the unperturbed system, *i.e.*  $P_{ij}^{(0)} = \delta_{ij}$  and  $P_{ia}^{(0)} = P_{ai}^{(0)} = P_{ab}^{(0)} = 0$ , where  $i$  and  $j$  refer to occupied and  $a$  and  $b$  to virtual canonical orbitals in the unperturbed system. Next to the unperturbed components  $F_{pq}^{(0)}$ , the perturbed Kohn–Sham-matrix elements ( $F_{pq}$ ) also include the effect of the external time-dependent field  $V_{pq}(t)$  as well as of the first-order response in the density matrix.

$$F_{pq} = F_{pq}^{(0)} + F_{pq}^{(1)} = F_{pq}^{(0)} + V_{pq}(t) + \sum_{st} \frac{\partial F_{pq}^{(0)}}{\partial P_{st}} P_{st}^{(1)}. \quad (63)$$

Inserting eqn (63) and (62) into eqn (61) yields at resonance conditions (*i.e.* assuming  $\langle \phi_p | V_0 | \phi_q \rangle = 0$ ) the eigenvalue problem

$$\left[ \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^* & -\mathbf{A}^* \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \mathbf{0}. \quad (64)$$

Here, we have defined  $d_{ai} = X_{ai}$  and  $d_{ia} = Y_{ia}$ , while **A** and **B** have the elements

$$A_{ai,bj} = \delta_{ij} \delta_{ab} (\varepsilon_a - \varepsilon_i) + B_{ai,jb} \quad (65)$$

$$B_{ai,bj} = \frac{\partial F_{ai}^{(0)}}{\partial P_{jb}}. \quad (66)$$

The quantities  $\varepsilon_a$  and  $\varepsilon_i$  are orbital energies of the unperturbed system. The eigenvalues of eqn (64) are the excitation energies  $\omega$  and its eigenvectors  $(\mathbf{X}^\omega \mathbf{Y}^\omega)^T$  define the transition density upon an electronic transition as

$$\delta\rho^\omega(\mathbf{r}) = \sum_{ia} (X_{ia}^\omega \phi_i(\mathbf{r}) \phi_a^*(\mathbf{r}) + Y_{ia}^\omega \phi_a(\mathbf{r}) \phi_i^*(\mathbf{r})). \quad (67)$$

With the vacuum equations in place, we can now investigate the effect of adding an embedding potential to  $\hat{f}_{\text{iso}}^{(0)}$ , so that

$$\hat{f}_{\text{tot}}^{(0)} = \hat{f}_{\text{iso}}^{(0)} + \nu_{\text{emb}}, \quad (68)$$

where  $\nu_{\text{emb}}$  is either in form of eqn (44) or (57) for FDE or PE, respectively. Accordingly, a matrix element of the unperturbed Kohn–Sham matrix is defined as

$$F_{\text{tot},pq}^{(0)} = F_{\text{iso},pq}^{(0)} + V_{\text{emb},pq}, \quad (69)$$

with the matrix elements

$$V_{\text{emb},pq} = \langle \phi_p | \nu^{\text{emb}} | \phi_q \rangle, \quad (70)$$

being the only difference from the isolated system.

If we carry out the same manipulations as above, but with  $\hat{f}_{\text{iso}}^{(0)}$  replaced by  $\hat{f}_{\text{tot}}^{(0)}$  (cf. eqn (68)), we likewise obtain an eigenvalue of the form of eqn (64). However, **A** and **B** terms will be altered through (i) altered orbital energies due to introduction of the embedding potential in the unperturbed Kohn–Sham matrix, entering directly in eqn (65) and (ii) an addition to the B-term in eqn (66) with no equivalent in the eigenvalue equations for the isolated system

$$B_{ai,jb}^{\text{emb}} = \frac{\partial V_{\text{emb},ai}}{\partial P_{bj}}. \quad (71)$$

For FDE, this term is obtained by<sup>98,99</sup> employing eqn (44) in eqn (71)

$$\begin{aligned} B_{ai,jb}^{\text{FDE}} &= \left\langle \phi_a(\mathbf{r}_a) \phi_b(\mathbf{r}'_a) \left| \frac{\delta \nu_A^{\text{FDE}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} \right| \phi_i(\mathbf{r}_a) \phi_j(\mathbf{r}'_a) \right\rangle \\ &= \int d\mathbf{r}_a \int d\mathbf{r}'_a \rho_{ai}^t(\mathbf{r}_a) \frac{\delta \nu_A^{\text{FDE}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} \rho_{bj}^t(\mathbf{r}'_a), \end{aligned} \quad (72)$$

where  $\rho_{ai}^t(\mathbf{r}) = \phi_a^*(\mathbf{r})\phi_i(\mathbf{r})$ . The kernel can be obtained as<sup>67,98</sup>

$$\frac{\delta v_A^{\text{FDE}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} = \frac{\delta^2 E^{\text{xc}}[\rho_{\text{tot}}]}{\delta \rho_{\text{tot}}(\mathbf{r}_a)\delta \rho_{\text{tot}}(\mathbf{r}'_a)} - \frac{\delta^2 E^{\text{xc}}[\rho_A]}{\delta \rho_A(\mathbf{r}_a)\delta \rho_A(\mathbf{r}'_a)} + \frac{\delta^2 T_s[\rho_{\text{tot}}]}{\delta \rho_{\text{tot}}(\mathbf{r}_a)\delta \rho_{\text{tot}}(\mathbf{r}'_a)} - \frac{\delta^2 T_s[\rho_A]}{\delta \rho_A(\mathbf{r}_a)\delta \rho_A(\mathbf{r}'_a)}, \quad (73)$$

where  $\rho_{\text{tot}} = \sum_I \rho_I$  is the electron density of the complete system and  $\rho_A$  that of the active subsystem. The  $E^{\text{xc}}$  and  $T_s$  terms are the exchange–correlation and kinetic energy functionals. Note that the embedding contribution to the LR-TD-DFT kernel contains no Coulomb term, as the functional derivative of eqn (50) with respect to  $\rho_A$  is zero for a frozen environmental density  $\rho_{\text{env}}$ . The remaining QM embedding contributions to the response kernel are typically small<sup>29,100</sup> and can often be neglected.<sup>29</sup>

In PE, we arrive through an analogous derivation (see Appendix A.2 for details) at

$$B_{ai,bj}^{\text{PE}} = \left\langle \phi_a(\mathbf{r}_a)\phi_b(\mathbf{r}'_a) \left| \frac{\delta v_A^{\text{PE}}[\rho_A(\mathbf{r}'_a)](\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} \right| \phi_i(\mathbf{r}_a)\phi_j(\mathbf{r}'_a) \right\rangle \\ = \int d\mathbf{r}_a \int d\mathbf{r}'_a \rho_{ai}^t(\mathbf{r}_a) \frac{\delta v_A^{\text{PE}}[\rho_A(\mathbf{r}'_a)](\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} \rho_{bj}^t(\mathbf{r}'_a), \quad (74)$$

with

$$\frac{\delta v_A^{\text{PE}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} = - \sum_t \sum_s \mathbf{T}_{at}^{(1)}(\mathbf{r}'_a) \mathbf{R}_{ts} \mathbf{T}_{as}^{(1)}(\mathbf{r}_a). \quad (75)$$

Eqn (74) is sometimes said to describe the *dynamical response* of the environment due to the perturbation.<sup>80,84</sup> It can also be understood as an (approximate) description of *differential polarization*. We define differential polarization as *mutual response* of the environmental and active subsystem densities to a local change of density in the active subsystem caused by an electronic excitation.

Even though we are mainly concerned with embedding models combined with linear-response methods, we also briefly comment on the alternative state-specific description of excited states. In this case, differential polarization may be accounted for by means of state-specific embedding potentials. That means every electronic state has a dedicated embedding potential, explicitly optimized to fit the excited state wave function or density. Within FDE with freeze-and-thaw cycles, this approach formally corresponds to incorporating differential polarization to infinite order (see also Section 4.1.2).<sup>40</sup> State-specific embedding models are not restricted to state-specific electronic structure methods for the local excited state, but can also be combined with response approaches for the local excitation.<sup>43</sup>



Within the PE models, the physical interpretation of the terms in eqn (74) and (75) are still discussed and compared to the state-specific optimization. In the framework of a continuum model (whose theoretical foundation is equivalent to the PE models<sup>39</sup>), it has been shown that embedded state-specific calculations and embedded response calculations cover the differential polarization in a rather different manner,<sup>101</sup> *i.e. via* static and dynamical contribution, respectively.<sup>101,102</sup> This difference originates from the assumption that the overall state can be described as a product wave function (see Section 2.1).

A topic related to the state-specific discussion is the applicability of PE for emission from an *excited state* using response methods for the active region with the *ground state* as reference state. This method has occasionally been questioned.<sup>103,104</sup> This concern arises due to the presence of non-linear embedding operators, leading to difference in results, expected for ground- and excited-state embedding calculations (such difference is not present for an exact state description). However, since the vertical emission and absorption energies are just energy differences between two electronic states at a given nuclear conformation, we see no reason, why one of the interpretations of this energy difference (the absorption energy) should be well described while the other (emission energy) is not. That means that emission energies may be obtained with the same PE-like embedding approaches as absorption energies. The only difference is then the nuclear conformation (or a set of these) the static calculations are performed in. Studies following this approach have been reported in ref. 52, 105 and 106.

## 4 Method comparison and extensions

Having shown the details of the PE and FDE models in a common framework above, we can now summarize the main formal differences (with reference to the underlying equations in the previous chapter). Key points are collected in Table 2. We also highlight extensions to both methods (also collected in Table 2). We further present a short summary of numerical studies performed with PE and/or FDE for selected target systems, that are small molecules and the GFP.

### 4.1 Formal comparison of the original methodologies and latest extensions

In Table 2, PE and FDE features are classified in the categories (i) classical part of the environment, (ii) QM embedding part of the environment, and (iii) differential polarization. Table 2 further lists possible extensions that address short-comings of the original forms of PE and FDE.

#### 4.1.1 Comparison

**4.1.1.1 Classical embedding contribution.** In FDE, the classical (Coulomb) part in eqn (50) is evaluated directly, employing the real-space density, typically expressed on an integration grid. The polarization is encoded as density changes upon mutual optimization (freeze-and-thaw) approaches. Contrary to this, the representation of the environment is discretized in PE models, employing a multipole expansion with localized

**Table 2** Formal comparison of polarizable embedding and frozen-density embedding approaches and selected extensions for local excitations.

Environment	Polarizable embedding	Frozen-density embedding
(i) Classical part ( $v^C$ , see eqn (50))	Multipole expansion with localized multipoles (cf. eqn (51)) and localized static polarizabilities (cf. eqn (56)).	Real-space densities and mutual optimization <i>via</i> freeze-and-thaw cycles.
(ii) QM part ( $v^{QM}$ , see eqn (43))	Neglected.	Orbital-free DFT (see ref. 25, 63 and 64).
Extensions	Tailored ECPs. <sup>107</sup>	Tailored ECPs or long-distance correction added to $v^{QM}$ in eqn (43). <sup>108,109</sup>
	Polarizable density embedding (PDE; corresponds to adding explicit density (eqn (50) with $\rho_{env}$ replaced by $\rho_{env}^{(0)}$ ) and a repulsive projection operator). <sup>110–113</sup>	Projection techniques. <sup>73,114–125</sup>
	Exchange repulsion terms in EFP (see ref. 88 and references therein).	MP2-based correction terms to exchange–correlation contribution of eqn (43). <sup>126</sup>
(iii) Differential polarization in response methods	Localized static polarizabilities (cf. eqn (74) and (75)).	Neglected.
Extensions		State-specific FDE potentials. <sup>40,43</sup> Effective local description of coupled FDE. <sup>67,127</sup>

multipoles to approximate eqn (50). The order of the expansion is usually up to quadru- or octopoles, but higher orders are possible. In PE approaches, the mutual polarization of the density is approximated through the induced dipole in the environment, obtained by solving eqn (53) self-consistently.

**4.1.1.2 Quantum mechanical embedding contribution.** Since FDE approaches originate from a full quantum description, QM contributions to the embedding energy are (approximately) covered. The PE models in the original formulation, do not include these effects, as discussed in Section 2. In FDE, these kinetic and exchange–correlation contributions (cf. eqn (43)) are often covered by approximate orbital-free DFT.<sup>25,63,64</sup> The resulting QM potential also contains repulsion terms,<sup>63,76,108</sup> which reduce the electron spill-out problem<sup>128</sup> for FDE compared to classical models.<sup>108</sup> Due to the approximate kinetic embedding functionals, however, electron spill-out effects can also occur in FDE.<sup>108,129,130</sup>

**4.1.1.3 Differential polarization.** The formulation of FDE is generally more rigorous than PE for ground states, but in the original linear-response FDE approach, the differential polarization is neglected. In the PE scheme this is included *via* localized static polarizabilities in the response kernel (eqn (75)).

**4.1.2 Extensions.** Within the last years, several extensions to the original FDE and PE models have been suggested. Some of these extensions

to PE are inspired by FDE and *vice versa*, so that the resulting methods can occasionally be considered hybrids. We describe those according to the categories in Table 2.

**4.1.2.1 Quantum-mechanical embedding contribution.** To reduce the electron spill-out problem, several extensions have been made to both PE and FDE. The FDE method (approximately) accounts for repulsive QM effects, and these effects have also been introduced to PE: in the so-called polarizable density embedding (PDE) methods, this has been done by adding explicit densities and using a projection-operator to mimic the repulsion for environment fragments in close proximity of the QM region (this is sometimes denoted exchange or Pauli repulsion). In contrast to the original FDE scheme, PDE employs projection techniques to enforce mutual orthogonality of the orbitals of the different subsystems.<sup>131</sup> Such orthogonality constraints have also been used to avoid the need of an orbital-free kinetic energy functional in density embedding approaches, often in combination with wave function methods for the central subsystem.<sup>73,114–125</sup> Projection-based embedding techniques even allow the construction of conjugated systems from subsystems.<sup>73</sup> In projection-based embedding schemes, the exchange–correlation contribution can significantly contribute to the error and MP2-based corrections have been suggested to reduce this error.<sup>126</sup> Projection-based density embedding schemes have further been extended to excited states.<sup>132–135</sup>

A somewhat more pragmatic way to reduce the electron spill-out is the use of additional potentials in the form of tailored effective core potentials (ECPs), as has been done for both PE<sup>107</sup> and FDE.<sup>109</sup> In the case of FDE, a long-distance correction to the kinetic energy functional has also been employed for this purpose.<sup>108</sup>

**4.1.2.2 Differential polarization in response methods.** The neglect of differential polarization in standard FDE response methods can be remedied through introduction of state-specific embedding potentials.<sup>40</sup> While conceptually easy, this approach requires excited-state electron densities, which are not generally easily accessible in practice, in particular for the embedded case using response theory for the active subsystem.<sup>43</sup> Further, the change of the environmental energy through different polarization has to be accounted for.<sup>40</sup>

An alternative to recover the differential polarization is the coupled FDE approach,<sup>67,99,127,136–141</sup> which attempts to reformulate the full response equations by means of subsystems. The original formulation of coupled FDE aimed at excitonically coupled subsystems,<sup>67,99,127,136–138</sup> *i.e.*, non-local excitations. However, also non-resonant effects can be formulated *via* an effective, frequency-dependent contribution to the eigenvalue equation in eqn (64).<sup>67</sup>

**4.1.2.3 Further extensions.** In Section 3, all contributions from the environment were exclusively included in **A** and **B** terms in eqn (65) and (66). However, the introduction of the external field has an additional effect, namely to induce a field within the environment, which is not accounted for by the modifications to **A** and **B**. Such local field effects may

be covered by so-called effective external field (EEF) approaches, which have been reported for PE<sup>60,142</sup> as well as a variant of FDE, denoted polarizable FDE<sup>143</sup> (see below). Note that the EEF effects do not alter the transition energies, but can have a significant effect on transition moments (and therefore also polarizabilities and oscillator strengths).<sup>142,143</sup>

The polarizable FDE method is a combination of FDE with a PE-like description of the environment.<sup>143,144</sup> Here, the rather expensive freeze-and-thaw cycles are replaced by a purely electrostatic PE-like treatment of the polarizability. For the corresponding linear-response equations, the introduction of the PE-like polarization operator incorporates differential polarization effects in the resulting response kernel.<sup>143</sup>

Another addition to the PE models concerns the employed polarizabilities: in the original PE approach, static polarizabilities are employed. However, an electronic excitation is induced by a frequency-dependent electromagnetic field (*cf.* Section 3) and it can be argued that this should entail use of frequency-dependent polarizabilities. This was done in ref. 145, although the effect is not very large for the investigated systems.

For PE additionally damped response methodologies are reported,<sup>146</sup> as well as application to higher-order response properties.<sup>147–150</sup> Finally, also real-time TD-DFT approaches are reported both for FDE<sup>151–153</sup> and PE<sup>154</sup> approaches, and the same holds for extensions to relativistic wave functions.<sup>29,52,53,155</sup>

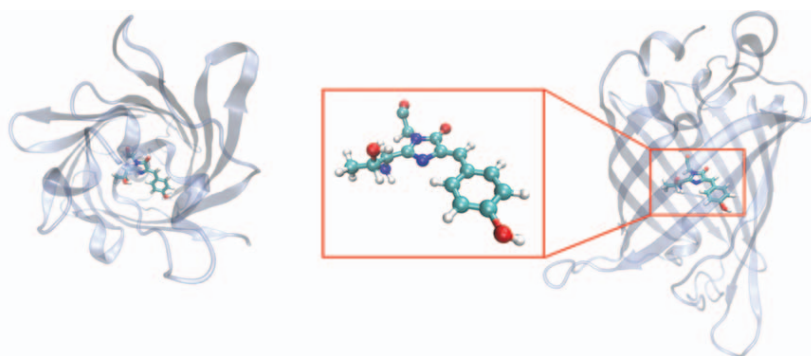
The methods discussed so far dealt with local excitations. It has, however, been argued that the inclusion of charge-transfer excitations between the dye and its environment can be required for good agreement with experiment.<sup>135,156–158</sup> This contribution may be included by exact subsystem LR-TD-DFT<sup>139–141</sup> or (combination of embedding schemes with) molecular fragmentation many-body expansion (MBE)-like approaches.<sup>135,159,160</sup> Another direction for PE models has been to derive correction terms to account for the asymmetry between state-specific and linear-response methods discussed in Section 3. These methods have been denoted corrected PE (see ref. 161 for an overview).

## 4.2 Numerical comparison

**4.2.1 Small model systems.** In an early study from 2006, Jacob *et al.*<sup>100</sup> compared a polarizable embedding model (DRF, see Table 1) to FDE for water in water. This study concludes that both approaches perform similarly for ground-state dipole and quadrupole moments, but differently for electronic excitation energies and frequency-dependent polarizabilities obtained *via* response methods. In a later review by Gomes and Jacob,<sup>66</sup> different studies on acetone in water are compared. The discussed studies employ either QM embedding or classical polarizable embedding, but do not compare them in a single study. It is concluded that differential polarization only has small effects on the solvent shift, as, for instance, found in ref. 31. This is in line with a good agreement to the experimental solvent shift obtained with FDE analogues neglecting this contribution compared to the experimental values.<sup>29,162</sup> Studies on microsolvated formaldehyde and *p*-nitroaniline reach the same conclusion (or at least find that ground-state polarization effects are larger than

the contribution due to differential polarization).<sup>163,164</sup> One should, however, be aware, that the impact of differential polarization is highly dependent on the molecule, size of solvent shell, transition and even the employed snapshot if dynamics are included: indeed, ref. 31 and 163 find an increase of the differential polarization when including two, four, and six water molecules in the same microsolvated model for *p*-nitroaniline. The analysis is expanded up to a 12 Å radius (239 water molecules) in the paper by Sneksov *et al.*<sup>31</sup> By averaging over 100 snapshots, they obtain an average contribution from differential polarization to the solvochromatic shift of *p*-nitroaniline in water of about 13%, while the contribution of ground-state polarization amounts to 19% and 21% for coupled-cluster and LR-TD-DFT methods (CAM-B3LYP), respectively. They further investigate *p*-nitrophenolate, where the impact of differential polarization on the excitation energies is above 60%. Another conclusion from ref. 31 is that different snapshots exhibit somewhat varying impact of differential polarization. This was also found within state-specific FDE in ref. 40. Here, the differential polarization for the  $\pi \rightarrow \pi^*$  transition of a *p*-nitroaniline–water system has been estimated for three microsolvated snapshots. The differential polarization for the different snapshots has been found in the ranges between 0.01–0.22 or –0.02–0.15 eV, depending on the description of the excited-state density.

**4.2.2 Example application: green fluorescent protein.** The green fluorescent protein (GFP) shown in Fig. 1<sup>165,168–173</sup> has become indispensable in non-invasive *in vivo* bioimaging.<sup>9,174,175</sup> The optical properties of GFP proteins are therefore actively studied experimentally as well as theoretically. These properties have origin in the  $\pi$ -conjugated chromophore *p*-hydroxybenzylideneimidazolidinone (HBDI), which in its isolated form is a non-fluorescent species due to a conical intersection reducing the excited-state lifetime.<sup>176–181</sup> However, as a covalently-bound, intrinsic part of the protein scaffold, the quantum yield increases by several orders of magnitude.<sup>170,182,183</sup> In equilibrium, the neutral form (A form) of the protein-embedded chromophore dominates over the anionic B form where the chromophore's phenol moiety



**Fig. 1** Neutral GFP chromophore (middle) extracted from the  $\beta$ -barrel wtGFP 1EMB structure<sup>165</sup> in top view (left) and front view (right). Graphics were created with Visual Molecular Dynamics 1.9.3.<sup>166,167</sup>

is deprotonated. Since these occur in a 6:1 ratio the experimental absorption spectrum displays two maxima, the first, higher maximum at about 395–398 nm corresponding to the A form and the second at about 475–478 nm corresponding to the B form.<sup>165,170,184–191</sup> There have been many studies to computationally reproduce the absorption properties of the system incorporating the structural changes in the chromophore pocket.<sup>17–19,107,133,137,159,160,192–212</sup> However, the size of the environment so far prohibits a full-system QM calculation.

In 2003, Marques *et al.*<sup>213</sup> introduced a QM/MM model for mechanically restricting the HBDI in the protein and subsequently performing a LR-TD-DFT calculation for the isolated chromophore. They were followed by several studies electrostatically embedding (EE) the chromophore.<sup>17,192,194,195,197,199–201,203,214–219</sup>

Beerepoot *et al.*<sup>204</sup> were able to show that the polarization effects of the environment in a distance of up to 20 Å influence the excitation energy of the chromophore by approximately 0.14 eV. In a study comparing QM/MM, QM/MMpol and (state-specific) FDE embedding schemes Daday *et al.*<sup>18</sup> confirmed the need of a protein description exceeding pure electrostatics. Several additional PE studies have been performed<sup>19,145,197,202,204,205,207,211</sup> as well as studies for different embedding schemes increasing the QM region.<sup>18,19,160,199–201,203,207,210,211,217</sup>

In the following, more recent studies will be shortly summarized. Selected excitation energies are condensed in Table 3: In ref. 207, a PE model was employed for studying the effect of different QM system sizes. The authors conclude that a PE model is essential to qualitatively capture environmental effects when comparatively small QM regions are employed. Additionally, the importance of using an EEF in the calculation of absorption intensities could be shown.<sup>207</sup> Further, Grabarek *et al.*<sup>210</sup> were able to show that including polarizability of the environment *via* PE allowed for a decreased size of the QM region in reproducing accurate results in comparison to electrostatic embedding (especially the intensities). Notably, a large effect of environmental polarization on the absorption energies is found for a GFP mutant in the same study. In an additional study, Grabarek *et al.*<sup>211</sup> were able to eliminate errors occurring in the previous PE study with diffuse function basis sets by including an effective-core potential (ECP) to avoid electron spill-out effects. This effect was also studied by Khah *et al.*<sup>107</sup> using PE(TD-HF) and PE(ADC(2)) models in combination with pseudopotentials, also distinctly stressing the impact of the electronic structure method for the active site, as similarly concluded earlier in a comparison of TD-DFT and CASPT2 methods.<sup>208</sup> Furthermore, the PE scheme was also adapted with a cost-effective polarizable protein potential (CP<sup>3</sup>) replacing the PE potential at a certain cut-off distance and achieving high accuracy while vastly decreasing the cost in the calculation of excitation energies of GFP.<sup>209</sup>

Another interesting point is the effect of ground-state polarization and differential polarization. Nifosi *et al.*<sup>208</sup> performed a study comparing QM/classical schemes with different treatment of the polarization contributions. These are non-polarized EE, ground-state polarized

**Table 3** Selected recent studies for GFP in their A form and B form employing quantum-derived embedding methods and their results for the 0 → 1 excitation energy dependent on the QM region size.

	Ref.	QM	Embedding	$\Delta E_A$ /eV	$\Delta E_B$ /eV	QM atoms A/B
Classical (polarizable) embedding	207	CAM-B3LYP/6-31G*	EE (AMBER)	—	3.15	—/96
		CAM-B3LYP/6-31G*	PE (B3LYP/6-31G*)	—	3.00	—/317
				—	2.98	—/35
					3.02	—/96
					2.97	—/317
	210	BHandHLYP/aug-cc-pVDZ	EE (AMBER)	3.28	3.10	40/39
				3.23	3.07	84/81
				3.19	3.00	141/154
		BHandHLYP/cc-pVDZ	PE (B3LYP/ANO-L-VDZP)	3.26	3.10	40/39
				3.24	3.08	84/81
				3.22	3.03	141/154
	211	BHandHLYP/aug-cc-pVDZ	PE(ECP) (B3LYP/ANO-L-VDZP)	3.22	—	40/—
	107	TD-HF/aug-cc-pVDZ	PE(ECP) (B3LYP/ANO-aug-cc-pVDZ)	—	3.32	—/35
		ADC(2)/aug-cc-pVDZ	PE(ECP) (B3LYP/ANO-aug-cc-pVDZ)	—	2.39	—/35
	209	CAM-B3LYP/pseg-2	AMOERA	—	3.48	—/39
		CAM-B3LYP/pseg-2	CP <sup>3</sup> (from PBE0/ANO-aug-cc-pVDZ)	—	3.14	—/39
	208	CAM-B3LYP/6-31 + G*	EE (AMBER99)	—	3.14	—/43
			MMpol (AMBER pol12)	—	3.23 (polGS)	—/43
					3.00 (polLR)	
					2.99 (polLR + SS)	
	SA-CASSCF/CASPT2/ANO	EE (AMBER99)	—	2.85	—/43	
		MMpol (AMBER pol12)	—	3.05 (polGS)	—/43	
				2.96 (polSS)		
				2.73 (polLR + SS)		

QM embedding	133	B3LYP/def2-TZVP	Projection-based embedding	3.27	—	49/—
			(B3LYP/def2-TZVP)	3.18		96/—
MBE-type	159	@B97X/6-31G*	EE (AMBER ff14SB)	3.38	3.07	36/35 <sup>a</sup>
			Polarized protein-specific charge (RESP from B3LYP/6-31G*)	3.31	2.88	36/35 <sup>a</sup>
	160	CAM-B3LYP/def2-SVP	sDFT + MBE (CAM-B3LYP/def2-SVP, $E^{\text{madd},\text{sc}}$ ; BLYP, $E^{\text{madd},\text{kin}}$ ; LLP91)	3.31	3.07	40/39 <sup>b</sup>
				3.29	3.04	152/151 <sup>b</sup>
165, 170 and 184–191	Experimental comparison		3.12–3.14	2.59–2.61	—	

<sup>a</sup>In the employed fragmentation approach the system is divided into smaller subsystems that are all calculated at QM level. However, in this case the value represents the number of atoms in the GFP chromophore. The whole system considered in the EE-GMFCC approach contains 776 atoms.

<sup>b</sup>Although technically all subsystems were determined at QM quality, in this case the value displays the number of atoms of the active subsystem in the FDE-IR-TD-DFT calculation. The full system included 278 atoms in the small system and 538 atoms in the large system.



embedding (polGS) as well as a linear-response description of differential polarization (polLR) *via* the MMPol method (see Table 1). The latter was further extended by a correction for the lack of contributions included in the state-specific approach (polLR + SS) through the correction terms briefly discussed at the end of Section 4.1.2. They show that with ground-state polarization, the excitation energies often exceeded the experimental results. A fine-tuning with differential and state-specific polarization is necessary since it is leading to a partial cancellation of the ground-state polarization effect.<sup>208</sup> Wen *et al.*<sup>133</sup> employed a projection-based density embedding model and confirmed the ground-state polarization as a major contribution. Subsequently, they performed a step-wise increase of the QM region including up to 96 atoms, which shows the residue-specific, systematic improvement in the agreement with the experiment.

Going beyond local embedding, a many-body expansion (MBE)-FDE hybrid achieved accurate results at second order (in the MBE) for the A form and at third order (in the MBE) for the B form. This approach also allowed an analysis of the specific residues dominant in response.<sup>160</sup> In another fragmentation study<sup>159</sup> the necessity of polarization effects in the stabilization of hydrogen bond interactions was revealed.

## 5 Summary and conclusions

We have presented a common derivation of the original frozen-density embedding (FDE) as well as polarizable embedding (PE) equations. We first compared to ground-state equations and then moved on to compare how optical properties are described within a linear-response framework.

This common framework enabled us to compare and discuss the two approaches' theoretical basis: The main differences for ground-state calculations are the incorporation of quantum-mechanical (QM) effects as well as a real-space representation of the electron densities in FDE, while the original form of PE uses a purely classical multipole description of the environment. For response properties an additional difference occurs: For PE, the embedding contribution to the response kernel contains only terms to approximate differential polarization effects, which are absent in the FDE response kernel. The latter only contains QM contributions. In summary, FDE accounts for QM effects and is therefore more rigorous and less prone to electron spill-out issues, while PE incorporates differential polarization effects in response calculations. Thus, the performance of the different methods will depend on how important electronic spill-out effects and differential polarization are, respectively.

We have surveyed a number of numerical investigations regarding the importance of differential polarization, but this contribution varies rather drastically with the system. Estimates ranges from above 60% of the solvchromatic shift for *p*-nitro-phenolate in water to negligible for acetone in water. We have further reviewed recent work on the green fluorescent protein, where embedding schemes that include polarization

effects seem indispensable for accurate calculation of absorption properties.

We have also discussed several recent extensions of the FDE and PE models; some of these extensions draw on experience from the other model, hoping to combine the advantages of both methodologies. This has in many cases been successful, yielding methods that better incorporate the correct physical description of the excitation process in a complex environment.

Within this book chapter, we neither discuss the nuclear and dynamic contributions to the optical spectra nor technical aspects related to the efficiency of the computational expressions. In a vertical approximation, introduction of nuclear effects requires a suitable set of conformations representing and covering all accessible conformations. These can be challenging to obtain, especially in case where quantum or excited-state descriptions are required for (part of) the system. This is, for example, the case, when targeting fluorescence spectra (see ref. 220–222 and references therein). Such studies indeed pose challenges to the described quantum-derived embedding methodologies, both regarding the required efficiency of spectra calculation, and the theoretical description. The latter holds particularly when vibronic effects of the spectrum need to be considered along with environmental effects. This is currently an active field of research (see ref. 223 and references therein), but considerably more work should be done in this direction.

## A Appendix

### A.1 Interaction operators and Taylor expansion of a potential

We define a distance between two vectors  $|\mathbf{R}_i - \mathbf{R}_j|$ . The Taylor expansion of  $|\mathbf{R}_i - \mathbf{R}_j|^{-1}$  around an expansion point,  $\mathbf{R}_O$ , using that

$$\left( \frac{\partial}{\partial R_{j,\alpha}} \frac{1}{|\mathbf{R}_i - \mathbf{R}_j|} \right)_{\mathbf{R}_j = \mathbf{R}_O} = - \left( \frac{\partial}{\partial R_{i,\alpha}} \frac{1}{|\mathbf{R}_i - \mathbf{R}_O|} \right) \quad (76)$$

gives

$$\begin{aligned} \frac{1}{|\mathbf{R}_i - \mathbf{R}_j|} &= \frac{1}{|\mathbf{R}_i - \mathbf{R}_O|} - \sum_{\alpha} \left( \frac{\partial}{\partial R_{i,\alpha}} \frac{1}{|\mathbf{R}_i - \mathbf{R}_O|} \right) (R_{j,\alpha} - R_{O,\alpha}) + \dots \\ &= T_{O_i}^{(0)} - \sum_{\alpha} T_{O_i,\alpha}^{(1)} (R_{j,\alpha} - R_{O,\alpha}) + \dots, \end{aligned} \quad (77)$$

where we have defined the interaction operators<sup>71</sup>  $T_{ji}^{(0)} = |\mathbf{R}_i - \mathbf{R}_j|^{-1}$ ,  $T_{ji,\alpha}^{(1)} = \frac{\partial}{\partial R_{i,\alpha}} |\mathbf{R}_i - \mathbf{R}_j|^{-1}$  and so on.

The Taylor expansion in eqn (77) can be written in a compact form if we use a three-dimensional multi-index  $k = (k_x, k_y, k_z)$ .<sup>80</sup> The norm of the multi-index is  $|k| = k_x + k_y + k_z$  and the  $k$ 'th power of the vectors

in eqn (76) and (77) is defined  $\mathbf{R}^k = x^{k_x} y^{k_y} z^{k_z}$ . We can then define a derivative

$$\frac{\partial^{|\mathbf{k}|}}{\partial \mathbf{R}^{\mathbf{k}}} = \frac{\partial^{k_x+k_y+k_z}}{\partial x^{k_x} \partial y^{k_y} \partial z^{k_z}} \quad (78)$$

The multi-index notation can thus be used to rewrite eqn (77) into

$$\begin{aligned} \frac{1}{|\mathbf{R}_i - \mathbf{R}_j|} &= \sum_{|\mathbf{k}|=0} \frac{(-1)^{|\mathbf{k}|}}{k!} \left( \frac{\partial^{|\mathbf{k}|}}{\partial \mathbf{R}_i^{\mathbf{k}}} \frac{1}{|\mathbf{R}_i - \mathbf{R}_0|} \right) (\mathbf{R}_i - \mathbf{R}_0)^{\mathbf{k}} \\ &= \sum_{|\mathbf{k}|=0} \frac{(-1)^{|\mathbf{k}|}}{k!} \mathbf{T}_{0i}^{(\mathbf{k})} (\mathbf{R}_i - \mathbf{R}_0)^{\mathbf{k}}. \end{aligned} \quad (79)$$

The interaction operator is defined as a tensor with Einstein summation over coordinate indices implied,  $\mathbf{T}_{ji}^{(\mathbf{k})} = \frac{\partial^{k_x+k_y+k_z}}{\partial x_i^{k_x} \partial y_i^{k_y} \partial z_i^{k_z}} |\mathbf{R}_i - \mathbf{R}_j|^{-1}$ .

### A.2 Polarizable embedding kernel in linear-response time-dependent density-functional theory

To derive the contributions of the PE operator to the LR-TD-DFT response matrix (eqn (64)), within the framework sketched in Section 3, we apply the matrix element of the PE operator in the form (cf. eqn (57))

$$V_{ai}^{\text{PE}} = \langle \phi_a(\mathbf{r}_a) | (v^{\text{mult}}(\mathbf{r}_a) + v^{\text{pol}}[\rho_A](\mathbf{r}_a)) | \phi_i(\mathbf{r}_a) \rangle \quad (80)$$

and insert it as  $V_{ai}^{\text{emb}}$  in eqn (63). The electrostatic part  $v^{\text{mult}}(\mathbf{r}_a)$  (cf. eqn (51)) is independent of the QM electron density therefore does not contribute to the response kernel. Thus, we can focus on the  $v^{\text{pol}}[\rho_A](\mathbf{r}_a)$  contribution; the matrix element in eqn (80) reads

$$V_{ai}^{\text{PE}}[\rho_A] = \langle \phi_a(\mathbf{r}_a) | v^{\text{pol}}[\rho_A](\mathbf{r}_a) | \phi_i(\mathbf{r}_a) \rangle. \quad (81)$$

Inserting eqn (81) into eqn (71) as  $V_{ai}^{\text{emb}} = V_{ai}^{\text{PE}}$  yields

$$\frac{\partial}{\partial P_{bj}} \langle \phi_a(\mathbf{r}_a) | v^{\text{pol}}[\rho_A](\mathbf{r}_a) | \phi_i(\mathbf{r}_a) \rangle = \langle \phi_a(\mathbf{r}_a) | \frac{\partial v^{\text{pol}}[\rho_A](\mathbf{r}_a)}{\partial P_{bj}} | \phi_i(\mathbf{r}_a) \rangle \quad (82)$$

The derivative is evaluated as a functional derivative: in general, the derivative of a functional  $G[g]$  with respect to the parameter  $\lambda$  can be expressed as,<sup>78</sup>

$$\frac{\partial G}{\partial \lambda} = \int \frac{\delta G}{\delta(g(x))} \frac{\partial g(x)}{\partial \lambda} dx \quad (83)$$

The derivative of  $v^{\text{pol}}$  in eqn (82) thus becomes

$$\frac{\partial v^{\text{pol}}[\rho_A](\mathbf{r}_a)}{\partial P_{bj}} = \int \frac{\delta v^{\text{pol}}[\rho_A(\mathbf{r}'_a)](\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} \frac{\partial \rho_A(\mathbf{r}'_a)}{\partial P_{bj}} d\mathbf{r}'_a, \quad (84)$$

where the kernel can be obtained by inserting the expression for  $v^{\text{pol}}$  (cf. eqn (54) and (56)) with

$$\begin{aligned} \frac{\delta v^{\text{pol}}[\rho_A(\mathbf{r}'_a)](\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}'_a)} &= - \sum_s \frac{\delta \boldsymbol{\mu}_s^{\text{ind}}[\rho_A(\mathbf{r}'_a)]}{\delta \rho_A(\mathbf{r}'_a)} \boldsymbol{\varepsilon}_{A,s}^e(\mathbf{r}_a) \\ &= - \sum_s \sum_t \left( \frac{\delta \mathcal{E}_{A,t}^e[\rho_A(\mathbf{r}'_a)]}{\delta \rho_A(\mathbf{r}'_a)} \right) \mathbf{R}_{ts} \boldsymbol{\varepsilon}_{A,s}^e(\mathbf{r}_a) \\ &= - \sum_s \sum_t \boldsymbol{\varepsilon}_{A,t}^e(\mathbf{r}'_a) \mathbf{R}_{ts} \boldsymbol{\varepsilon}_{A,s}^e(\mathbf{r}_a) = - \sum_s \sum_t \mathbf{T}_{at}^{(1)}(\mathbf{r}'_a) \mathbf{R}_{ts} \mathbf{T}_{as}^{(1)}(\mathbf{r}_a). \end{aligned} \quad (85)$$

Note, that only the electronic part of the electric field in eqn (52) contributes, as this is the only part dependent on  $\rho_A$ . We further realize that  $\boldsymbol{\varepsilon}_{A,t}^e(\mathbf{r}_a) = \mathbf{T}_{at}^{(1)}(\mathbf{r}_a)$ , as can be deduced from eqn (30) and (47), and explicitly write the  $\mathbf{r}_a$  dependence of  $\mathbf{T}_{at}^{(1)}(\mathbf{r}_a)$ . This form of the kernel given in eqn (75), thus, eqn (82) becomes

$$\begin{aligned} \langle \phi_a(\mathbf{r}_a) \left| \frac{\partial v^{\text{pol}}[\rho_A](\mathbf{r}_a)}{\partial P_{bj}} \right| \phi_i(\mathbf{r}_a) \rangle &= - \sum_s \sum_t \langle \phi_a(\mathbf{r}_a) \left| \int \mathbf{T}_{at}^{(1)}(\mathbf{r}'_a) \mathbf{R}_{ts} \frac{\partial \rho(\mathbf{r}'_a)}{\partial P_{bj}} d\mathbf{r}'_a \mathbf{T}_{as}^{(1)}(\mathbf{r}_a) \right| \phi_i(\mathbf{r}_a) \rangle. \quad (86) \\ &= - \sum_s \sum_t \langle \phi_a(\mathbf{r}_a) \phi_b(\mathbf{r}'_a) \left| \mathbf{T}_{at}^{(1)}(\mathbf{r}'_a) \mathbf{R}_{ts} \mathbf{T}_{as}^{(1)}(\mathbf{r}_a) \right| \phi_j(\mathbf{r}'_a) \phi_i(\mathbf{r}_a) \rangle. \end{aligned}$$

which is the equation given in eqn (74).

## Acknowledgements

EDH thanks The Villum Foundation, Young Investigator Program (grant no. 29412), the Swedish Research Council (grant no. 2019-04205), and Independent Research Fund Denmark (grant no. 0252-00002B) for support. CK acknowledges support from the Deutsche Forschungsgemeinschaft (DFG) through the Emmy Noether Young Group Leader Programme (project KO 5423/1-1).

## References

- 1 T. Terai and T. Nagano, *Curr. Opin. Chem. Biol.*, 2008, **12**, 515–521.
- 2 T. Ueno and T. Nagano, *Nat. Methods*, 2011, **8**, 642–645.
- 3 Z. Guo, S. Park, J. Yoon and I. Shin, *Chem. Soc. Rev.*, 2014, **43**, 16–29.
- 4 L. Y. Niu, Y. Z. Chen, H. R. Zheng, L. Z. Wu, C. H. Tung and Q. Z. Yang, *Chem. Soc. Rev.*, 2015, **44**, 6143–6160.
- 5 Y. Fu and N. S. Finney, *RSC Adv.*, 2018, **8**, 29051–29061.
- 6 M. Tian, Y. Ma and W. Lin, *Acc. Chem. Res.*, 2019, **52**, 2147–2157.
- 7 A. Åslund, *et al.*, *ACS Chem. Biol.*, 2009, **4**, 673–684.
- 8 M. Zimmer, *Chem. Rev.*, 2002, **102**, 759–781.

- 9 D. M. Chudakov, M. V. Matz, S. Lukyanov and K. A. Lukyanov, *Physiol. Rev.*, 2010, **90**, 1103–1163.
- 10 U. N. Morzan, D. J. Alonso De Armiño, N. O. Foglia, F. Ramírez, M. C. González Lebrero, D. A. Scherlis and D. A. Estrin, *Chem. Rev.*, 2018, **118**, 4071–4113.
- 11 A. V. Nemukhin, B. L. Grigorenko, M. G. Khrenova and A. I. Krylov, *J. Phys. Chem. B*, 2019, **123**, 6133–6149.
- 12 M. Rossano-Tapia and A. Brown, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1557.
- 13 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3093.
- 14 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227–249.
- 15 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
- 16 M. Bondanza, M. Nottoli, L. Cupellini, F. Lipparini and B. Mennucci, *Phys. Chem. Chem. Phys.*, 2020, **22**, 14433–14448.
- 17 C. Filippi, F. Buda, L. Guidoni and A. Sinicropi, *J. Chem. Theory Comput.*, 2012, **8**, 112–124.
- 18 C. Daday, C. Curutchet, A. Sinicropi, B. Mennucci and C. Filippi, *J. Chem. Theory Comput.*, 2015, **11**, 4825–4839.
- 19 T. Schwabe, M. T. Beerepoot, J. M. H. Olsen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2015, **17**, 2582–2588.
- 20 J. Olsen and P. Jørgensen, *J. Chem. Phys.*, 1984, **82**, 3235–3264.
- 21 M. E. Casida, *Recent Advances in Density Functional Methods*, 1995, pp. 155–192.
- 22 P. Norman, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20519–20535.
- 23 T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen and K. Ruud, *Chem. Rev.*, 2012, **112**, 543–631.
- 24 P. Norman, K. Ruud and T. Saue, *Principles and practices of molecular properties: Theory, modeling and simulations*, John Wiley & Sons Ltd., 2017.
- 25 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050–8053.
- 26 T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.*, 1996, **248**, 71–76.
- 27 N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 1999, **110**, 7677–7688.
- 28 T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 2002, **116**, 42–54.
- 29 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353–5362.
- 30 T. A. Wesolowski, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2008, **77**, 12504.
- 31 K. Sneskov, T. Schwabe, O. Christiansen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2011, **13**, 18551–18560.
- 32 C. Huang, M. Pavone and E. A. Carter, *J. Chem. Phys.*, 2011, **134**, 154110.
- 33 C. Huang and E. A. Carter, *J. Chem. Phys.*, 2011, **135**, 194104.
- 34 T. Schwabe, K. Sneskov, J. M. Haugaard Olsen, J. Kongsted, O. Christiansen and C. Hättig, *J. Chem. Theory Comput.*, 2012, **8**, 3274–3283.
- 35 S. Höfener, A. Severo Pereira Gomes and L. Visscher, *J. Chem. Phys.*, 2012, **136**, 44104.
- 36 S. Höfener and L. Visscher, *J. Chem. Phys.*, 2012, **137**, 204120.
- 37 S. Höfener, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2013, **139**, 104106.
- 38 E. D. Hedegård, N. H. List, H. J. A. Jensen and J. Kongsted, *J. Chem. Phys.*, 2013, **139**, 44101.
- 39 N. H. List, H. J. A. Jensen, J. Kongsted and E. D. Hedegård, *Adv. Quantum Chem.*, 2013, **66**, 195–238.
- 40 C. Daday, C. König, O. Valsson, J. Neugebauer and C. Filippi, *J. Chem. Theory Comput.*, 2013, **9**, 2355–2367.
- 41 E. D. Hedegård, H. J. A. Jensen and J. Kongsted, *Int. J. Quantum Chem.*, 2014, **114**, 1102–1107.

- 
- 42 S. Höfener, *J. Comput. Chem.*, 2014, **35**, 1716–1724.
- 43 C. Daday, C. König, J. Neugebauer and C. Filippi, *ChemPhysChem*, 2014, **15**, 3892.
- 44 F. Libisch, C. Huang and E. A. Carter, *Acc. Chem. Res.*, 2014, **47**, 2768–2775.
- 45 K. Yu, F. Libisch and E. A. Carter, *J. Chem. Phys.*, 2015, **143**, 102806.
- 46 E. D. Hedegård, J. M. H. Olsen, S. Knecht, J. Kongsted and H. J. A. Jensen, *J. Chem. Phys.*, 2015, **142**, 114113.
- 47 T. Dresselhaus, J. Neugebauer, S. Knecht, S. Keller, Y. Ma and M. Reiher, *J. Chem. Phys.*, 2015, **142**, 44111.
- 48 T. A. Wesolowski, S. Shedge and X. Zhou, *Chem. Rev.*, 2015, **115**, 5891–5928.
- 49 E. D. Hedegård and M. Reiher, *J. Chem. Theory Comput.*, 2016, **12**, 4242–4253.
- 50 S. Prager, A. Zech, F. Aquilante, A. Dreuw and T. A. Wesolowski, *J. Chem. Phys.*, 2016, **144**, 204103.
- 51 R. Guareschi, H. Zulfikri, C. Daday, F. M. Floris, C. Amovilli, B. Mennucci and C. Filippi, *J. Chem. Theory Comput.*, 2016, **12**, 1674–1683.
- 52 K. Krause, M. Bauer and W. Klopper, *J. Chem. Theory Comput.*, 2016, **12**, 2853–2860.
- 53 E. D. Hedegård, R. Bast, J. Kongsted, J. M. H. Olsen and H. J. A. Jensen, *J. Chem. Theory Comput.*, 2017, **13**, 2870–2880.
- 54 K. Yu, C. M. Krauter, J. M. Dieterich and E. A. Carter, *Density and Potential Functional Embedding: Theory and Practice*, John Wiley & Sons, Hoboken, NJ, 2017.
- 55 J. Heuser and S. Höfener, *J. Chem. Phys.*, 2018, **148**, 141101.
- 56 M. Böckers and J. Neugebauer, *J. Chem. Phys.*, 2018, **149**, 74102.
- 57 M. Scheurer, M. F. Herbst, P. Reinholdt, J. M. H. Olsen, A. Dreuw and J. Kongsted, *J. Chem. Theory Comput.*, 2018, **14**, 4870–4883.
- 58 N. Schieschke, T. Bodenstein and S. Höfener, *Mol. Phys.*, 2020, **118**, 1665726.
- 59 N. Schieschke, T. Bodenstein and S. Höfener, *J. Chem. Phys.*, 2021, **154**, 84120.
- 60 N. H. List, J. M. H. Olsen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2016, **18**, 20234–20250.
- 61 D. Loco, L. Lagardère, O. Adjoua and J. P. Piquemal, *Acc. Chem. Res.*, 2021, **54**, 2812–2822.
- 62 F. Lipparini and B. Mennucci, *Chem. Phys. Rev.*, 2021, **2**, 041303.
- 63 C. R. Jacob and J. Neugebauer, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 325–362.
- 64 A. Krishtal, D. Sinha, A. Genova and M. Pavanello, *J. Phys.: Condens. Matter*, 2015, **27**, 183202.
- 65 Q. Sun and G. K. L. Chan, *Acc. Chem. Res.*, 2016, **49**, 2705–2712.
- 66 A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2012, **108**, 222–277.
- 67 J. Neugebauer, *Phys. Rep.*, 2010, **489**, 1–87.
- 68 M. S. Gordon, L. Slipchenko, H. Li and J. H. Jensen, *Annu. Rep. Comput. Chem.*, 2007, **3**, 177–193.
- 69 J. Gao, D. G. Truhlar, Y. Wang, M. J. Mazack, P. Löffler, M. R. Provorose and P. Rehak, *Acc. Chem. Res.*, 2014, **47**, 2837–2845.
- 70 C. Steinmann, P. Reinholdt, M. S. Nørby, J. Kongsted and J. M. H. Olsen, *Int. J. Quantum Chem.*, 2019, **119**, e25717.
- 71 A. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, 2013.
- 72 K. Szalewicz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 254–272.
- 73 D. V. Chulhai and L. Jensen, *J. Chem. Theory Comput.*, 2015, **11**, 3080–3088.

- 
- 74 J. Gao, *J. Phys. Chem. B*, 1997, **101**, 657–663.  
75 W. Xie and J. Gao, *J. Chem. Theory Comput.*, 2007, **3**, 1890–1900.  
76 S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher and M. Reiher, *J. Chem. Phys.*, 2010, **132**, 164101.  
77 D. G. Artiukhin, C. R. Jacob and J. Neugebauer, *J. Chem. Phys.*, 2015, **142**, 234101.  
78 R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1989.  
79 R. Ahlrichs, *Theor. Chim. Acta*, 1976, **41**, 7–15.  
80 J. M. H. Olsen and J. Kongsted, in *Advances in Quantum Chemistry*, Elsevier, 2011, vol. 61, pp. 107–143.  
81 J. Applequist, J. R. Carl and K. K. Fung, *J. Am. Chem. Soc.*, 1972, **94**, 2952–2960.  
82 L. Jensen, P. T. Van Duijnen and J. G. Snijders, *J. Chem. Phys.*, 2003, **119**, 12998–13006.  
83 L. Jensen, P. T. Van Duijnen and J. G. Snijders, *J. Chem. Phys.*, 2003, **118**, 514–521.  
84 J. M. Olsen, K. Aidas and J. Kongsted, *J. Chem. Theory Comput.*, 2010, **6**, 3721–3734.  
85 L. Gagliardi, R. Lindh and G. Karlström, *J. Chem. Phys.*, 2004, **121**, 4494–4500.  
86 M. A. Thompson, *J. Phys. Chem.*, 1996, **100**, 14492–14507.  
87 L. X. Dang, J. E. Rice, J. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1991, **113**, 2481–2486.  
88 M. S. Gordon, Q. A. Smith, P. Xu and L. V. Slipchenko, *Annu. Rev. Phys. Chem.*, 2013, **64**, 553–578.  
89 D. Loco, É. Polack, S. Caprasecca, L. Lagardère, F. Lipparini, J. P. Piquemal and B. Mennucci, *J. Chem. Theory Comput.*, 2016, **12**, 3654–3661.  
90 A. Stone and M. Alderton, *Mol. Phys.*, 1985, **56**, 1047–1064.  
91 B. Thole, *Chem. Phys.*, 1981, **59**, 341–350.  
92 R. Bauernschmitt and R. Ahlrichs, *Chem. Phys. Lett.*, 1996, **256**, 454–464.  
93 A. Dreuw and M. Head-Gordon, *Chem. Rev.*, 2005, **105**, 4009–4037.  
94 O. Christiansen, P. Jørgensen and C. Hättig, *Int. J. Quantum Chem.*, 1998, **68**, 1–52.  
95 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *Mol. Phys.*, 2002, **100**, 1813–1828.  
96 J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *J. Chem. Phys.*, 2003, **118**, 1620–1633.  
97 K. Sneskov, T. Schwabe, J. Kongsted and O. Christiansen, *J. Chem. Phys.*, 2011, **134**, 104108.  
98 M. E. Casida and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2004, **96**, 577–588.  
99 C. König, N. Schlüter and J. Neugebauer, *J. Chem. Phys.*, 2013, **138**, 34104.  
100 C. R. Jacob, J. Neugebauer, L. Jensen and L. Visscher, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2349–2359.  
101 R. Cammi, S. Corni, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, 2005, **122**, 104513.  
102 T. Schwabe, *J. Chem. Phys.*, 2016, **145**, 154105.  
103 R. Improta, G. Scalmani, M. J. Frisch and V. Barone, *J. Chem. Phys.*, 2007, **127**, 74504.  
104 J. Sjöqvist, M. Linares, K. V. Mikkelsen and P. Norman, *J. Phys. Chem. A*, 2014, **118**, 3419–3428.  
105 C. Gustafsson, M. Linares and P. Norman, *J. Phys. Chem. A*, 2020, **124**, 875–888.

- 
- 106 D. Loco, N. Gelfand, S. Jurinovich, S. Protti, A. Mezzetti and B. Mennucci, *J. Phys. Chem. A*, 2018, **122**, 390–397.
- 107 A. Marefat Khah, P. Reinholdt, J. M. H. Olsen, J. Kongsted and C. Hättig, *J. Chem. Theory Comput.*, 2020, **16**, 1373–1381.
- 108 C. R. Jacob, S. M. Beyhan and L. Visscher, *J. Chem. Phys.*, 2007, **126**, 234116.
- 109 R. S. Trefß, C. Hättig and S. Höfener, *J. Chem. Theory Comput.*, 2022, **18**, 1737–1747.
- 110 J. M. H. Olsen, C. Steinmann, K. Ruud and J. Kongsted, *J. Phys. Chem. A*, 2015, **119**, 6928.
- 111 P. Reinholdt, J. Kongsted and J. M. H. Olsen, *J. Phys. Chem. Lett.*, 2017, **8**, 5949–5958.
- 112 P. Reinholdt, F. K. Jørgensen, J. Kongsted and J. M. H. Olsen, *J. Chem. Theory Comput.*, 2020, **16**, 5999–6006.
- 113 D. Hršak, J. M. H. Olsen and J. Kongsted, *J. Chem. Theory Comput.*, 2018, **14**, 1351–1360.
- 114 J. D. Goodpaster, N. Ananth, F. R. Manby and T. F. Miller, *J. Chem. Phys.*, 2010, **133**, 84103.
- 115 F. R. Manby, M. Stella, J. D. Goodpaster and T. F. Miller, *J. Chem. Theory Comput.*, 2012, **8**, 2564–2568.
- 116 Y. G. Khait and M. R. Hoffmann, *Annu. Rep. Comput. Chem.*, 2012, **8**, 53–70.
- 117 B. Hégyely, P. R. Nagy, G. G. Ferenczy and M. Kállay, *J. Chem. Phys.*, 2016, **145**, 64107.
- 118 D. V. Chulhai and J. D. Goodpaster, *J. Chem. Theory Comput.*, 2017, **13**, 1503–1508.
- 119 T. Culpitt, K. R. Brorsen and S. Hammes-Schiffer, *J. Chem. Phys.*, 2017, **146**, 100201.
- 120 F. Ding, F. R. Manby and T. F. Miller, *J. Chem. Theory Comput.*, 2017, **13**, 1605–1615.
- 121 S. J. Lee, M. Welborn, F. R. Manby and T. F. Miller, *Acc. Chem. Res.*, 2019, **52**, 1359–1368.
- 122 D. S. Graham, X. Wen, D. V. Chulhai and J. D. Goodpaster, *J. Chem. Theory Comput.*, 2020, **16**, 2284–2295.
- 123 D. S. Graham, X. Wen, D. V. Chulhai and J. D. Goodpaster, *J. Chem. Phys.*, 2022, **156**, 54112.
- 124 M. Bensberg and J. Neugebauer, *J. Chem. Phys.*, 2019, **150**, 184104.
- 125 M. Bensberg and J. Neugebauer, *J. Chem. Phys.*, 2019, **150**, 214106.
- 126 J. D. Goodpaster, T. A. Barnes, F. R. Manby and T. F. Miller, *J. Chem. Phys.*, 2014, **140**, 18A507.
- 127 J. Neugebauer, C. Curutchet, A. Muñoz-Losa and B. Mennucci, *J. Chem. Theory Comput.*, 2010, **6**, 1843–1851.
- 128 A. Laio, J. VandeVondele and U. Rothlisberger, *J. Chem. Phys.*, 2002, **116**, 6941–6947.
- 129 E. V. Stefanovich and T. N. Truong, *J. Chem. Phys.*, 1996, **104**, 2946–2955.
- 130 M. Duřak and T. A. Wesolowski, *J. Chem. Phys.*, 2006, **124**, 1–6.
- 131 S. Huzinaga and A. A. Cantu, *J. Chem. Phys.*, 1971, **55**, 5543–5549.
- 132 D. V. Chulhai and L. Jensen, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21032–21039.
- 133 X. Wen, D. S. Graham, D. V. Chulhai and J. D. Goodpaster, *J. Chem. Theory Comput.*, 2020, **16**, 385–398.
- 134 N. Niemeyer, J. Tölle and J. Neugebauer, *J. Chem. Theory Comput.*, 2020, **16**, 3104–3120.
- 135 L. Scholz, J. Tölle and J. Neugebauer, *Int. J. Quantum Chem.*, 2020, **120**, e26213.
- 136 J. Neugebauer, *J. Chem. Phys.*, 2007, **126**, 134116.



- 
- 137 J. Neugebauer, *J. Chem. Phys.*, 2009, **131**, 84104.  
138 S. Höfener and L. Visscher, *J. Chem. Theory Comput.*, 2016, **12**, 549–557.  
139 J. Tölle, M. Böckers and J. Neugebauer, *J. Chem. Phys.*, 2019, **150**, 181101.  
140 J. Tölle, M. Böckers, N. Niemeyer and J. Neugebauer, *J. Chem. Phys.*, 2019, **151**, 174109.  
141 J. Tölle and J. Neugebauer, *J. Phys. Chem. Lett.*, 2022, **13**, 1003–1018.  
142 N. H. List, H. J. A. Jensen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2016, **18**, 10070–10080.  
143 A. K. Harshan, M. J. Bronson and L. Jensen, *J. Chem. Theory Comput.*, 2022, **18**, 380–393.  
144 P. P. Pal, P. Liu and L. Jensen, *J. Chem. Theory Comput.*, 2019, **15**, 6588–6596.  
145 M. S. Nørby, O. Vahtras, P. Norman and J. Kongsted, *Mol. Phys.*, 2017, **115**, 39–47.  
146 M. N. Pedersen, E. D. Hedegård, J. M. H. Olsen, J. Kauczor, P. Norman and J. Kongsted, *J. Chem. Theory Comput.*, 2014, **10**, 1164–1171.  
147 K. Sneskov, J. M. H. Olsen, T. Schwabe, C. Hättig, O. Christiansen and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7567–7576.  
148 D. Hršak, L. Holmegaard, A. S. Poulsen, N. H. List, J. Kongsted, M. P. Denofrio, R. Erra-Balsells, F. M. Cabrerizo, O. Christiansen and P. R. Ogilby, *Phys. Chem. Chem. Phys.*, 2015, **17**, 12090–12099.  
149 A. H. Steindal, M. T. Beerepoot, M. Ringholm, N. H. List, K. Ruud, J. Kongsted and J. M. H. Olsen, *Phys. Chem. Chem. Phys.*, 2016, **18**, 28339–28352.  
150 S. I. Lu, *J. Phys. Chem. A*, 2019, **123**, 5334–5340.  
151 A. Krishtal, D. Ceresoli and M. Pavanello, *J. Chem. Phys.*, 2015, **142**, 154116.  
152 M. De Santis, L. Belpassi, C. R. Jacob, A. Severo Pereira Gomes, F. Tarantelli, L. Visscher and L. Storchi, *J. Chem. Theory Comput.*, 2020, **16**, 5695–5711.  
153 M. De Santis, V. Vallet and A. S. P. Gomes, *Front. Chem.*, 2022, **10**, 1–14.  
154 G. Donati, A. Wildman, S. Caprasecca, D. B. Lingerfelt, F. Lipparini, B. Mennucci and X. Li, *J. Phys. Chem. Lett.*, 2017, **8**, 5283–5289.  
155 J. Creutzberg and E. D. Hedegård, *J. Chem. Theory Comput.*, 2022, **18**, 3671–3686.  
156 L. Cupellini, S. Caprasecca, C. A. Guido, F. Müh, T. Renger and B. Mennucci, *J. Phys. Chem. Lett.*, 2018, **9**, 6892–6899.  
157 M. Nottoli, S. Jurinovich, L. Cupellini, A. T. Gardiner, R. Cogdell and B. Mennucci, *Photosynth. Res.*, 2018, **137**, 215–226.  
158 F. C. Ramos, M. Nottoli, L. Cupellini and B. Mennucci, *Chem. Sci.*, 2019, **10**, 9650–9662.  
159 X. Jin, W. J. Glover, X. He and X. He, *J. Chem. Theory Comput.*, 2020, **16**, 5174–5188.  
160 L. Scholz and J. Neugebauer, *J. Chem. Theory Comput.*, 2021, **17**, 6105–6121.  
161 H. Schröder and T. Schwabe, *J. Chem. Theory Comput.*, 2018, **14**, 833–842.  
162 J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski, *J. Chem. Phys.*, 2005, **122**, 94115.  
163 L. V. Slipchenko, *J. Phys. Chem. A*, 2010, **114**, 8824–8830.  
164 A. Defusco, N. Minezawa, L. V. Slipchenko, F. Zahariev and M. S. Gordon, *J. Phys. Chem. Lett.*, 2011, **2**, 2184–2192.  
165 K. Brejc, T. K. Sixma, P. A. Kitts, S. R. Kain, R. Y. Tsien, M. Ormö and S. J. Remington, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 2306–2311.  
166 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.  
167 J. Stone, An Efficient Library for Parallel Ray Tracing and Animation, Ph.D. thesis, Computer Science Department, University of Missouri-Rolla, 1995.  
168 O. Shimomura, F. H. Johnson and Y. Saiga, *Extraction, purification and properties of aequorin, a bioluminescent*, 1962, Tech. Rep.

- 
- 169 F. H. Johnson, O. Shimomura, Y. Saiga, L. C. Gershman, G. T. Reynolds and J. R. Waters, *J. Cell. Comp. Physiol.*, 1962, **60**, 85–103.
- 170 H. Morise, O. Shimomura, F. H. Johnson and J. Winant, *Biochemistry*, 1974, **13**, 2656–2662.
- 171 F. G. Prendergast and K. G. Mann, *Biochemistry*, 1978, **17**, 3448–3453.
- 172 F. Yang, L. G. Moss and G. N. Phillips, *Nat. Biotechnol.*, 1996, **14**, 1246–1251.
- 173 M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien and S. J. Remington, *Science*, 1996, **273**, 1392–1395.
- 174 R. N. Day and M. W. Davidson, *Chem. Soc. Rev.*, 2009, **38**, 2887–2921.
- 175 R. H. Newman, M. D. Fosbrink and J. Zhang, *Chem. Rev.*, 2011, **111**, 3614–3666.
- 176 H. Niwa, S. Inouye, T. Hirano, T. Matsuno, S. Kojima, M. Kubota, M. Ohashi and F. I. Tsuji, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13617–13622.
- 177 S. Kojima, H. Ohkawa, T. Hirano, S. Maki, H. Niwa, M. Ohashi, S. Inouye and F. I. Tsuji, *Tetrahedron Lett.*, 1998, **39**, 5239–5242.
- 178 A. Svendsen, H. V. Kiefer, H. B. Pedersen, A. V. Bochenkova and L. H. Andersen, *J. Am. Chem. Soc.*, 2017, **139**, 8766–8771.
- 179 C. W. West, A. S. Hudson, S. L. Cobb and J. R. Verlet, *J. Chem. Phys.*, 2013, **139**, 071104.
- 180 M. E. Martin, F. Negri and M. Olivucci, *J. Am. Chem. Soc.*, 2004, **126**, 5452–5464.
- 181 A. V. Bochenkova and L. H. Andersen, *Faraday Discuss.*, 2013, **163**, 297–319.
- 182 R. Y. Tsien, *The green fluorescent protein*, 1998, Tech. Rep., pp. 509–544.
- 183 S. R. Meech, *Chem. Soc. Rev.*, 2009, **38**, 2922–2934.
- 184 W. W. Ward, C. W. Cody, R. C. Hart and M. J. Cormier, *Photochem. Photobiol.*, 1980, **31**, 611–615.
- 185 R. Heim, D. C. Prasher and R. Y. Tsien, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 12501–12504.
- 186 A. B. Cubitt, R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross and R. Y. Tsien, *Trends Biochem. Sci.*, 1995, **20**, 448–455.
- 187 H. Lossau, A. Kummer, R. Heinecke, F. Pöllinger-Dammer, C. Kompa, G. Bieser, T. Jonsson, C. M. Silva, M. M. Yang, D. C. Youvan and M. E. Michel-Beyerle, *Chem. Phys.*, 1996, **213**, 1–16.
- 188 M. Chattoraj, B. A. King, G. U. Bublitz and S. G. Boxer, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 8362–8367.
- 189 T. M. H. Creemers, A. J. Lock, V. Subramaniam, T. M. Jovin and S. Völker, *Nat. Struct. Biol.*, 1999, **6**, 706.
- 190 T. M. Creemers, A. J. Lock, V. Subramaniam, T. M. Jovin and S. Völker, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 2974–2978.
- 191 G. Jung, J. Wiehler and A. Zumbusch, *Biophys. J.*, 2005, **88**, 1932–1947.
- 192 C. Filippi, M. Zaccheddu and F. Buda, *J. Chem. Theory Comput.*, 2009, **5**, 2074–2087.
- 193 E. Epifanovsky, I. Polyakov, B. Grigorenko, A. Nemukhin and A. I. Krylov, *J. Chem. Theory Comput.*, 2009, **5**, 1895–1906.
- 194 J. Y. Hasegawa, T. Ise, K. J. Fujimoto, A. Kikuchi, E. Fukumura, A. Miyawaki and Y. Shiro, *J. Phys. Chem. B*, 2010, **114**, 2971–2979.
- 195 K. B. Bravaya, M. G. Khrenova, B. L. Grigorenko, A. V. Nemukhin and A. I. Krylov, *J. Phys. Chem. B*, 2011, **115**, 8296–8303.
- 196 M. Wanko, P. García-Risueño and A. Rubio, *Phys. Status Solidi B*, 2012, **249**, 392–400.
- 197 A. H. Steindal, J. M. H. Olsen, K. Ruud, L. Frediani and J. Kongsted, *Phys. Chem. Chem. Phys.*, 2012, **14**, 5440–5451.
-

- 
- 198 K. B. Bravaya, B. L. Grigorenko, A. V. Nemukhin and A. I. Krylov, *Acc. Chem. Res.*, 2012, **45**, 265–275.
- 199 V. R. Kaila, R. Send and D. Sundholm, *Phys. Chem. Chem. Phys.*, 2013, **15**, 4491–4495.
- 200 A. Petrone, P. Caruso, S. Tenuta and N. Rega, *Phys. Chem. Chem. Phys.*, 2013, **15**, 20536–20544.
- 201 B. L. Grigorenko, A. V. Nemukhin, I. V. Polyakov, D. I. Morozov and A. I. Krylov, *J. Am. Chem. Soc.*, 2013, **135**, 11541–11549.
- 202 M. T. Beerepoot, A. H. Steindal, J. Kongsted, B. O. Brandsdal, L. Frediani, K. Ruud and J. M. H. Olsen, *Phys. Chem. Chem. Phys.*, 2013, **15**, 4735–4743.
- 203 P. Amat and R. Nifosi, *J. Chem. Theory Comput.*, 2013, **9**, 497–508.
- 204 M. T. Beerepoot, A. H. Steindal, K. Ruud, J. M. H. Olsen and J. Kongsted, *Comput. Theor. Chem.*, 2014, **1040–1041**, 304–311.
- 205 A. Pikulska, A. H. Steindal, M. T. Beerepoot and M. Pecul, *J. Phys. Chem. B*, 2015, **119**, 3377–3386.
- 206 J. W. Park and Y. M. Rhee, *Phys. Chem. Chem. Phys.*, 2016, **18**, 3944–3955.
- 207 L. J. Nãbo, J. M. H. Olsen, T. J. Martínez and J. Kongsted, *J. Chem. Theory Comput.*, 2017, **13**, 6230–6236.
- 208 R. Nifosi, B. Mennucci and C. Filippi, *Phys. Chem. Chem. Phys.*, 2019, **21**, 18988–18998.
- 209 P. Reinholdt, E. R. Kjellgren, C. Steinmann and J. M. H. Olsen, *J. Chem. Theory Comput.*, 2020, **16**, 1162–1174.
- 210 D. Grabarek and T. Andruniów, *J. Chem. Theory Comput.*, 2020, **16**, 6439–6455.
- 211 D. Grabarek and T. Andruniów, *J. Chem. Phys.*, 2020, **153**, 215102.
- 212 D. Grabarek and T. Andruniów, *ChemPhysChem*, 2022, **23**, e202200003.
- 213 M. A. Marques, X. López, D. Varsano, A. Castro and A. Rubio, *Phys. Rev. Lett.*, 2003, **90**, 4.
- 214 A. Sinicropi, T. Andruniów, N. Ferré, R. Basosi and M. Olivucci, *J. Am. Chem. Soc.*, 2005, **127**, 11534–11535.
- 215 A. V. Nemukhin, I. A. Topol and S. K. Burt, *J. Chem. Theory Comput.*, 2006, **2**, 292–299.
- 216 J. Y. Hasegawa, K. Fujimoto, B. Swerts, T. Miyahara and H. Nakatsuji, *J. Comput. Chem.*, 2007, **28**, 2443–2452.
- 217 B. L. Grigorenko, A. V. Nemukhin, D. I. Morozov, I. V. Polyakov, K. B. Bravaya and A. I. Krylov, *J. Chem. Theory Comput.*, 2012, **8**, 1912–1920.
- 218 L. Ding, L. W. Chung and K. Morokuma, *J. Phys. Chem. B*, 2013, **117**, 1075–1084.
- 219 F. J. Avila Ferrer, M. D. Davari, D. Morozov, G. Groenhof and F. Santoro, *ChemPhysChem*, 2014, **15**, 3246–3257.
- 220 E. Brunk and U. Rothlisberger, *Chem. Rev.*, 2015, **115**, 6217–6263.
- 221 G. Díaz Mirón and M. C. González Lebrero, *J. Phys. Chem. A*, 2020, **124**, 9503–9512.
- 222 M. Nottoli, B. Mennucci and F. Lipparini, *Phys. Chem. Chem. Phys.*, 2020, **22**, 19532–19541.
- 223 T. J. Zuehlsdorff, S. V. Shedge, S. Y. Lu, H. Hong, V. P. Aguirre, L. Shi and C. M. Isborn, *Annu. Rev. Phys. Chem.*, 2020, **72**, 165–188.

### A.2.4.2. Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. J. *Phys. Chem. A* **2023**, *127*, 5689–5703

Reprinted with permission from Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. J. *Phys. Chem. A* **2023**, *127*, 5689–5703. Copyright 2023 American Chemical Society.



pubs.acs.org/JPCA

Article

## Theoretical and Numerical Comparison of Quantum- and Classical Embedding Models for Optical Spectra

Published as part of *The Journal of Physical Chemistry A* virtual special issue "Roland Lindh Festschrift".

Marina Jansen, Peter Reinholdt, Erik D. Hedegård,\* and Carolin König\*

Cite This: *J. Phys. Chem. A* **2023**, *127*, 5689–5703

Read Online

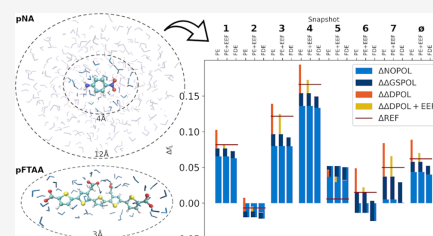
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Quantum-mechanical (QM) and classical embedding models approximate a supermolecular quantum-chemical calculation. This is particularly useful when the supermolecular calculation has a size that is out of reach for present QM models. Although QM and classical embedding methods share the same goal, they approach this goal from different starting points. In this study, we compare the polarizable embedding (PE) and frozen-density embedding (FDE) models. The former is a classical embedding model, whereas the latter is a density-based QM embedding model. Our comparison focuses on solvent effects on optical spectra of solutes. This is a typical scenario where super-system calculations including the solvent environment become prohibitively large. We formulate a common theoretical framework for PE and FDE models and systematically investigate how PE and FDE approximate solvent effects. Generally, differences are found to be small, except in cases where electron spill-out becomes problematic in the classical frameworks. In these cases, however, atomic pseudopotentials can reduce the electron-spill-out issue.



### INTRODUCTION

Quantum-mechanical (QM) methods are indispensable for the calculation of optical spectra, but their use often becomes computationally too demanding for large systems. Embedding schemes have been introduced to circumvent the full, super-system QM calculation by including large environments through an effective embedding operator.

The definition of an embedding model requires that the system is split into an active system and the remaining part ("the environment").<sup>1–4</sup> Embedding models can be divided into two main classes: (i) QM–classical embedding approaches describe the active system by a QM method, whereas all interactions between the active system and environment (as well as the environment itself) are treated by a classical description. (ii) QM–QM embedding describes both active system and environment with QM methods (either on the same or different footings). In this case, the interaction between active system and the environment also contains QM contributions.

For optical properties, the electrostatic interaction between the active system and the environment is often the dominating embedding contribution. In traditional QM–classical approaches, this contribution is modeled through (atomic) point charges in the environment.<sup>4</sup> The point-charge model is, however, insufficient in many cases.<sup>1,5–9</sup> Therefore, a large

number of more advanced embedding schemes have been developed over the years.<sup>3,7,10–23</sup>

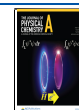
In this work, we employ an advanced QM–classical embedding model, namely the polarizable embedding (PE) model.<sup>20</sup> In this model, point charges in the environment are replaced by a multipole expansion. Additionally, PE incorporates the environment polarization through anisotropic electronic dipole–dipole polarizabilities. The parameters for the environment (*i.e.*, multipoles and polarizabilities) are obtained from QM calculations on isolated fragments. If the environment is a solvent, these fragments are most naturally defined as solvent molecules.

In the class of QM–QM embedding methods, the total system is expressed by means of fragments or subsystems:<sup>24–27</sup> Within density functional theory (DFT) this is known as subsystem DFT.<sup>28</sup> All subsystems are described by their electron densities, which are obtained by quantum-chemical calculations. The interaction of the environment subsystems

Received: April 17, 2023

Revised: June 10, 2023

Published: July 3, 2023



ACS Publications

© 2023 American Chemical Society

5689

<https://doi.org/10.1021/acs.jpca.3c02540>  
*J. Phys. Chem. A* **2023**, *127*, 5689–5703

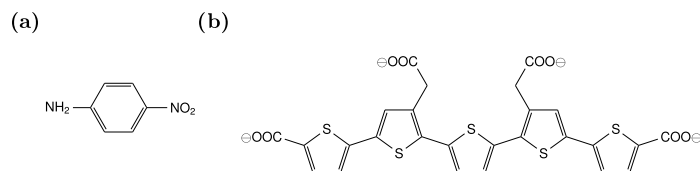


Figure 1. Dyes considered in this work: (a) *para*-nitroaniline (b) pentameric formyl thiophene acetic acid.

with the active subsystem is then recovered through an embedding potential, which contains quantum-mechanical contributions. This potential is dependent on all other subsystem's electron densities.<sup>28,29</sup> In practice, the environmental electron densities are commonly kept frozen, so that only the active subsystem's electron density is polarized. In this frozen-density embedding (FDE) approach,<sup>29</sup> the effect of the environment's polarizability can be incorporated by self-consistently cycling through all subsystems in a so-called freeze-and-thaw scheme.<sup>25,27–30</sup>

Both embedding classes share, hence, a common goal and have been employed to show that polarization effects originating from the environment can play a significant role in the accurate calculation of local optical properties.<sup>8,31–34</sup> Yet, direct numerical comparisons have been rare.<sup>35,36</sup> We recently developed a common theoretical framework<sup>17</sup> encompassing both fragmentation-based QM–QM and QM–classical embedding methods with a special focus on FDE and PE. This framework was employed to dissect how the two classes of embedding models describe the interactions between the active system and the environment. We here continue this comparison by quantifying how the theoretical differences manifest numerically for optical properties of two solvated systems, employing a supermolecular calculation as a reference.

Our target systems are two fluorescent dyes whose excited-state properties are known to be sensitive to solvent effects: The first target system is *para*-nitroaniline (*p*NA, see Figure 1a), which has been studied with several different embedding schemes.<sup>7,8,37–40</sup> Yet, the performance of the QM–QM and QM–classical embedding schemes for *p*NA have never been compared in a combined study. The second test case is pentameric formyl thiophene acetic acid (pFTAA, see Figure 1b), a luminescent biomarker developed for fluorescence imaging for amyloid proteins.<sup>41</sup> The mechanism occurring with the chromophore embedded in the protein is not fully understood yet, but it is known that its properties strongly depend on the solvent–solute interactions and the conformation of the molecule.<sup>42–45</sup> Notably, pFTAA is an anionic system and therefore poses somewhat different challenges in the description of the solute–solvent interaction than *p*NA.

This paper is organized as follows. We first briefly introduce the PE and FDE scheme in the common theoretical framework as derived in previous work<sup>17</sup> (Theoretical Background Section). In particular, we point out similarities and differences between these schemes. We then describe the Computational Setup, developed to enable us to compare the embedding methods on equal footing before we outline the Computational Details. Subsequently, the results are presented and discussed (Results and Discussion Section). Finally, the last section provides the Summary and Conclusions.

## THEORETICAL BACKGROUND

In this section, we give a brief overview of the density-based QM–QM embedding and classical PE methods. For a more detailed derivation, reviews, and further extensions of the presented models we refer to refs 14–17, 28, 36, 46–48. Here, we use the common framework developed in ref 17.

Embedding schemes commonly focus on a selected, active subsystem A. The remaining parts are labeled the environment (env). The total energy can be written as the sum of the energy functionals of the respective electron densities  $\rho_A$  and  $\rho_{env}$

$$E_{tot}[\rho_{tot}] = E_A[\rho_A] + E_{env}[\rho_{env}] + E_{int}[\rho_A, \rho_{env}] \quad (1)$$

where  $E_{tot}[\rho_{tot}]$  and  $E_A[\rho_A]$  denote the total energy functional and the energy functional of the active subsystem A, respectively. The  $E_{int}[\rho_A, \rho_{env}]$  term defines the interaction energy, i.e.,  $E_{int}[\rho_A, \rho_{env}] = E_{tot}[\rho_{tot}] - E_A[\rho_A] - E_{env}[\rho_{env}]$ , where we assume  $\rho_{tot} = \rho_A + \rho_{env}$ . The environment density can analogously be approximated as a sum of all environmental (subsystem) densities,  $\rho_X$  by  $\rho_{env} = \sum_{X \neq A} \rho_X$ . We further decompose the interaction energy into Coulomb (classical),  $E_{int}^C[\rho_A, \rho_{env}]$ , and QM contributions,  $E_{int}^{QM}[\rho_A, \rho_{env}]$ ,

$$E_{int}[\rho_A, \rho_{env}] = E_{int}^C[\rho_A, \rho_{env}] + E_{int}^{QM}[\rho_A, \rho_{env}] \quad (2)$$

where both contributions can be obtained analogously to eq 1.

The Coulomb part is included in both density-based and classical embedding schemes. In the density-based QM–QM embedding schemes this part is expressed as

$$\begin{aligned} E_{int}^C[\rho_A, \rho_{env}] = & \sum_{X \neq A} \int \int \frac{\rho_A(\mathbf{r}_a) \rho_X(\mathbf{r}_x)}{|\mathbf{r}_a - \mathbf{r}_x|} d\mathbf{r}_a d\mathbf{r}_x \\ & - \sum_{X \neq A} \sum_{I \in A} \int \frac{Z_I \rho_X(\mathbf{r}_x)}{|\mathbf{R}_I - \mathbf{r}_x|} d\mathbf{r}_x \\ & - \sum_{X \neq A} \sum_{J \in X} \int \frac{Z_J \rho_A(\mathbf{r}_a)}{|\mathbf{R}_J - \mathbf{r}_a|} d\mathbf{r}_a \\ & + \sum_{X \neq A} \sum_{I \in A} \sum_{J \in X} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \end{aligned} \quad (3)$$

where  $\mathbf{R}_{I/J}$  denote nuclear coordinates,  $\mathbf{r}_{a/x}$  denote electronic coordinates, and  $Z_{I/J}$  are nuclear charges. The interaction of the environment with the active subsystem is included via Coulomb [and possibly quantum-mechanical (QM)] contributions through an effective embedding operator. The effective Hamiltonian for subsystem A can be expressed as

$$\hat{H}_A^{\text{eff}} = \hat{H}_A + \hat{v}_A^{\text{emb}} = \hat{H}_A + \int \hat{\rho}_A(\mathbf{r}_a) v_A^{\text{emb}}(\mathbf{r}_a) d\mathbf{r}_a \quad (4)$$

where  $\hat{H}_A$  denotes the Hamiltonian of the isolated subsystem A incorporating the kinetic energy and intra-subsystem Coulomb interaction terms. The density operator  $\hat{\rho}_A(\mathbf{r}_a) = \sum_{i \in A} \delta(\mathbf{r}_i -$

$\mathbf{r}_a$ ) defines the connection between the Hamiltonian with electron-based coordinates and density-based expressions with real-space coordinates. We define the embedding operator via a real-space potential

$$v_A^{\text{emb}}(\mathbf{r}_a) = \frac{\delta}{\delta \rho_A} (E_{\text{tot}}[\rho_{\text{tot}}] - E_A[\rho_A]) \\ = \frac{\delta E_{\text{tot}}^{\text{C}}[\rho_{\text{tot}}]}{\delta \rho_A} - \frac{\delta E_A^{\text{C}}[\rho_A]}{\delta \rho_A} + \frac{\delta E_{\text{tot}}^{\text{QM}}[\rho_{\text{tot}}]}{\delta \rho_A} - \frac{\delta E_A^{\text{QM}}[\rho_A]}{\delta \rho_A} \quad (5)$$

In FDE, the environmental density  $\rho_{\text{env}}$  is kept frozen so that the total embedding potential resulting from eq 5 becomes

$$v_A^{\text{FDE}}(\mathbf{r}_a) = v_A^{\text{C(FDE)}}[\rho_{\text{env}}](\mathbf{r}_a) + v_A^{\text{nadd,kin}}[\rho_A, \rho_{\text{env}}](\mathbf{r}_a) \\ + v_A^{\text{nadd,xc}}[\rho_A, \rho_{\text{env}}](\mathbf{r}_a) \quad (6)$$

with the Coulomb potential only depending on  $\rho_A$  and the (frozen) densities of the environment

$$v_A^{\text{C(FDE)}}[\rho_{\text{env}}](\mathbf{r}_a) = \frac{\delta E_{\text{int}}^{\text{C}}[\rho_A, \rho_{\text{env}}]}{\delta \rho_A} \\ = - \sum_{X \neq A} \sum_{J \in X} \frac{Z_J}{|\mathbf{r}_a - \mathbf{R}_J|} \\ + \sum_{X \neq A} \int \frac{\rho_X(\mathbf{r}_x)}{|\mathbf{r}_a - \mathbf{r}_x|} d\mathbf{r}_x \quad (7)$$

where  $E_{\text{int}}^{\text{C}}[\rho_A, \rho_{\text{env}}]$  is defined in eq 3. The QM contributions from eq 5 are comprised of kinetic and an exchange–correlation (xc) parts, represented by  $v_A^{\text{nadd,kin}}$  and  $v_A^{\text{nadd,xc}}$  in eq 6. In practical calculations, these contributions are often approximated by orbital-free DFT methodologies,<sup>28,29,47</sup> though for  $v_A^{\text{nadd,kin}}$  also orbital-dependent projection schemes have been reported.<sup>49–61</sup>

The application of a fixed  $\rho_{\text{env}}$  in FDE leads to several possible choices of frozen densities. The crudest approximation is to use the density from the isolated fragments. We denote these densities  $\{\rho_X^{(0)}\}$ . Likewise, we also can define  $\rho_A^{(0)}$ . Allowing the active subsystem A to relax by submitting  $\rho_A$  to a self-consistent-field optimization in the frozen environment density,  $\rho_{\text{env}}^{(0)}$ , leads to a relaxed density  $\rho_A^{(1)}$ . In terms of density-based embedding schemes, this approach directly refers to FDE.<sup>29</sup> It yields a relaxed energy for the active subsystem  $E_A[\rho_A^{(1)}]$ .

The relaxation of  $\rho_A^{(0)}$  to  $\rho_A^{(1)}$  can be done for all fragments in a step-wise manner until self-consistency to obtain the relaxed densities  $\rho_X^{(2)}$  and  $\rho_{\text{env}}^{(2)}$ . This is denoted a freeze-and-thaw procedure.<sup>62</sup> Formally, the mutual polarization of the densities in the ground state of the super system is recovered when performing a sufficient number of freeze-and-thaw cycles. This mutual optimization does not only account for electronic polarization but also affects non-electrostatic contributions through the density-dependent description of the kinetic part of the interaction energy.<sup>46</sup>

In contrast to that, the PE methods approximate both static electrostatics and polarization solely based on frozen densities in the environment,  $\rho_{\text{env}}^{(0)} = \sum_{X \neq A} \rho_X^{(0)}$ . An expression for the

total energy comparable to eq 1 can then be obtained through Rayleigh–Schrödinger perturbation theory,

$$E_{\text{tot}} \approx E^{(0)} + E^{(1)} + E^{(2)} \quad (8)$$

By expressing the interaction between the subsystems as perturbations of the energy, the zeroth-order perturbation can be identified as the isolated subsystem energies. Thus, from eq 1 we identify  $E^{(0)} = E_A[\rho_A^{(0)}] + E_{\text{env}}[\rho_{\text{env}}^{(0)}]$  and the interaction energy ( $E_{\text{int}}[\rho_A^{(0)}, \rho_{\text{env}}^{(0)}]$ ) must therefore come through the higher-order energy corrections. Indeed, the first-order correction corresponds to eq 3 with frozen densities, i.e.,  $E_{\text{int}}^{\text{C}}[\rho_A^{(0)}, \rho_{\text{env}}^{(0)}]$ . The PE model further approximates  $E_{\text{int}}^{\text{C}}[\rho_A^{(0)}, \rho_{\text{env}}^{(0)}]$  through a multipole expansion,<sup>34,63,64</sup> i.e.,

$$E^{(1)} = E_{\text{int}}^{\text{C}}[\rho_A^{(0)}, \rho_{\text{env}}^{(0)}] \approx E_{\text{int}}^{\text{mult}}[\rho_A^{(0)}, \rho_{\text{env}}^{(0)}] \quad (9)$$

the multipole expansion employs individual atoms of the subsystems/fragments as expansion points ( $\{\mathbf{R}_s\}$  or in short sites,  $s$ ). The multipole expansion can thus be written as

$$E_{\text{int}}^{\text{mult}}[\rho_A^{(0)}, \{\rho_X^{(0)}\}] \\ = \sum_{X \neq A} \sum_{s \in X} \left( - \int \rho_A^{(0)}(\mathbf{r}_a) T_{sa}^{(0)} d\mathbf{r}_a + \sum_{I \in A} Z_I T_{sI}^{(0)} \right) q_I[\rho_X^{(0)}] \\ - \sum_{X \neq A} \sum_{s \in X} \left( - \int \rho_A^{(0)}(\mathbf{r}_a) T_{sa}^{(1)} d\mathbf{r}_a + \sum_{I \in A} Z_I T_{sI}^{(1)} \right) \mu_I^s[\rho_X^{(0)}] \\ + \dots \quad (10)$$

In the above equation, we have defined the interaction operators to  $k$ 'th order in a multi-index notation<sup>17</sup>

$\mathbf{T}_{sa}^{(k)} = \frac{\partial^{k_x+k_y+k_z}}{\partial x_s^{k_x} \partial y_s^{k_y} \partial z_s^{k_z}} |\mathbf{R}_a - \mathbf{R}_s|^{-1}$ . Moreover, the multipole moment operator of  $k$ 'th order on site  $s$  is defined as  $\mathbf{Q}^{(k)}[\rho_X^{(0)}] = (\Psi_X^{(0)} | \hat{\mathbf{Q}}^{(k)} | \Psi_X^{(0)})$ . The term for zeroth-order moments represents the charge contribution  $\hat{\mathbf{Q}}_s^{(0)} = \hat{q}_s$ , the first-order term denotes the dipole contribution  $\hat{\mathbf{Q}}_s^{(1)} = \hat{\boldsymbol{\mu}}_s$ , and so on. In the following, we combine the two sums over fragments  $X$  and sites  $s$  into one sum over all sites  $s$ .

The mutual polarization effects are approximately covered by the second-order correction<sup>64</sup>

$$E^{(2)} = E_A^{\text{pol}} + E_{\text{env}}^{\text{pol}} + E^{\text{disp}} \quad (11)$$

We focus on the following only on the polarization part, while neglecting the dispersion contribution,  $E^{\text{disp}}$ . The environment polarization energy,  $E_{\text{env}}^{\text{pol}}$ , can be described as

$$E_{\text{env}}^{\text{pol}}[\rho_A^{(0)}] = - \frac{1}{2} \mathcal{E}^T[\rho_A^{(0)}] \cdot \boldsymbol{\mu}^{\text{ind}}[\rho_A^{(0)}] \quad (12)$$

and  $E_A^{\text{pol}}$  can in principle be obtained analogously. This part is, however, inherently included in the QM model for the active system. The field  $\mathcal{E}$  is defined as the sum of the fields from electrons in system A, nuclei in system A, and the multipoles in the environment

$$\mathcal{E}[\rho_A^{(0)}] = \mathcal{E}_A^{\text{e}}[\rho_A^{(0)}] + \mathcal{E}_A^{\text{n}} + \mathcal{E}_{\text{env}}^{\text{mult}} \quad (13)$$

The induced dipole moment on site  $s$  can then be obtained as<sup>65</sup>

$$\boldsymbol{\mu}_s^{\text{ind}}[\rho_A^{(0)}] = \boldsymbol{\alpha}_s \left( \mathcal{E}_s[\rho_A^{(0)}] + \sum_{s' \neq s} \mathbf{T}_{ss'}^{(2)} \boldsymbol{\mu}_{s'}^{\text{ind}} \right) \quad (14)$$

where  $\alpha_s$  is the (static) point-polarizability localized on site  $s$  and  $\mathcal{E}_s[\rho_A^{(0)}]$  the field in eq 13 on site  $s$ . Note that the induced dipole on site  $s$  depends on the field generated from the induced dipoles on all remaining sites. Thus, a self-consistent optimization is required to obtain the induced dipole moment,  $\mu_s^{\text{ind}}$ . This optimization problem can be written as

$$\mu_s^{\text{ind}}[\rho_A^{(0)}] = \sum_t \mathbf{R}_{st} \mathcal{E}_{A,s}[\rho_A^{(0)}] \quad (15)$$

where the so-called classical response matrix,  $\mathbf{R}$ , is given as

$$\mathbf{R} = \begin{pmatrix} \alpha_1^{-1} & -\mathbf{T}_{12}^{(2)} & \dots & -\mathbf{T}_{1S}^{(2)} \\ -\mathbf{T}_{21}^{(2)} & \alpha_2^{-1} & \dots & -\mathbf{T}_{2S}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{T}_{S1}^{(2)} & -\mathbf{T}_{S2}^{(2)} & \dots & \alpha_S^{-1} \end{pmatrix}^{-1} \quad (16)$$

The total energy (in eq 1) is now defined by combining eqs 8–11,

$$E_{\text{tot}}^{\text{PE}}[\rho_A, \rho_{\text{env}}^{(0)}] = E_A^{\text{PE}}[\rho_A, \rho_{\text{env}}^{(0)}] + E_{\text{env}}^{\text{PE}}[\rho_{\text{env}}^{(0)}, \rho_A] + E_{\text{int}}^{\text{mult}}[\rho_A, \rho_{\text{env}}^{(0)}] \quad (17)$$

here we skip the superscript for  $\rho_A$  to denote that it is subject to change in the self-consistent-field (SCF) procedure performed during the optimization of the QM system (note that eq 15 will then have to be solved within each SCF cycle). The environment density  $\rho_{\text{env}}^{(0)}$  remains the isolated density of the environment fragments (represented by a multipole expansion). For consistency with the definition in eq 1, we have written the total energy in eq 17 as a sum of the energies of the active system A and the environment plus an interaction energy, where we have combined the energy of the isolated subsystem and polarization in the term  $E_A^{\text{PE}}[\rho_A, \rho_{\text{env}}^{(0)}] = E_A[\rho_A] + E_{\text{env}}^{\text{pol}}[\rho_{\text{env}}^{(0)}]$ . The term  $E_{\text{env}}^{\text{PE}}[\rho_{\text{env}}^{(0)}, \rho_A]$  is defined in a similar fashion.

With this starting point, a PE embedding potential according to eq 5 is derived to be<sup>17</sup>

$$v_A^{\text{PE}} = \frac{\delta E_{\text{tot}}^{\text{PE}}[\rho_A, \rho_{\text{env}}^{(0)}]}{\delta \rho_A} - \frac{\delta E_A^{\text{PE}}[\rho_A]}{\delta \rho_A} = v^{\text{mult}} + v^{\text{pol}} \quad (18)$$

where the two operators are defined as

$$v^{\text{mult}} = \sum_s \sum_{k=0} \frac{(-1)^{|k|}}{k!} \mathbf{T}_{sa}^{(k)}(\mathbf{r}_s) \mathbf{Q}_s^{(k)} \quad (19)$$

and

$$v^{\text{pol}} = \sum_s (\mu_s^{\text{ind}}[\rho_A])^T \mathbf{e}_{A,s}^e(\mathbf{r}_s) \quad (20)$$

The field potential  $\mathbf{e}_{A,s}^e(\mathbf{r}_s)$  is the component of the electronic part of the electric field operator on the site  $s$ , defined in real-space coordinates as,

$$\mathbf{e}_{A,s}^e(\mathbf{r}_s) = \frac{\delta \mathcal{E}_{A,s}^e[\rho_A(\mathbf{r}_s)]}{\delta \rho_A(\mathbf{r}_s)} \quad (21)$$

where  $\mathcal{E}_{A,s}^e[\rho_A(\mathbf{r}_s)]$  is the electronic component of the field in system A on site  $s$  (cf. eq 13). Thus,  $v^{\text{mult}}$  corresponds to a multipole approximation of eq 7, where only  $\rho_X^{(0)}$  are employed. Similarly,  $v^{\text{pol}}$  approximates the effect of mutual

polarization, *i.e.*, moving from  $\rho_X^{(0)}$  to  $\rho_X^{(2)}$  in eq 7. The QM contributions from eq 6 are not included in standard PE.

Optical spectra are in this work obtained by linear response theory in the form of time-dependent DFT (TD-DFT). To incorporate embedding contributions for the models discussed above, we add the embedding potential  $v_{\text{emb}}$  to the Kohn-Sham operator of the vacuum system ( $\hat{f}_{\text{iso}}$ )

$$\hat{f}_{\text{tot}} = \hat{f}_{\text{iso}} + v_{\text{emb}} \quad (22)$$

with  $v_{\text{emb}}$  being eqs 6 or 18 for FDE and PE, respectively. Replacing  $\hat{f}_{\text{iso}}$  with  $\hat{f}_{\text{tot}}$  in the derivation of the response equations leads to a set of modified response equations, that are,

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^* & -\mathbf{A}^* \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \mathbf{0} \quad (23)$$

with the excitation energies  $\omega$  and the eigenvectors  $(\mathbf{X}, \mathbf{Y})^T$ . The elements of  $\mathbf{A}$  and  $\mathbf{B}$  are given as

$$A_{ai,bj} = \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i) + B_{ai,jb} \quad (24)$$

$$B_{ai,bj} = \frac{\partial F_{ai}^{\text{iso}}}{\partial P_{bj}} + \frac{\partial \langle \phi_a | v^{\text{emb}} | \phi_j \rangle}{\partial P_{bj}} \quad (25)$$

Here,  $F_{ai}^{\text{iso}}$  denotes a Fock matrix element of the isolated system and  $P_{bj}$  is an element of the density matrix. The quantities  $\epsilon_a$  and  $\epsilon_i$  are orbital energies, where occupied orbitals are labeled with  $i$  or  $j$  and the virtual ones with  $a$  or  $b$ . The orbital energies are eigenvalues of the *total* Fock operator  $\hat{f}_{\text{tot}}$  in eq 22. Thus, part of the environmental contribution (the static electrostatics and ground-state polarization) enters through these energies.

For FDE,  $v^{\text{emb}}$  can be chosen to rely only on  $\{\rho_X^{(0)}\}$  densities, which we denote FDE NOPOL. An equivalent contribution can be defined for PE (which we denote PE NOPOL) if only  $v^{\text{mult}}$  of eq 18 is included in  $v_{\text{emb}}$ , but not  $v^{\text{pol}}$ . Employing the relaxed densities  $\{\rho_X^{(2)}\}$  in eq 5 corresponds to including the ground-state polarization and we denote this model FDE GSPOL. The corresponding model in the PE framework (denoted PE GSPOL) corresponds to employing both  $v^{\text{mult}}$  and  $v^{\text{pol}}$  of eq 18 in  $v_{\text{emb}}$ , while neglecting the  $v_{\text{emb}}$  part of  $B_{ai,bj}$  in eqs 24 and 25.

The second term of  $\mathbf{B}$  requires more attention since the physical content between PE and FDE models is rather different.<sup>17</sup> The term can be identified as

$$\frac{\partial \langle \phi_a | v^{\text{emb}} | \phi_j \rangle}{\partial P_{bj}} = \left\langle \phi_a(\mathbf{r}_a) \phi_b(\mathbf{r}_a) \left| \frac{\delta v_A^{\text{emb}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}_a)} \right| \phi_j(\mathbf{r}_a) \phi_j(\mathbf{r}_a) \right\rangle \quad (26)$$

The functional derivative  $\left( \frac{\delta v_A^{\text{emb}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}_a)} \right)$  for the corresponding embedding scheme can be derived from eqs 6 and 18 for FDE and PE, respectively.

In the case of FDE, we have a static (ground-state) potential and the Coulomb terms vanish. Thus, only quantum-mechanical terms contribute,<sup>66</sup> so that

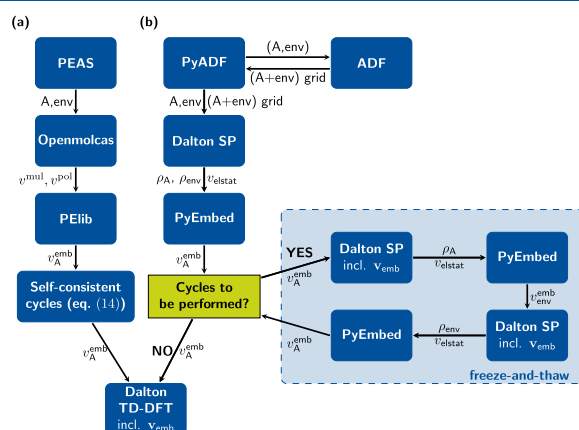
$$\frac{\delta v_A^{\text{FDE}}(\mathbf{r}_a)}{\delta \rho_A(\mathbf{r}_a)} = \frac{\delta^2 E^{\text{QM}}[\rho_{\text{tot}}]}{\delta \rho_{\text{tot}}(\mathbf{r}_a) \delta \rho_{\text{tot}}(\mathbf{r}_a)} - \frac{\delta^2 E^{\text{QM}}[\rho_A]}{\delta \rho_A(\mathbf{r}_a) \delta \rho_A(\mathbf{r}_a)} \quad (27)$$

The remaining QM embedding contributions to the response kernel, however, are often small, so that the major environ-

Table 1. Overview of the Included Contributions in the Different Embedding Models Considered Here

class	label	static	G.s. pol	QM <sup>a</sup>	diff. pol.	EEF	comments
QM/classical	PE NOPOL	⊕	×	×	×	×	$v_A^{\text{mult}}$ (based on $\{\rho_A^{(0)}\}$ ), see eqs 18 and 19.
QM/QM	FDE NOPOL	⊕	×	⊕	×	×	$v_A^{\text{FDE}}$ ( $r_A$ ) in eq 6 based on $\{\rho_A^{(0)}\}$ .
QM/classical	PE GSPOL	⊕	⊕	×	×	×	$v_A^{\text{mult}}$ (based on $\{\rho_A^{(0)}\}$ ) and $v^{\text{pol}}$ , see eqs 18–20.
QM/QM	FDE GSPOL	⊕	⊕	⊕	×	×	$v_A^{\text{FDE}}$ ( $r_A$ ) (eq 6) based on $\{\rho_A^{(2)}\}$ .
QM/classical	PE DPOL	⊕	⊕	×	⊕	×	PE GSPOL and additionally eq 28.
QM/classical	PE DPOL + EEF	⊕	⊕	×	⊕	⊕	PE DPOL with modified dipole transition moments, eq 32

<sup>a</sup>Can be included via orbital-free DFT. Note that we do not consider the response kernel, eq 27, in this work.



**Figure 2.** Flowchart of the performed workflows for molecule (A) in the environment of the molecule (env). The steps in dark blue boxes stand for subprograms used for a task. The large arrows indicate results that are passed in between the different programs. In either workflow, the supermolecular structure is passed to PEAS or PyADF, respectively, and split up into subsystems. (a) PE workflow utilizing PEAS and PElib<sup>78</sup> (PEAS calls Openmolcas for subsystem calculations). (b) The FDE workflow uses the PyADF scripting framework, that is calling all programs and managing all results mentioned in this workflow.

mental effect in the electronic transitions and oscillator strengths are results of differences in the canonical orbitals and orbital energies.<sup>35,67</sup> The QM contributions to the embedding in the response kernel are, hence, not included in the numerical examples of the present work.

For PE, the functional derivative is obtained from eq 18 as<sup>17</sup>

$$\frac{\delta v_A^{\text{PE}}(r_A)}{\delta \rho_A(r_A')} = -\sum_t \sum_s \mathbf{T}_{st}^{(1)}(r_A') \mathbf{R}_t \mathbf{T}_{st}^{(1)}(r_A) \quad (28)$$

This contribution can be understood as an (approximate) treatment of differential polarization, *i.e.*, the difference in the interaction between the ground-state and excited-state densities with the environment densities. We denote PE models with this effect included as PE DPOL. There is no corresponding term for FDE models, although extensions have been suggested that include differential polarization.<sup>8,32,68–72</sup> Most of them are, however, rather computationally demanding<sup>68</sup> or require embedded excited-state densities,<sup>8</sup> which are somewhat tedious to obtain for TD-DFT methods.<sup>69</sup>

While excitation energies are a fundamental part of a UV–vis spectrum, the associated intensities are often highly important for assignments. The intensity is usually obtained from the oscillator strength which can also be extracted from eq 23; for transition  $n$ , the oscillator strength,  $f_n$ , can be calculated as

$$f_n = \frac{2}{3} \omega_n \mu_n^2 \quad (29)$$

where  $\omega_n$  is the excitation frequency and  $\mu_n$  is the transition dipole moment. The latter can (for the  $\alpha$ -component) be obtained from the converged response vectors in eq 23 as<sup>73</sup>

$$\mu_n^\alpha = \frac{1}{\sqrt{\omega_n}} \mathbf{M}^\alpha (\mathbf{X} + \mathbf{Y}) \quad (30)$$

where the  $\mathbf{M}^\alpha$  is a vector comprised of the  $\alpha = x, y$ , and  $z$  components with the elements

$$M_{ii}^\alpha = -\langle \phi_i | r_i | \phi_i \rangle \quad (31)$$

Since the introduction of  $v^{\text{emb}}$  in the response equations (eqs 23–25) also affects the eigenvectors, the embedding also influences the calculated oscillator strengths. However, the external field employed to excite the solute also generates an induced dipole on environment sites. This effect is not included in the standard local embedding schemes: In the FDE scheme, applied here, solely the response of the active system in the frozen environment is determined, while the coupling to an environmental response is neglected. The latter may be incorporated via the non-local, so-called coupled FDE scheme, which affects both excitation energies and oscillator strengths.<sup>68,74–76</sup> This, however, goes beyond the scope of



the present work on local descriptions. Also in the standard PE scheme, the local field effect is not incorporated. It's effect on the oscillator strengths can, however, approximately be accounted for by adding an external effective field (EEF) term,  $\langle \phi_a | \hat{V}^{\text{loc}} | \phi_i \rangle$ , to  $\mathbf{M}$ .<sup>22,77</sup> The required operator is defined as

$$\hat{V}^{\text{loc}} = \sum_{Is} \mathbf{T}_{Ia}^{(1)} \mathbf{R}_{Is} \mathcal{E}_s^{\text{uni}} = \sum_t \mathbf{T}_{Ia}^{(1)} \mu_{\text{ext},t}^{\text{ind}} \quad (32)$$

where  $\mu_{\text{ext},t}^{\text{ind}}$  is the induced dipole to a unit field,  $\mathcal{E}_s^{\text{uni}}$ . This approximate description has thereby an effect on the oscillator strengths only, but does not alter the excitation energies.

We have summarized the contributions considered in the different embedding approaches in Table 1 and refer to the labels used in this table in the following sections.

**Computational Setup.** The common theoretical comparison of density-based QM–QM embedding and PE is only a first step. We also aim for a setup that allows a one-to-one comparison between the two embedding models in practical calculations. Our setup is shown in Figure 2.

The general procedure for the PE model involves the construction of the embedding potential. For this purpose, we employ the PE Assistance Script (PEAS).<sup>79</sup> The script divides the environment into subsystems/fragments and constructs the densities  $\{\rho_X^{(0)}\}$  for the individual fragments, employing DFT calculations with Openmolcas.<sup>80</sup> From these densities localized multipoles,  $\{Q[\rho_X^{(0)}]\}$ , and static polarizabilities,  $\{\alpha_X^s\}$ , can be derived from the LoProp<sup>81</sup> method, implemented in Openmolcas. PEAS collects multipoles and polarizabilities in a potential file that is employed in the calculation of the excitation energies and oscillator strengths with TD-DFT. Optimization of the ground-state as well as solving TD-DFT equations [eqs 23–25] are done while including the PE potential; this means that the induced dipole moments in eq 28 are self-consistently optimized along with the SCF/linear response iterations. PELib handles this self-consistent calculation of the induced dipole moments and adds the resulting PE contributions to the Kohn–Sham or Kohn–Sham-like matrices used in the SCF or response calculations.<sup>78</sup>

We employed four different models of increasing accuracy. (i) PE NOPOL, which only contains multipoles up to quadrupoles and no polarizabilities (eq 18), (ii) PE GSPOL, where full ground-state polarization is included, but differential polarization (eq 28) is ignored in the TD-DFT calculation, (iii) PE DPOL including both multipoles, ground-state, and differential polarization (eq 28), and (iv) PE DPOL with modified dipole transition moments (eq 30) due to the effect of the external effective field in eq 32. The last model (iv) is equivalent to (iii) for excitation energies but leads to a change in oscillator strengths. We denote this model PE DPOL+EEF.

For the FDE calculations, we employed the PyADF scripting framework (Figure 2b).<sup>82,83</sup> A supermolecular DFT integration grid was obtained via the ADF program from AMS2020.103 program suite<sup>84</sup> to ensure that the grids include the full area of all subsystems to preempt grid artifacts that could affect our comparison. Subsequently, ground-state calculations of all subsystems were performed individually in Dalton.<sup>85</sup> The resulting molecular orbitals from these calculations were then translated to electron density and electrostatic potentials on the initially generated grid, using the DensityEvaluator module of PyADF. The non-additive kinetic and the exchange–correlation term (eq 6) were then evaluated by PyEmbed

module of PyADF on the same grid. Finally, the embedding potential was obtained by adding the environmental electrostatic potential and the environmental non-additive potential for the kinetic and exchange–correlation term.

For the FDE calculations, we employed two potentials: (i) A static embedding potential from isolated environmental densities (skipping the performance of FDE cycles, see Figure 2b: blue underlaid box, FDE NOPOL, eq 6). (ii) A potential including mutual polarization via freeze-and-thaw cycles of the active subsystem with environmental fragments in the ground state (FDE GSPOL, eq 6). In the freeze-and-thaw procedure, we employ the DensityEvaluator to write updated density and electrostatic potential of every ground-state subsystem calculation in Dalton on the integration grid. This results in an updated embedding potential when evaluating the embedding potential with PyEmbed.

It should be noted, that the current implementation in Dalton includes the non-additive parts of the embedding potential in the SCF process, but not in the response kernel for the TD-DFT calculation.

The above-described framework enables us to dissect the embedding contributions and quantify the different approximations discussed above. For this, the presented embedding models were to a large degree implemented in Dalton to allow a fair side-by-side comparison and the stepwise inclusion of polarization effects.

## COMPUTATIONAL DETAILS

The solvated pNA and pFTAA molecules are dynamic and different solvent configurations may lead to different results. To ensure that our comparison is not biased towards one specific configuration of solvent molecules, we employed a few different configurations. These configurations (*i.e.*, solute and water atomic coordinates) were extracted as snapshots from a time-dependent simulation of the solvated molecules, employing classical molecular dynamics (MDs). The snapshots of pNA were taken from an MD simulation, using the AMBER software.<sup>86</sup> We parameterized the pNA molecule with the General AMBER force field (GAFF)<sup>87</sup> and RESP charges<sup>88</sup> calculated with B3LYP<sup>89–91</sup> 6-31+G\* basis set<sup>92–94</sup> (with PCM<sup>95</sup> using the dielectric constant of water). The system was set up with tleap of the Amber package and pNA was solvated with 3160 water molecules, represented by the OPC model.<sup>96</sup> We first ran a minimization using 10000 steps of steepest descent, followed by 10,000 steps of conjugate gradient minimization. We next equilibrated the system by running a 1 ns (in the NPT ensemble), heating the system from 0 to 298 K (at 1 atm. pressure) over the first 20 ps. This was followed by a 100 ns production run, using the NPT ensemble (at 298 K), a Langevin thermostat, and a Monte Carlo barostat. Electrostatics were treated with Particle Mesh Ewald,<sup>97</sup> and non-bonded interactions were cut-off at 12 Å. The hydrogen bonds were constrained with the SHAKE algorithm.<sup>98,99</sup> For further calculations, we arbitrarily selected seven out of the total of one hundred obtained snapshots. For these snapshots, we constructed systems where all environment molecules within 3, 4, 5, and 12 Å of pNA were included.

The snapshots of pFTAA were taken from an MD simulation, using the GROMACS software.<sup>100–107</sup> We parameterized the pFTAA molecule with an adapted CHARMM force field.<sup>42,43,108,109</sup> The pFTAA molecule was solvated with 4028 water molecules, represented by the TIP3P model.<sup>110</sup> We first ran a minimization using 50,000 steps of steepest descent.

We next equilibrated the system for 10 ns (in the NPT ensemble), heating the system from 0 to 300 K (at 1 atm. pressure) over the first 0.2 ps. All employed snapshots were taken from a 100 ns production run, using the NVT ensemble (at 300 K), a velocity-rescaling thermostat,<sup>111</sup> a Berendsen barostat<sup>112</sup> and electrostatics were treated with Particle Mesh Ewald,<sup>97,113</sup> and non-bonded interactions were cut-off at 10 Å. All bonds were constrained with the LINCS algorithm.<sup>114</sup> For pFTAA, we only consider solvated models with a 3 Å water environment for a selection of eight independent snapshots. We note that, for some snapshots, there are sodium ions in the 3 Å environment at an average distance of 2.3 Å to the chromophore, while for others there are no sodium ions in close proximity of the dye.

The reference calculations were performed with Dalton 2020<sup>85</sup> in a supermolecular TD-DFT calculation with the CAM-B3LYP<sup>115</sup> xc functional. The workflow for the embedding calculation is shown in Figure 2. The construction of the environment potential in the polarizable embedding approach was performed with LoProp<sup>81</sup> in Openmolcas<sup>80</sup> in combination with the Polarizable Embedding Assistant Script (PEAS).<sup>116</sup> In these fragment calculations, the B3LYP<sup>89–91</sup> xc functional was employed together with ANO-type reconstructions of the aug-cc-pVDZ and aug-cc-pVTZ basis set, respectively.<sup>117–120</sup> For sodium ions, the ANO-L basis sets were applied.<sup>121</sup> The linear response calculation for pNA was then carried out with Dalton including the constructed PE potential using the PElib<sup>78</sup> module in Dalton. For PE pFTAA calculations with sodium counterions, the QM core region was adapted to account for the lack of repulsion via transferable atomic all-electron pseudopotentials for the sodium ions.<sup>122</sup>

In the FDE approach, the supermolecular grid with a “good” Becke grid quality<sup>123</sup> was obtained with the ADF<sup>84</sup> code and the TD-DFT calculations with the Dalton code via the PyADF scripting environment.<sup>82,83</sup> In line with the calculations for the PE approach, the linear response calculations for pNA were performed with the CAM-B3LYP xc functional whereas for the environment molecules, a B3LYP xc functional was employed. In all FDE calculations, the additive xc functional BP86<sup>89,124</sup> and the kinetic energy functional PW91k<sup>125,126</sup> were applied for the non-additive contributions to the embedding potential. Three freeze-and-thaw cycles have been used throughout as this setting had been found to generally yield sufficient results.<sup>127</sup>

All calculations for pNA were performed with an aug-cc-pVDZ [see Supporting Information (SI)] or aug-cc-pVTZ basis set. For pFTAA, an aug-cc-pVDZ basis set was employed in all calculations. After calculating the five lowest excitations for the reference, as well as for the embedding calculations, they were sorted by the oscillator strength of the transition. The strongest  $\pi \rightarrow \pi^*$  transition (ensured *via* inspection of response vectors and orbitals) was chosen to be compared with other results.

## RESULTS AND DISCUSSION

We numerically compare calculated excitation energies and oscillator strengths for the models NOPOL, GSPOL, PE DPOL, and PE DPOL+EEF introduced in the Theoretical Background Section (see Table 1 for an overview). We generally report on shifts, *i.e.*, differences in excitation energy or oscillator strength of a solvation model to the vacuum case with the same structure of the dye. We denote these shifts as solvatochromic ( $\mathcal{S}$ )-shifts and  $\mathcal{F}$ -shifts for excitation energies and oscillator

strengths, respectively. We compare the shifts from embedding models to reference shifts obtained as the difference of a full quantum-chemical result to the vacuum case ( $\Delta_{\text{REF}}$ ). We generally denote these shifts by  $\Delta$ , *i.e.*,  $\Delta_{\text{NOPOL}}$  is the shift obtained with the NOPOL approximation and  $\Delta_{\text{GSPOL}}$  and  $\Delta_{\text{DPOL}}$  are defined analogously. The individual contributions are then defined with respect to the next lower model, *i.e.*,  $\Delta\Delta_{\text{GSPOL}} = \Delta_{\text{GSPOL}} - \Delta_{\text{NOPOL}}$ ,  $\Delta\Delta_{\text{DPOL}} = \Delta_{\text{DPOL}} - \Delta_{\text{GSPOL}}$ , and  $\Delta\Delta_{\text{EEF}} = \Delta(\text{DPOL+EEF}) - \Delta_{\text{DPOL}}$ . Additionally, we define  $\Delta\Delta(\text{DPOL+EEF}) = \Delta(\text{DPOL+EEF}) - \Delta_{\text{GSPOL}}$ .

Regarding the proportion of the single contributions ( $\Delta_{\text{NOPOL}}$ ,  $\Delta\Delta_{\text{GSPOL}}$ ,  $\Delta\Delta_{\text{DPOL}}$ ) to the total shift, we reference to the total supermolecular shift ( $\Delta_{\text{REF}}$ ) for both models, whenever available, and to the total DPOL shift ( $\Delta_{\text{DPOL}}$ ) when a supermolecular reference is unavailable (*cf.* Figure 5).

**para-Nitroaniline.** Our first test system is *para*-nitroaniline (pNA) in different water environments (see Figure 3). First,

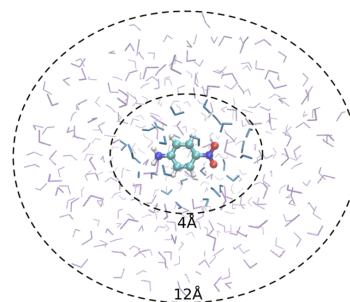
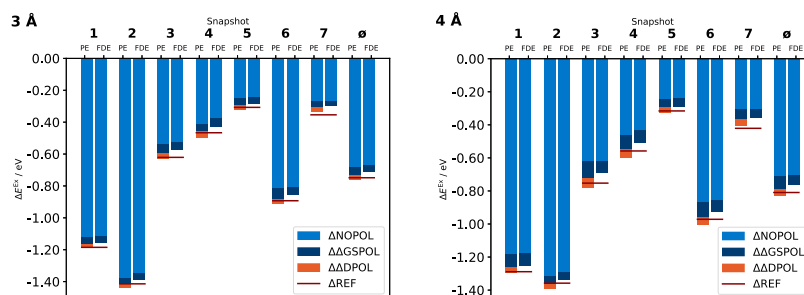


Figure 3. Example MD configuration of *para*-nitroaniline in a 4 Å and 12 Å water environment selection.

we investigate environment sizes of 3 and 4 Å for seven snapshots. Figure 4 shows the contributions to the total  $\mathcal{S}$ -shifts of pNA for the different solvent models and compares them to the supermolecular reference. It can be seen, that the total  $\mathcal{S}$ -shift varies largely for the different snapshots, independent of the environment size or embedding scheme used. Both, PE and FDE models reproduce the changes obtained in the reference calculations qualitatively correctly: For both, FDE and PE,  $\Delta_{\text{NOPOL}}$  is the largest contribution to the  $\mathcal{S}$ -shift (for 4 Å on average a proportion of 86 and 88% for FDE and PE, respectively). Thus, the  $\Delta_{\text{GSPOL}}$  contribution ( $\Delta\Delta_{\text{GSPOL}}$ ) is small (7% and 9% of the supermolecular  $\mathcal{S}$ -shift for 4 Å and FDE and PE, respectively). The PE  $\Delta\Delta_{\text{DPOL}}$  proportion lies below 5%. Thus,  $\Delta_{\text{NOPOL}}$  and  $\Delta_{\text{GSPOL}}$  for PE as well as FDE are in very good agreement with  $\Delta_{\text{REF}}$  (the largest average differences are  $-0.01$  eV, see Tables S3, S4, S7 and S8 in the SI).

Ultimately, the total  $\mathcal{S}$ -shifts for  $\Delta_{\text{GSPOL}}$  and  $\Delta_{\text{DPOL}}$  are in good agreement with  $\Delta_{\text{REF}}$  (Figure 4).  $\Delta_{\text{GSPOL}}$  on average slightly underestimates the total  $\mathcal{S}$ -shift, whereas adding the  $\Delta_{\text{DPOL}}$  leads to an (equally small) overestimation: For most snapshots, the  $\Delta_{\text{DPOL}}$  from PE is slightly higher than for the  $\Delta_{\text{REF}}$ , with an average deviation of  $-0.01$  and  $-0.02$  eV for the 3 Å and 4 Å system, respectively (see Tables S3 and S4 in the SI for the 3 Å system and Tables S7 and S8 in



**Figure 4.** Contributions from the different models to the total  $S$ -shifts and their average for different configurations of pNA in 3 and 4 Å environments of water obtained from an MD simulation and subsequently calculated in a PE and FDE framework and different orders of polarization contributions obtained in calculations with a aug-cc-pVTZ basis set.

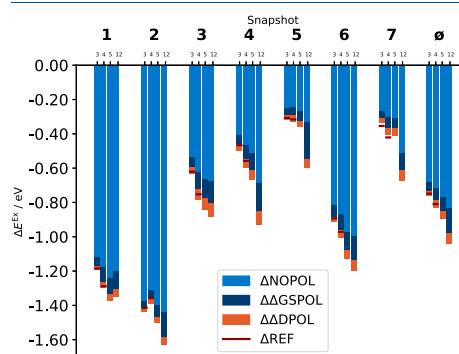
the SI for the 4 Å system). These deviations are smaller than those we expect from the differences in the applied xc functionals (the reference calculation is a full CAM-B3LYP calculation and in the determination of the PE embedding potential, B3LYP was employed for environment fragments).

It should be noted that for individual snapshots, the  $\Delta\Delta\text{GSPOL}$  proportion can exceed the average considerably. This is most pronounced for snapshot 5, which shows an overall small total  $S$ -shift: Here, the  $\Delta\Delta\text{GSPOL}$  proportion for both FDE and PE constitutes between 13% and 16% of the supermolecular  $S$ -shift for 3 and 4 Å environments, respectively. PE  $\Delta\Delta\text{DPOL}$  takes a proportion of 10% and 13% in the 3 and 4 Å environments, respectively.

We further observe a slight change in the proportions when increasing the environment size: When going from 3 to 4 Å environment, the average  $\Delta\Delta\text{GSPOL}$  proportion remains around 7% of the supermolecular shift for FDE and slightly increases from 7 to 9% for PE. The PE  $\Delta\Delta\text{DPOL}$  proportion on average increases from 4 to 5%. In absolute values, however, these contributions for all environment sizes are rather low, *i.e.*, at most  $-0.10$  eV for  $\Delta\Delta\text{GSPOL}$  for both PE and FDE and  $-0.06$  eV for  $\Delta\Delta\text{DPOL}$  in PE.

For PE, we also extended the environment further to 5 and 2 Å. The results for the  $S$ -shifts from all these calculations are depicted in Figure 5. The overall trend is a distinct increase in the size of the total  $S$ -shifts when extending from a 3 to a 12 Å environment (*i.e.*, the shift becomes more negative): on average it increases by  $-0.27$  eV. Again,  $\Delta\text{NOPOL}$  is the largest contribution and it increases with enlarged environment size: With the extension from the 3 to 4 Å environment it increases by  $-0.02$  eV on average, from the 4 to the 5 Å environment it increases by  $-0.06$  eV on average, and by  $-0.07$  eV when further extending to the 12 Å environment.  $\Delta\Delta\text{GSPOL}$  also increases: the increase is on average  $-0.02$  eV from the 3 Å to 4 Å environment, additional  $-0.02$  eV from the 4 Å to the 5 Å environment, and  $-0.05$  eV when extending to the 12 Å environment.  $\Delta\Delta\text{DPOL}$  also shows an increase when going from a 3 Å to a 12 Å environment. The absolute contribution on average increases from  $-0.03$  to  $-0.06$  eV. However, the average proportion of the total shift does not steadily increase: From a 3 to a 4 Å environment it changes from  $-0.03$  eV (4% of  $\Delta\text{DPOL}$ ) to  $-0.04$  eV (5% of  $\Delta\text{DPOL}$ ), for a 5 Å environment it decreases to  $-0.04$  eV (4% of  $\Delta\text{DPOL}$ ) and

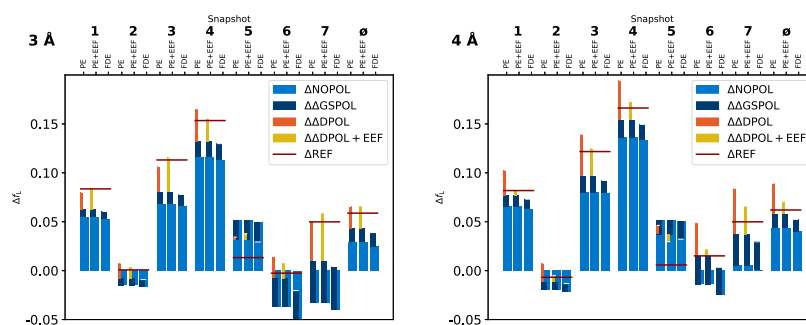
increases for a 12 Å environment to  $-0.06$  eV (6% of  $\Delta\text{DPOL}$ ).



**Figure 5.** Contributions from the different models to the total  $S$ -shifts and their average for different configurations of pNA in 3, 4, 5, and 12 Å environments of water obtained from an MD simulation and subsequently calculated in a PE framework and different orders of polarization contributions obtained in calculations with a aug-cc-pVTZ basis set.

As discussed above, for the snapshots with smaller total shifts,  $\Delta\Delta\text{GSPOL}$  can exceed the average considerably: Here, we again look at snapshot 5 for which the  $\Delta\Delta\text{GSPOL}$  proportion changes from 13% ( $-0.05$  eV) in a 3 Å environment to 15% ( $-0.05$  eV) in a 4 Å environment, 14% ( $-0.05$  eV) in a 5 Å environment, and 37% ( $-0.22$  eV) in a 12 Å environment, where all percentages refer to the  $\Delta\text{DPOL}$  shift. Thus, for this particular snapshot, the  $\Delta\text{NOPOL}$  accounts for 55% ( $-0.33$  eV) of  $\Delta\text{DPOL}$  for a 12 Å environment model.

In a previous study on pNA in a water environment (2–6 molecules) using an EOM-CCSD/EFP scheme, Slipchenko et al.<sup>39</sup> found the  $\Delta\text{NOPOL}$  proportion of the excitation energy to be of similar amount (80%) as was obtained in our calculations (86–88%). The  $\Delta\Delta\text{DPOL}$  proportion was determined to be 3–8% which is in good agreement with our result of 4–5%. In their study, increasing the number of water molecules used in



**Figure 6.** Contributions from the different models to the total  $\mathcal{F}$ -shifts and their average for different configurations of pNA in 3 and 4 Å environments of water obtained from an MD simulation and subsequently calculated in a PE and FDE framework, different orders of polarization contributions and added EEF effects obtained in calculations with a aug-cc-pVTZ basis set. (Full data in Tables S15, S16, S19, and S20).

the solvation led to an increase in  $\Delta\Delta\text{DPOL}$ , similar to the increase observed in our results.

In a study by Sneskov et al.,<sup>40</sup> the average polarization contribution was obtained from 100 snapshots for pNA in a 12 Å environment of water and for PE only. Both the  $\Delta\Delta\text{GSPOL}$  and the  $\Delta\Delta\text{DPOL}$  proportion are higher than in the present study's result: 19–21 and 13% for the  $\Delta\Delta\text{GSPOL}$  and  $\Delta\Delta\text{DPOL}$  proportion, respectively. These authors also noted that the variation of these values is dependent on the individual snapshot, similar to what is seen in Figure 5. In an FDE context, absolute values for  $\Delta\Delta\text{DPOL}$  were obtained from mutual optimization with excited-state densities at the study of Daday et al.<sup>8</sup> The magnitudes (0.01–0.22 or  $-0.02$ – $-0.15$  eV, depending on the description of the excited-state density) are similar to our results for  $\Delta\Delta\text{DPOL}$  ranging between 0.02–0.06 eV.

Equivalently to the  $\mathcal{S}$ -shifts discussed above, Figure 6 displays the change in the oscillator strength ( $\mathcal{F}$ -shift) of the strongest ( $\pi \rightarrow \pi^*$ ) transition for the various PE and FDE solvation models. We find that the  $\mathcal{F}$ -shift of the reference calculations displays larger sensitivity than the  $\mathcal{S}$ -shifts with respect to both, the size of the system and snapshot. This is in line with previous comparisons of different electronic structure methods, showing oscillator strengths to be more sensitive to the employed electronic structure methods.<sup>128</sup>

In contrast to the discussion of  $\mathcal{S}$ -shifts, we here omit the presentation of single contributions ( $\Delta\text{NOPOL}$ ,  $\Delta\Delta\text{GSPOL}$ , and  $\Delta\Delta\text{DPOL}$ ) as percentage of the total shift since the oscillator strengths are generally smaller than the excitations energies and even small changes can lead to large percent-wise changes.

The reference  $\mathcal{F}$ -shifts are on average 0.06 for both 3 and 4 Å. The FDE and PE  $\text{NOPOL}$  models both give an average  $\mathcal{F}$ -shift of 0.03 for 3 Å and 0.04 for 4 Å. Generally,  $\Delta\text{NOPOL}$  is estimated similarly by FDE and PE, the largest deviation being 0.01.  $\Delta\text{NOPOL}$  is often the largest contribution to the total  $\mathcal{F}$ -shift, but is much less dominant compared to the  $\mathcal{S}$ -shift. Notably, the  $\Delta\text{NOPOL}$  results alone are often rather far from the total shifts (most obvious in snapshots 6 and 7 for both 3 and 4 Å).  $\Delta\Delta\text{GSPOL}$  generally improves the results for PE and FDE similarly: The largest deviation between FDE and PE amounts to less than 0.01 for both the 3 and 4 Å systems.

In contrast to the  $\mathcal{S}$ -shift,  $\Delta\Delta\text{DPOL}$  can be rather large for  $\mathcal{F}$ -shifts: In some cases (see snapshots 2 and 6) both  $\Delta\Delta\text{GSPOL}$  (for FDE and PE) and  $\Delta\Delta\text{DPOL}$  (for PE) correct the  $\mathcal{F}$ -shift in the opposite direction of  $\Delta\text{NOPOL}$ . Especially for the 4 Å environment, we observe large  $\Delta\Delta\text{DPOL}$  values often leading to larger deviation of  $\Delta\text{DPOL}$  from the  $\Delta\text{REF}$  than  $\Delta\text{GSPOL}$ . This over-correction of  $\Delta\text{DPOL}$  led us to investigate local field effects on the oscillator strength by means of effective external field (Figure 6). While for the 3 Å system on average only a small increase in  $\mathcal{F}$ -shift can be observed (below 0.01), the total  $\mathcal{F}$ -shift for the 4 Å system decreases significantly (for all snapshots), leading to an improved result compared to the reference: The average deviation is 0.03 for  $\Delta\text{DPOL}$  compared to and less than 0.01 for  $\Delta\text{DPOL} + \text{EEF}$ .

We finally note that all the discussed results are obtained with aug-cc-pVTZ but for calculations with an aug-cc-pVDZ basis set, the same trends can be observed (see Figures S1, S3 and Tables S1–S24 in the SI). The proportions of the single contributions are in line with those obtained with an aug-cc-pVTZ basis set.

In summary, we observe similar results for FDE  $\text{GSPOL}$  and PE  $\text{GSPOL}$ , suggesting that the additional quantum-mechanical contribution and real-space treatment in FDE have only a minor effect in this case. The supermolecular reference of excitation energies of pNA in the 3 and 4 Å water environment is well in line with FDE results as well as the PE values with or without including differential polarization effects. When going to larger systems sizes within the PE model, we observe an increased  $\Delta\Delta\text{GSPOL}$  proportion to the  $\mathcal{S}$ -shift, but an unclear trend for the differential polarization.

The observations for the oscillator strengths are similar, though not identical to those for the electronic excitation energies. In particular, we observe a larger snapshot dependence and the differential polarization contribution in the PE calculations is larger than for the excitation energies. We further observe an overcorrection for the 4 Å environments due  $\Delta\Delta\text{DPOL}$ , which can be largely canceled by accounting for local field effects (EEF).

**Pentameric formyl thiophene acetic acid (pFTAA).** Our second test case, pFTAA, is in contrast to pNA highly negatively charged (4 $-$ ). The solvated system is shown in Figure 7. It is hence, more challenged by possible electron-

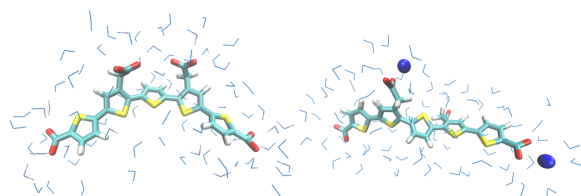


Figure 7. Selected snapshots of pFTAA in a pure 3 Å water environment (left, snapshot 1) and additionally including sodium ions in close vicinity to pFTAA (right, snapshot 2).

Table 2. Contributions from the Different Models to the Total  $\mathcal{S}$ -Shift in eV for Different Configurations of pFTAA in a 3 Å Environment of Water Obtained from an MD Simulation and Subsequently Calculated in a PE and FDE Framework and Different Orders of Polarization Contributions<sup>a</sup>

S	FDE			PE					$\Delta_{\text{REF}}$
	$\Delta_{\text{NOPOL}}$	$\Delta\Delta_{\text{GSPOL}}$	$\Delta_{\text{GSPOL}}$	$\Delta_{\text{NOPOL}}$	$\Delta\Delta_{\text{GSPOL}}$	$\Delta\Delta_{\text{DPOL}}$	$\Delta_{\text{GSPOL}}$	$\Delta_{\text{DPOL}}$	
1	0.15	0.01	0.17	0.16	0.01	-0.03	0.17	0.14	0.14
2**	0.04	0.02	0.07	0.01	0.05	-0.08	0.07	0.04	0.03

<sup>a</sup>All snapshots containing (two) sodium ions in close proximity are marked with \*\* and incorporate pseudopotentials in the PE calculations.

Table 3. Contributions From the Different Models to the Total  $\mathcal{F}$ -Shifts for Different Configurations of pFTAA in a 3 Å Environment of Water Obtained from an MD Simulation and Subsequently Calculated in a PE and FDE Framework and Different Orders of Polarization Contributions Obtained with a Aug-cc-pVDZ Basis Set<sup>a</sup>

$\mathcal{F}$	FDE			PE							$\Delta_{\text{REF}}$
	$\Delta_{\text{NOPOL}}$	$\Delta\Delta_{\text{GSPOL}}$	$\Delta_{\text{TOT}_{\text{GSPOL}}}$	$\Delta_{\text{NOPOL}}$	$\Delta\Delta_{\text{GSPOL}}$	$\Delta\Delta_{\text{DPOL}}$	$\Delta\Delta_{\text{EEF}}$	$\Delta_{\text{GSPOL}}$	$\Delta_{\text{DPOL}}$	$\Delta_{\text{DPOL}} + \text{EEF}$	
1	0.294	0.053	0.348	0.320	0.060	0.063	-0.029	0.380	0.443	0.414	0.388
2**	0.140	0.044	0.183	0.143	0.024	0.052	-0.076	0.167	0.219	0.143	-0.007

<sup>a</sup>All snapshots containing (two) sodium ions in close proximity are marked with \*\* and incorporate pseudopotentials in the PE calculations.

spill-out effects, which makes it difficult to describe classical models like PE. The FDE is expected to be less prone to electron-spill-out effects due to the approximate quantum contributions in the embedding potential [cf. eq 6].<sup>129,130</sup> Indeed, we found that in the five snapshots that contained sodium cations close to the pFTAA chromophore, the standard PE model broke down. The breakdown manifests in too low excitation energies and varying oscillator strengths for the excitation of interest. The electron spill-out was further revealed by analyzing the contributing orbitals in the response solution vectors. We counteracted the electron spill-out in these cases by placing atomic pseudopotentials on sodium ions<sup>122</sup> which in all cases led to meaningful results. The PE results for all these snapshots can be found in Tables S25 and S28 in the SI.

Here, we focus on the discussion of two, representative snapshots: one with two sodium ions and one without sodium ions in close proximity to the pFTAA solute. Again we compare the performance of the PE and FDE models (see Table 2). The supersystem reference shifts ( $\Delta_{\text{REF}}$ ) are well reproduced for both the FDE and PE models.  $\Delta_{\text{NOPOL}}$  is by far the largest contribution, while  $\Delta\Delta_{\text{GSPOL}}$  is small. Both,  $\Delta_{\text{NOPOL}}$  and  $\Delta\Delta_{\text{GSPOL}}$ , are close to identical for FDE and PE.  $\Delta\Delta_{\text{DPOL}}$  is of similar magnitude as  $\Delta\Delta_{\text{GSPOL}}$  but points in the opposite direction.

Table 3 shows the  $\mathcal{F}$ -shifts for the PE and FDE embedding model employing the two snapshots.  $\Delta_{\text{NOPOL}}$  and  $\Delta\Delta_{\text{GSPOL}}$  are similar for FDE and PE, where the  $\Delta\Delta_{\text{GSPOL}}$  contributions are significantly smaller than the  $\Delta_{\text{NOPOL}}$  contributions.  $\Delta\Delta_{\text{DPOL}}$  in the PE model for the

two investigated snapshots is also rather small and of similar magnitude as the  $\Delta\Delta_{\text{GSPOL}}$  contributions. The  $\Delta\Delta_{\text{EEF}}$  contribution is of similar magnitude as the  $\Delta\Delta_{\text{DPOL}}$  contribution but of opposing sign.

For snapshot 1, both  $\Delta_{\text{GSPOL}}$  and  $\Delta_{\text{DPOL}}$  are in reasonable agreement with the reference value of 0.388: For  $\Delta_{\text{GSPOL}}$  we obtain deviations of 0.04 and 0.01, respectively for the FDE and the PE models. As also seen for pNA,  $\Delta\Delta_{\text{DPOL}}$  overcorrects the  $\mathcal{F}$ -shift (leading to a deviation of 0.05 to the reference), whereas introducing EEF effects ( $\Delta_{\text{DPOL}} + \text{EEF}$ ) again brings the value closer to the reference.

## ■ SUMMARY AND CONCLUSIONS

We have investigated the influence of different approximations in quantum- and classical embedding schemes on the excitation energies and oscillator strengths for pNA and pFTAA in water. We chose Frozen-density embedding (FDE) and polarizable embedding (PE) schemes to represent quantum and a classical embedding scheme, respectively. Unlike previous comparisons,<sup>32,35</sup> we employed an FDE framework that complies to a large degree with the PE implementation in Dalton.<sup>34,116,131</sup> In particular, we performed the mutual polarization of subsystems within FDE in the PyADF scripting environment<sup>82,83</sup> using Dalton<sup>85</sup> for all (TD-)DFT calculations thus enabling a one-to-one comparison of the two methods.

With this computational setup at hand, we performed a detailed analysis of the different contributions, i.e., static

electrostatics (no polarization), ground-state polarization, differential polarization, and quantum-mechanical effects in the FDE and PE models: The target systems were pNA and pFTAA in an explicit water solvent for which we compared the solvent shift of excitation energies and oscillator strengths to supermolecular TDDFT calculations.

We find that FDE and PE perform similarly with the inclusion of static environmental densities and ground-state polarization, respectively. Since these two contributions are dominating, the solvochromatic ( $\mathcal{S}$ -)shift for both pNA and pFTAA, FDE and PE both achieve good agreement with the reference  $\mathcal{S}$ -shifts. This also holds when neglecting differential polarization effects.

The effect of different embedding contributions on the  $\mathcal{F}$ -shifts has not previously been investigated systematically. We find that the differential polarization effect on  $\mathcal{F}$ -shifts is more pronounced than that of the  $\mathcal{S}$ -shifts in standard PE. This effect on the  $\mathcal{F}$ -shifts is, however, reduced by the incorporation of external effective field effects, so that for a 4 Å environment of pNA, the average  $\mathcal{F}$ -shift is similarly well described with and without differential polarization. For individual snapshots, however, the effect of differential polarization can be sizable, both with and without including external effective field effects. In these cases, external field effects improve the agreement with the supersystem reference. We could further show that the severe electron-spill-out issues preventing traditional PE calculations on the highly anionic pFTAA dye with sodium ions in close proximity could be largely reduced by atomic pseudopotentials on the sodium ions.

In summary, we find a similar performance for FDE and PE on excitation energies as well as average oscillator strengths. For the excitation energies, this is similar to what has previously been found.<sup>32,35</sup> Our investigation on  $\mathcal{F}$ -shifts, however, lead us to recommend employing both differential polarization and external effective field (EEF) effects (or neglecting both), since accurate oscillator strengths with PE DPOL required the incorporation of EEF effects. Unlike previous investigations, we also compared PE and FDE for a challenging system, namely the highly charged anionic pFTAA. Here, PE and FDE can be brought to become similar, but pseudopotentials are essential on nearby cations to avoid electron-spill-out effects in PE. We thus recommend the use of such potentials in PE calculations on highly charged anions.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.3c02540>.

Excitation energies and  $\mathcal{S}$ -shifts as well as the oscillator strengths and  $\mathcal{F}$ -shifts for all calculated snapshots for pNA and pFTAA in all environment sizes (3, 4, 5, 12 Å for pNA and 3 Å for pFTAA) for the PE and FDE model obtained with different basis sets (aug-cc-pVDZ and aug-cc-pVTZ for pNA; aug-cc-pVDZ for pFTAA) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

Erik D. Hedegård – Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, 5230 Odense,

Denmark; [orcid.org/0000-0002-3868-1206](https://orcid.org/0000-0002-3868-1206);  
Email: [erdh@sdu.dk](mailto:erdh@sdu.dk)

Carolin König – Institute of Physical Chemistry and Electrochemistry, Leibniz University Hannover, 30167 Hannover, Germany; [orcid.org/0000-0001-8931-4337](https://orcid.org/0000-0001-8931-4337);  
Email: [carolin.koenig@pci.uni-hannover.de](mailto:carolin.koenig@pci.uni-hannover.de)

### Authors

Marina Jansen – Institute of Physical Chemistry and Electrochemistry, Leibniz University Hannover, 30167 Hannover, Germany; [orcid.org/0000-0002-0293-6959](https://orcid.org/0000-0002-0293-6959)

Peter Reinholdt – Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, 5230 Odense, Denmark; [orcid.org/0000-0003-2406-700X](https://orcid.org/0000-0003-2406-700X)

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.jpca.3c02540>

### Author Contributions

M.J.: methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, review and editing, visualization; P.R.: molecular dynamics for pNA in water; E.D.H.: conceptualization, validation, formal analysis, writing—review & editing; C.K.: conceptualization, validation, formal analysis, writing—review & editing, supervision, project administration.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) through the Emmy Noether Young Group Leader Programme (project KO 5423/1-1). EDH thanks The Villum Foundation, Young Investigator Program (grant no. 29412), the Swedish Research Council (grant no. 2019-04205), and Independent Research Fund Denmark (grant no. 0252-00002B and grant no. 2064-00002B) for support. The computations were carried out on the cluster system at the Leibniz University Hannover, Germany, which is funded by the Leibniz University Hannover, the Lower Saxony Ministry of Science and Culture (MWK) and the DFG.

## ■ REFERENCES

- Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- Singh, U. C.; Kollman, P. A. A Combined Ab Initio Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the CH<sub>3</sub>Cl + Cl<sup>-</sup> Exchange Reaction and Gas Phase Protonation of Polyethers. *J. Comp. Chem.* **1986**, *7*, 718–730.
- Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.
- Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- Thompson, M. A. QM/MMpol: A Consistent Model for Solute/Solvent Polarization. Application to the Aqueous Solvation and Spectroscopy of Formaldehyde, Acetaldehyde, and Acetone. *J. Phys. Chem. A* **1996**, *100*, 14492–14507.
- Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholles, G. D.; Mennucci, B. Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

- (7) Defusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. Modeling Solvent Effects on Electronic Excited States. *J. Phys. Chem. Lett.* **2011**, *2*, 2184–2192.
- (8) Daday, C.; König, C.; Valsson, O.; Neugebauer, J.; Filippi, C. State-Specific Embedding Potentials for Excitation-Energy Calculations. *J. Chem. Theory Comput.* **2013**, *9*, 2355–2367.
- (9) Boulanger, E.; Harvey, J. N. QM/MM Methods for Free Energies and Photochemistry. *Curr. Opin. Struct. Biol.* **2018**, *49*, 72–76.
- (10) Thompson, M. A.; Schenter, G. K. Excited States of the Bacteriochlorophyll b Dimer of *Rhodospseudomonas Viridis*: A QM/MM Study of the Photosynthetic Reaction Center That Includes MM Polarization. *J. Phys. Chem. B* **1995**, *99*, 6374–6386.
- (11) Gao, J.; Byun, K. Solvent Effects on the  $N\pi^*$  Transition of Pyrimidine in Aqueous Solution. *Theor. Chem. Acc.* **1997**, *96*, 151–156.
- (12) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (13) List, N. H.; Olsen, J. M. H.; Kongsted, J. Excited States in Large Molecular Systems through Polarizable Embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20234–20250.
- (14) Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable Embedding QM/MM: The Future Gold Standard for Complex (Bio)Systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433–14448.
- (15) Loco, D.; Lagardère, L.; Adjoua, O.; Piquemal, J. P. Atomistic Polarizable Embeddings: Energy, Dynamics, Spectroscopy, and Reactivity. *Acc. Chem. Res.* **2021**, *54*, 2812–2822.
- (16) Lipparini, F.; Mennucci, B. Hybrid QM/Classical Models: Methodological Advances and New Applications. *Chem. Phys. Rev.* **2021**, *2*, No. 041303.
- (17) Jansen, M.; Nguyen, T. M. N.; Hedegård, E. D.; König, C. Quantum-Derived Embedding Schemes for Local Optical Spectra. In *SPR Chemical Modelling*, 17th ed.; Bahmann, H.; Tremblay, J. C., Eds.; The Royal Society of Chemistry, 2022; pp 24–60.
- (18) Jensen, L.; Van Duijnen, P. T.; Sniijders, J. G. A Discrete Solvent Reaction Field Model for Calculating Frequency-Dependent Hyperpolarizabilities of Molecules in Solution. *J. Chem. Phys.* **2003**, *119*, 12998–13006.
- (19) Jensen, L.; Van Duijnen, P. T.; Sniijders, J. G. A Discrete Solvent Reaction Field Model within Density Functional Theory. *J. Chem. Phys.* **2003**, *118*, 514–521.
- (20) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.
- (21) Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. The Effective Fragment Potential: A General Method for Predicting Intermolecular Interaction. In *Annual Reports in Computational Chemistry*; Elsevier, 2007; Chapter 10, Vol. 3, pp 177–193.
- (22) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate First Principles Model Potentials for Intermolecular Interactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553–578.
- (23) Loco, D.; Polack, E.; Caprasecca, S.; Lagardère, L.; Lipparini, F.; Piquemal, J. P.; Mennucci, B. A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654–3661.
- (24) Yang, W. Direct Calculation of Electron Density in Density-Functional Theory. *Phys. Rev. Lett.* **1991**, *66*, 1438–1441.
- (25) Senatore, G.; Subbaswamy, K. R. Density Dependence of the Dielectric Constant of Rare-Gas Crystals. *Phys. Rev. B* **1986**, *34*, 5754–5757.
- (26) Johnson, M. D.; Subbaswamy, K. R.; Senatore, G. Hyperpolarizabilities of Alkali Halide Crystals Using the Local-Density Approximation. *Phys. Rev. B* **1987**, *36*, 9202–9211.
- (27) Cortona, P. Self-Consistently Determined Properties of Solids without Band-Structure Calculations. *Phys. Rev. B* **1991**, *44*, 8454–8458.
- (28) Jacob, C. R.; Neugebauer, J. Subsystem Density-Functional Theory. *WIREs Comput. Mol. Sci.* **2014**, *4*, 325–362.
- (29) Wesolowski, T. A.; Warshel, A. Frozen Density Functional Approach for Ab Initio Calculations of Solvated Molecules. *J. Phys. Chem. C* **1993**, *97*, 8050–8053.
- (30) Laricchia, S.; Fabiano, E.; Constantin, L. A.; Della Sala, F. Generalized Gradient Approximations of the Noninteracting Kinetic Energy from the Semiclassical Atom Theory: Rationalization of the Accuracy of the Frozen Density Embedding Theory for Nonbonded Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 2439–2451.
- (31) Filippi, C.; Buda, F.; Guidoni, L.; Sinicropi, A. Bathochromic Shift in Green Fluorescent Protein: A Puzzle for QM/MM Approaches. *J. Chem. Theory Comput.* **2012**, *8*, 112–124.
- (32) Daday, C.; Curutchet, C.; Sinicropi, A.; Mennucci, B.; Filippi, C. Chromophore-Protein Coupling beyond Nonpolarizable Models: Understanding Absorption in Green Fluorescent Protein. *J. Chem. Theory Comput.* **2015**, *11*, 4825–4839.
- (33) Schwabe, T.; Beerepoot, M. T.; Olsen, J. M. H.; Kongsted, J. Analysis of Computational Models for an Accurate Study of Electronic Excitations in GFP. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2582–2588.
- (34) Olsen, J. M. H.; Kongsted, J. Molecular Properties through Polarizable Embedding. In *Advances in Quantum Chemistry*; Elsevier, 2011; Vol. 61, pp 107–143.
- (35) Jacob, C. R.; Neugebauer, J.; Jensen, L.; Visscher, L. Comparison of Frozen-Density Embedding and Discrete Reaction Field Solvent Models for Molecular Properties. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2349–2359.
- (36) Gomes, A. S. P.; Jacob, C. R. Quantum-Chemical Embedding Methods for Treating Local Electronic Excitations in Complex Chemical Systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2012**, *108*, 222–277.
- (37) Fujisawa, T.; Terazima, M.; Kimura, Y.; Maroncelli, M. Resonance Raman Study of the Solvation of P-Nitroaniline in Supercritical Water. *Chem. Phys. Lett.* **2006**, *430*, 303–308.
- (38) Kosenkov, D.; Slipchenko, L. V. Solvent Effects on the Electronic Transitions of P-Nitroaniline: A QM/EFP Study. *J. Phys. Chem. A* **2011**, *115*, 392–401.
- (39) Slipchenko, L. V. Solvation of the Excited States of Chromophores in Polarizable Environment: Orbital Relaxation versus Polarization. *J. Phys. Chem. A* **2010**, *114*, 8824–8830.
- (40) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the Effects of Polarization in QM/MM Excited State Calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.
- (41) Klingstedt, T.; Shirani, H.; Mahler, J.; Wegenast-Braun, B. M.; Nyström, S.; Goedert, M.; Jucker, M.; Nilsson, K. P. R. Distinct Spacing between Anionic Groups: An Essential Chemical Determinant for Achieving Thiophene-Based Ligands to Distinguish  $\beta$ -Amyloid or Tau Polymorphic Aggregates. *Chem. - Eur. J.* **2015**, *21*, 9072–9082.
- (42) Sjöqvist, J.; Linares, M.; Lindgren, M.; Norman, P. Molecular Dynamics Effects on Luminescence Properties of Oligothiophene Derivatives: A Molecular Mechanics–Response Theory Study Based on the CHARMM Force Field and Density Functional Theory. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17532–17542.
- (43) Sjöqvist, J.; Linares, M.; Mikkelsen, K. V.; Norman, P. QM/MM-MD Simulations of Conjugated Polyelectrolytes: A Study of Luminescent Conjugated Oligothiophenes for Use as Biophysical Probes. *J. Phys. Chem. A* **2014**, *118*, 3419–3428.
- (44) Gustafsson, C.; Linares, M.; Norman, P. Quantum Mechanics/Molecular Mechanics Density Functional Theory Simulations of the Optical Properties Fingerprinting the Ligand-Binding of Pentameric Formyl Thiophene Acetic Acid in Amyloid- $\beta$ (1–42). *J. Phys. Chem. A* **2020**, *124*, 875–888.
- (45) Gustafsson, C.; Shirani, H.; Leira, P.; Rehn, D. R.; Linares, M.; Nilsson, K. P. R.; Norman, P.; Lindgren, M. Deciphering the Electronic Transitions of Thiophene-Based Donor-Acceptor-Donor Pentameric Ligands Utilized for Multimodal Fluorescence Microscopy of Protein Aggregates. *ChemPhysChem* **2021**, *22*, 323–335.

- (46) Wesolowski, T. A.; Shedde, S.; Zhou, X. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* **2015**, *115*, 5891–5928.
- (47) Krishtal, A.; Sinha, D.; Genova, A.; Pavanello, M. Subsystem Density-Functional Theory as an Effective Tool for Modeling Ground and Excited States, Their Dynamics and Many-Body Interactions. *J. Phys.: Condens. Matter* **2015**, *27*, No. 183202.
- (48) Sun, Q.; Chan, G. K. L. Quantum Embedding Theories. *Acc. Chem. Res.* **2016**, *49*, 2705–2712.
- (49) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F. Exact Nonadditive Kinetic Potentials for Embedded Density Functional Theory. *J. Chem. Phys.* **2010**, *133*, No. 84103.
- (50) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.
- (51) Khait, Y. G.; Hoffmann, M. R. On the Orthogonality of Orbitals in Subsystem Kohn-Sham Density Functional Theory. In *Annual Reports in Computational Chemistry*; Elsevier, 2012; Chapter 3, Vol. 8, pp 53–70.
- (52) Chulhai, D. V.; Jensen, L. Frozen Density Embedding with External Orthogonality in Delocalized Covalent Systems. *J. Chem. Theory Comput.* **2015**, *11*, 3080–3088.
- (53) Hégyely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. Exact Density Functional and Wave Function Embedding Schemes Based on Orbital Localization. *J. Chem. Phys.* **2016**, *145*, No. 64107.
- (54) Chulhai, D. V.; Goodpaster, J. D. Improved Accuracy and Efficiency in Quantum Embedding through Absolute Localization. *J. Chem. Theory Comput.* **2017**, *13*, 1503–1508.
- (55) Culpitt, T.; Brorsen, K. R.; Hammes-Schiffer, S. Communication: Density Functional Theory Embedding with the Orthogonality Constrained Basis Set Expansion Procedure. *J. Chem. Phys.* **2017**, *146*, No. 211101.
- (56) Ding, F.; Manby, F. R.; Miller, T. F. Embedded Mean-Field Theory with Block-Orthogonalized Partitioning. *J. Chem. Theory Comput.* **2017**, *13*, 1605–1615.
- (57) Lee, S. J. R.; Welborn, M.; Manby, F. R.; Miller, T. F. Projection-Based Wavefunction-in-DFT Embedding. *Acc. Chem. Res.* **2019**, *52*, 1359–1368.
- (58) Graham, D. S.; Wen, X.; Chulhai, D. V.; Goodpaster, J. D. Robust, Accurate, and Efficient: Quantum Embedding Using the Huzinaga Level-Shift Projection Operator for Complex Systems. *J. Chem. Theory Comput.* **2020**, *16*, 2284–2295.
- (59) Graham, D. S.; Wen, X.; Chulhai, D. V.; Goodpaster, J. D. Huzinaga Projection Embedding for Efficient and Accurate Energies of Systems with Localized Spin-Densities. *J. Chem. Phys.* **2022**, *156*, No. 54112.
- (60) Bensberg, M.; Neugebauer, J. Automatic Basis-Set Adaptation in Projection-Based Embedding. *J. Chem. Phys.* **2019**, *150*, No. 184104.
- (61) Bensberg, M.; Neugebauer, J. Direct Orbital Selection for Projection-Based Embedding. *J. Chem. Phys.* **2019**, *150*, No. 214106.
- (62) Wesolowski, T. A.; Weber, J. Kohn-Sham Equations with Constrained Electron Density: An Iterative Evaluation of the Ground-State Electron Density of Interacting Molecules. *Chem. Phys. Lett.* **1996**, *248*, 71–76.
- (63) Stone, A. J. Distributed Multipole Analysis, or How to Describe a Molecular Charge Distribution. *Chem. Phys. Lett.* **1981**, *83*, 233–239.
- (64) Stone, A. *The Theory of Intermolecular Forces*; Oxford University Press, 2013.
- (65) Applequist, J.; Carl, J. R.; Fung, K.-K. An Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*, 2952–2960.
- (66) Casida, M. E.; Wesolowski, T. A. Generalization of the Kohn-Sham Equations with Constrained Electron Density Formalism and Its Time-Dependent Response Theory Formulation. *Int. J. Quantum Chem.* **2004**, *96*, 577–588.
- (67) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. Calculation of Local Excitations in Large Systems by Embedding Wave-Function Theory in Density-Functional Theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353–5362.
- (68) Neugebauer, J.; Curutchet, C.; Muñoz-Losa, A.; Mennucci, B.; Mennucci, B. A Subsystem TDDFT Approach for Solvent Screening Effects on Excitation Energy Transfer Couplings. *J. Chem. Theory Comput.* **2010**, *6*, 1843–1851.
- (69) Daday, C.; König, C.; Neugebauer, J.; Filippi, C. Wavefunction in Density Functional Theory Embedding for Excited States: Which Wavefunctions Which Densities? *ChemPhysChem* **2014**, *15*, 3205–3217.
- (70) Pal, P. P.; Liu, P.; Jensen, L. Polarizable Frozen Density Embedding with External Orthogonalization. *J. Chem. Theory Comput.* **2019**, *15*, 6588–6596.
- (71) Scholz, L.; Neugebauer, J. Protein Response Effects on Cofactor Excitation Energies from First Principles: Augmenting Subsystem Time-Dependent Density-Functional Theory with Many-Body Expansion Techniques. *J. Chem. Theory Comput.* **2021**, *17*, 6105–6121.
- (72) Harshan, A. K.; Bronson, M. J.; Jensen, L. Local-Field Effects in Linear Response Properties within a Polarizable Frozen Density Embedding Method. *J. Chem. Theory Comput.* **2022**, *18*, 380–393.
- (73) Autschbach, J.; Ziegler, T. Calculating Molecular Electric and Magnetic Properties from Time-Dependent Density Functional Response Theory. *J. Chem. Phys.* **2002**, *116*, 891–896.
- (74) Neugebauer, J. Couplings between Electronic Transitions in a Subsystem Formulation of Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2007**, *126*, No. 134116.
- (75) Neugebauer, J. Chromophore-Specific Theoretical Spectroscopy: From Subsystem Density Functional Theory to Mode-Specific Vibrational Spectroscopy. *Phys. Rep.* **2010**, *489*, 1–87.
- (76) Scholz, L.; Tölle, J.; Neugebauer, J. Analysis of Environment Response Effects on Excitation Energies within Subsystem-Based Time-Dependent Density-Functional Theory. *Int. J. Quantum Chem.* **2020**, *120*, No. e26213.
- (77) List, N. H.; Jensen, H. J. A.; Kongsted, J. Local Electric Fields and Molecular Properties in Heterogeneous Environments through Polarizable Embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 10070–10080.
- (78) Olsen, J. M. H.; List, N. H.; Steinmann, C.; Steindal, A. H.; Nørby, M. S.; Reinholdt, P. P. Elib: The Polarizable Embedding library. **2020**.
- (79) Olsen, J. M. H.; List, N. H.; Steinmann, C. PEAS (Polarizable Embedding Assistant Script).
- (80) Galván, I. F.; Vacher, M.; Alavi, A.; et al. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925–5964.
- (81) Gagliardi, L.; Lindh, R.; Karlström, G. Local Properties of Quantum Chemical Systems: The LoProp Approach. *J. Chem. Phys.* **2004**, *121*, 4494–4500.
- (82) Jacob, C. R.; Beyhan, S. M.; Bulo, R. E.; Gomes, A. S. P.; Götz, A. W.; Kiewisch, K.; Sikkema, J.; Visscher, L. PyADF - A Scripting Framework for Multiscale Quantum Chemistry. *J. Comput. Chem.* **2011**, *32*, 2328–2338.
- (83) Jacob, C. *PyADF*; Version 0.98, 2021.
- (84) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, S. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.
- (85) Aidas, K.; Angeli, C.; Bak, K. L.; et al. The Dalton Quantum Chemistry Program System. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269–284.
- (86) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- (87) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.



- (88) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem. D* **1993**, *97*, 10269–10280.
- (89) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (90) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (91) Becke, A. D. Becke's 3 Parameter Functional Combined with the Non-Local Correlation LYP. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (92) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.
- (93) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (94) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- (95) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (96) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.
- (97) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N-Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (98) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (99) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (100) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. ROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (101) Páll, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E. In *Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS*, Solving Software Challenges for Exascale: International Conference on Exascale Applications and Software, EASC 2014, Stockholm, Sweden; Springer International Publishing, 2014; pp 3–27.
- (102) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.
- (103) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (104) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- (105) Lindahl, E.; Hess, B.; Van Der Spoel, D. GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Model.* **2001**, *7*, 306–317.
- (106) Berendsen, H. J. C.; vander Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
- (107) Lindahl, E.; Abraham, M. J.; Hess, B.; van der Spoel, D. *GROMACS 2020 Source code*; Zenodo: Geneva, Switzerland, 2019.
- (108) Feller, S. E.; MacKerell, A. D. An Improved Empirical Potential Energy Function for Molecular Simulations of Phospholipids. *J. Phys. Chem. B* **2000**, *104*, 7510–7515.
- (109) Klauda, J. B.; Brooks, B. R.; MacKerell, A. D.; Venable, R. M.; Pastor, R. W. An Ab Initio Study on the Torsional Surface of Alkanes and Its Effect on Molecular Simulations of Alkanes and a DPPC Bilayer. *J. Phys. Chem. B* **2005**, *109*, 5300–5311.
- (110) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (111) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, No. 14101.
- (112) Berendsen, H. J. C.; van Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A. R. H. J.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (113) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (114) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (115) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange-Correlation Functional Using the Coulomb-attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (116) Olsen, M. H. Development of Quantum Chemical Methods towards Rationalization and Optimal Design of Photoactive Proteins, Ph.D. thesis; University of Southern Denmark: Odense, Denmark, 2012.
- (117) Almlöf, J.; Taylor, P. R. General Contraction of Gaussian Basis Sets. I. Atomic Natural Orbitals for First- and Second-Row Atoms. *J. Chem. Phys.* **1987**, *86*, 4070–4077.
- (118) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (119) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (120) Olsen, J. M. H.; List, N. H.; Kristensen, K.; Kongsted, J. Accuracy of Protein Embedding Potentials: An Analysis in Terms of Electrostatic Potentials. *J. Chem. Theory Comput.* **2015**, *11*, 1832–1842.
- (121) Pou-Américo, R.; Merchán, M.; Nebot-Gil, I.; Widmark, P. O.; Roos, B. O. Density Matrix Averaged Atomic Natural Orbital (ANO) Basis Sets for Correlated Molecular Wave Functions - III. First Row Transition Metal Atoms. *Theor. Chim. Acta* **1995**, *92*, 149–181.
- (122) Khah, A. M.; Reinholdt, P.; Olsen, J. M. H.; Kongsted, J.; Hättig, C. Avoiding Electron Spill-out in QM/MM Calculations on Excited States with Simple Pseudopotentials. *J. Chem. Theory Comput.* **2020**, *16*, 1373–1381.
- (123) Franchini, M.; Philipsen, P. H. T.; Visscher, L. The Becke Fuzzy Cells Integration Scheme in the Amsterdam Density Functional Program Suite. *J. Comput. Chem.* **2013**, *34*, 1819–1827.
- (124) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (125) Perdew, J. *Electronic Structure of Solids 91*, Edited by Ziesche, P. and Eschrig, H. (Berlin: Akademie-verlag) p. 11; Perdew, J. P. and Wang, Y., 1992. *Phys. Rev. B* **1991**, *45*, No. 244.
- (126) Lembarki, A.; Chermette, H. Obtaining a Gradient-Corrected Kinetic-Energy Functional from the Perdew-Wang Exchange Functional. *Phys. Rev. A* **1994**, *50*, 5328–5331.
- (127) Kiewisch, K.; Eickerling, G.; Reiher, M.; Neugebauer, J. Topological Analysis of Electron Densities from Kohn-Sham and Subsystem Density Functional Theory. *J. Chem. Phys.* **2008**, *128*, No. 44114.

(128) Hedegård, E. D. Assessment of Oscillator Strengths with Multiconfigurational Short-Range Density Functional Theory for Electronic Excitations in Organic Molecules. *Mol. Phys.* **2017**, *115*, 26–38.

(129) Stefanovich, E. V.; Truong, T. N. Embedded Density Functional Approach for Calculations of Adsorption on Ionic Crystals. *J. Chem. Phys.* **1996**, *104*, 2946–2955.

(130) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian Electrostatic Coupling Scheme for Hybrid Car-Parrinello Molecular Dynamics Simulations. *J. Chem. Phys.* **2002**, *116*, 6941–6947.

(131) Gao, B.; Thorvaldsen, A. J.; Ruud, K. Gen1Int: A Unified Procedure for the Evaluation of One-Electron Integrals over Gaussian Basis Functions and Their Geometric Derivatives. *Int. J. Quantum Chem.* **2011**, *111*, 858–872.

## Recommended by ACS

### Constrained Nuclear-Electronic Orbital Density Functional Theory with a Dielectric Continuum Solvent Model

Xi Xu.

JULY 20, 2023  
THE JOURNAL OF PHYSICAL CHEMISTRY A

READ [↗](#)

### Predict the Polarizability and Order of Magnitude of Second Hyperpolarizability of Molecules by Machine Learning

Guoxiang Zhao, Jian Zhang, *et al.*

JULY 14, 2023  
THE JOURNAL OF PHYSICAL CHEMISTRY A

READ [↗](#)

### Embedding Beyond Electrostatics: The Extended Polarizable Density Embedding Model

Willem Van den Heuvel, Jacob Kongsted, *et al.*

APRIL 01, 2023  
THE JOURNAL OF PHYSICAL CHEMISTRY B

READ [↗](#)

### Development, Validation, and Pilot Application of a Generalized Fluctuating Charge Model for Computational Spectroscopy in Solution

Vincenzo Barone, Marco Pagliani, *et al.*

APRIL 06, 2022  
ACS OMEGA

READ [↗](#)

Get More Suggestions >

# B | Global Properties

## B.1. MOF Guest Placer Script

In order to introduce guest molecules into the ZIFs, the following script was developed enabling the introduction of molecules by center of mass. The MOF Guest Placer is called *via* the main.py and employs the different classes: VectorRotation, Molecule, Coordinate, Plane. The code for all of these is shown below (listings B.1 to B.14). <sup>||</sup>

**Listing B.1:** MOF Guest Placer main.py

---

```
#!/usr/bin/env python3
from flags import *
from Molecule import *
from Plane import *
from sympy.geometry import *
from Coordinate import *

if args.debug > 0:
    print(args)

# determine input files
supermol_file = None
if args.input and len(args.input) >= 2:
    input_files = args.input
    plane_file = input_files[0]
    guest_file = input_files[1]
    try:
        supermol_file = input_files[2]
    except:
        print("No supermol file found.")
        pass
else:
    raise FileNotFoundError('Error reading input files.
                            Looking for: -r Plane Guest Supermol')

if not args.output:
    args.output = 'out'
```

---

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022. Pascal Czember assisted with writing the functions for reading and writing pdb files, overall organizational assistance and fruitful discussions.

**Listing B.2:** MOF Guest Placer main.py continued

---

```
# create molecule instances from input files
plane_mol = Molecule(plane_file)
guest_mol = Molecule(guest_file)
super_mol = None

if supermol_file:
    if args.debug > 0:
        print(supermol_file)
    super_mol = Molecule(supermol_file)

# example for what can be done
if super_mol:
    x, y, z = super_mol.atoms[0].get_coords() # get coords of first atom

#####
## TRANSLATE TO ORIGIN #
#####
# calculate plane from plane_mol
plane_mol.get_atomic_vectors()
plane_plane = Molplane(mol=plane_mol, center=plane_mol.com)
first_t = plane_mol.com # save for later
plane_mol.translate(first_t, invert=True)

# guest mol: make plane
guest_mol.get_atomic_vectors()
guest_plane = Molplane(mol=guest_mol, center=guest_mol.com)
guest_mol.translate(guest_mol.com, invert=True)

if args.debug > 1:
    new_mol = plane_mol + guest_mol
    new_mol.write_xyz("1_first_translation")

#####
# ROTATE AROUND INTERSECTION #
#####
# calculate plane from plane_mol
plane_mol.get_atomic_vectors()
plane_plane = Molplane(mol=plane_mol, center=plane_mol.com)
# guest mol: make plane
guest_mol.get_atomic_vectors()
guest_plane = Molplane(mol=guest_mol, center=guest_mol.com)
common_line = plane_plane.intersection(guest_plane)
common_vector = common_line[0].points[1]-common_line[0].points[0] # get origin vec
```

---

**Listing B.3:** MOF Guest Placer main.py continued

---

```

rot_vec = Coordinate([[float(common_vector.x), float(common_vector.y),
                                                                float(common_vector.z)]])

phi = float(guest_plane.angle_between(plane_plane))
if args.invert_rot:
    phi = -1 * phi
guest_mol.rotates(rot_vec, phi=phi)
if args.debug > 1:
    new_mol = plane_mol + guest_mol
    new_mol.write_xyz("2_rotation")
#####
# ROTATE IN PLANE #
#####
# rotation of guest around com to match planemol
guest_mol.get_atomic_vectors()
guest_plane = Molplane(mol=guest_mol, center=guest_mol.com)
nv = Coordinate([[float(guest_plane.normal_vector[0]),
                    float(guest_plane.normal_vector[1]), float(guest_plane.normal_vector[2])]])
nv_norm = nv/math.sqrt(nv.x**2 + nv.y**2+ nv.z**2)
nv_norm_float = Coordinate([[float(nv_norm[0]), float(nv_norm[1]), float(nv_norm[2])]])
a = plane_mol.atoms[0].coords
b = guest_mol.atoms[0].coords
c = float(a.dot(b)/(a.len*b.len)) # angle between first atom of guest and plane
phi = np.arccos(c)
if args.invert_2_rot:
    phi = -1 * phi
guest_mol.rotates(nv_norm_float, phi=-1*phi)
if args.debug > 1:
    new_mol = plane_mol + guest_mol
    new_mol.write_xyz("3_planerotatation")
#####
# ROTATE IN PLANE (USER) #
#####
# rotation of plane around com
guest_mol.get_atomic_vectors()
guest_plane = Molplane(mol=guest_mol, center=guest_mol.com)
nv = Coordinate([[float(guest_plane.normal_vector[0]),
                    float(guest_plane.normal_vector[1]), float(guest_plane.normal_vector[2])]])
nv_norm = nv/math.sqrt(nv.x**2 + nv.y**2+ nv.z**2)
nv_norm_float = Coordinate([[float(nv_norm[0]), float(nv_norm[1]), float(nv_norm[2])]])
phi = np.radians(args.angle)
guest_mol.rotates(nv_norm_float, phi=phi)
if args.debug > 1:
    new_mol = plane_mol + guest_mol
    new_mol.write_xyz("4_userrotation")

```

---

**Listing B.4:** MOF Guest Placer main.py continued

---

```
#####  
## TRANSLATE OFFSET #  
#####  
# offset  
if args.invert:  
    offset = -1*args.offset  
else:  
    offset = args.offset  
  
# shift by offset vector (normalvector)  
nv = Coordinate([[float(guest_plane.normal_vector[0]),  
                float(guest_plane.normal_vector[1]), float(guest_plane.normal_vector[2])]])  
nv_norm = offset * nv/math.sqrt(nv.x**2 + nv.y**2+ nv.z**2)  
nv_norm_float = Coordinate([[float(nv_norm[0]), float(nv_norm[1]), float(nv_norm[2])]])  
guest_mol.translate(nv_norm_float)  
  
if args.debug > 1:  
    new_mol = plane_mol + guest_mol  
    new_mol.write_xyz("5_offset_translation")  
  
#####  
# TRANSLATE BACK #  
#####  
plane_mol.translate(first_t)  
guest_mol.translate(first_t)  
if args.debug > 1:  
    new_mol = plane_mol + guest_mol  
    new_mol.write_xyz("6_back_translation")  
  
if super_mol:  
    final_mol = super_mol + guest_mol  
else:  
    final_mol = plane_mol + guest_mol  
  
if args.outputformat == "pdb":  
    final_mol.write_pdb(args.output)  
else:  
    final_mol.write_xyz(args.output)
```

---

**Listing B.5:** MOF Guest Placer VectorRotation.py

---

```

import numpy as np
import math
from Coordinate import *

def rot_matrix(u, phi=None, phi_deg=None):
    # u: rotation um diesen vektor
    # rotation gegen uhrzeigersinn um winkel in radians
    if phi_deg:
        phi = np.radians(phi_deg)
    u_norm = u/math.sqrt(u.x**2 + u.y**2+ u.z**2)

    a = u_norm[0,0]
    b = u_norm[0,1]
    c = u_norm[0,2]
    sin = np.sin(phi)
    cos = np.cos(phi)
    m = np.array(
        [[a**2*(1-cos)+cos , a*b*(1-cos)-c*sin , a*c*(1-cos)+b*sin ],
         [a*b*(1-cos)+c*sin, b**2 *(1-cos)+cos , b*c*(1-cos) - a*sin],
         [a*c*(1-cos)-b*sin, b*c*(1-cos)+a*sin , c**2 * (1-cos)+cos ]]
    )
    return m

def rotate(v, u, phi=None, phi_deg=None):
    # v: vector to rotate, u: rotation axis, phi: angle to rotate in radians,
    # phi_deg: angle to rotate in degrees
    if args.debug > 1:
        print(v)
    if u.rows == 1 and u.cols == 3:
        pass
    else:
        raise TypeError("u needs to be a np array of form 1x3")
    if phi:
        m = rot_matrix(u, phi=phi)
    elif phi_deg:
        m = rot_matrix(u, phi_deg=phi_deg)
    else:
        raise Exception("Define angle for rotation: phi in radians or
                        phi_deg in degrees.")

```

---

**Listing B.6:** MOF Guest Placer Plane.py continued

---

```
if v.rows == 1 and v.cols == 3:
    x = v.T
elif v.rows == 3 and v.cols == 1:
    x = v
else:
    raise TypeError("v needs to be a np array of form 1x3 or 3x1")
if args.debug > 2:
    print("m: ", m, "\n v", x)
product = m.dot(x)
out_vec = Coordinate([[product[0,0], product[1,0], product[2,0]]])
if args.debug > 2:
    print(out_vec)
return out_vec
```

---

**Listing B.7:** MOF Guest Placer Plane.py

---

```
from flags import *
from sympy import Plane
from sympy import Point3D

class Molplane(Plane):
    def __new__(cls, p1=None, p2=None, p3=None, normal_vector=None,
               mol=None, center=None):

        if mol and center:
            atom_list = mol.sorted_atom_list_by_com
            p1 = atom_list[0].coords
            p2 = atom_list[1].coords-atom_list[0].coords
            p3 = atom_list[2].coords-atom_list[0].coords
            normal_vector = p2.cross(p3)
            if args.debug > 2:
                print("P1: ", p1, "\nNormal vector: ", normal_vector)
            return super().__new__(cls, p1=p1, normal_vector=normal_vector)
        elif p1 and p2 and p3:
            normal_vector = p2.cross(p3)
            return super().__new__(cls, p1=p1, normal_vector=normal_vector)
        elif p1 and normal_vector:
            return super().__new__(cls, p1=p1, normal_vector=normal_vector)
        else:
            raise Exception('Neither p1 and normal_vector nor mol and center defined.
                             No plane could be created.')
```

---



**Listing B.8:** MOF Guest Placer Molecule.py

```
from flags import *
from collections import namedtuple
import sympy as s
from VectorRotation import *
from Coordinate import *

class Molecule:
    def __init__(self, filename=None):
        self.com = None
        self.center = None
        self.atoms = None
        self.number_of_atoms = None
        self.comment = None
        self.sorted_atom_list_by_com = None

        if filename is not None:
            self.read(filename)

    def __add__(self, second_mol):
        added_mol = Molecule()
        added_mol.number_of_atoms = self.number_of_atoms + second_mol.number_of_atoms
        added_mol.atoms = self.atoms + second_mol.atoms

        return added_mol

    def translate(self, t, invert=False): # translation vector t
        for atom in self.atoms:
            if invert:
                new_coords = atom.coords - t
            else:
                new_coords = atom.coords + t
            atom.coords = new_coords

    def rotates(self, u, phi=None, phi_deg=None):
        atoms = []
        for atom in self.atoms:
            if phi:
                new_coord = rotate(atom.coords, u=u, phi=phi)
            elif phi_deg:
                new_coord = rotate(atom.coords, u=u, phi_deg=phi_deg)
            else:
                new_coord = atom.coords
            atoms.append(Atom(element=atom.element, coords=new_coord))
        self.atoms = atoms
```

**Listing B.9:** MOF Guest Placer Molecule.py continued

---

```
def read(self, filename):
    try:
        f = open(filename, 'r')
    except OSError:
        raise FileNotFoundError('Error reading file: {}'.format(filename))

    if filename.endswith('.xyz'):
        self.readxyz(f)
    elif filename.endswith('.pdb'):
        self.readpdb(f)
    else:
        raise TypeError("Filename should be of type .xyz or .pdb
                        but is not for: ", filename)

def readxyz(self, f):
    if args.debug > 0:
        print("read xyz")
    lines = f.readlines()
    self.number_of_atoms = int(lines[0])
    self.comment = lines[1]
    if args.debug > 2:
        print(self.number_of_atoms)
    atoms = []
    for line in lines[2:]:
        data = line.split()
        atoms.append(Atom(number_of_atoms=len(lines[2:]),
                        element=data[0].lower().title(),
                        coords=Coordinate([[float(data[1]), float(data[2]),
                        float(data[3])]]), unit='angstrom'))
    self.atoms = atoms

def readpdb(self, f):
    if args.debug > 0:
        print("read pdb")
    lines = f.readlines()
    self.number_of_atoms = int(lines[-2].split()[1])
    if args.debug > 0:
        print(self.number_of_atoms)
    atoms = []
    for line in lines[1:-1]:
        data = line.split()
        atoms.append(Atom(number_of_atoms=len(lines[-2:1]),
                        element=data[2].lower().title(),
                        coords=Coordinate([[float(data[5]), float(data[6]),
                        float(data[7])]]), unit='angstrom'))
    self.atoms = atoms
```

---

**Listing B.10:** MOF Guest Placer Molecule.py continued

```
def get_com(self):
    import periodictable as pt
    if args.debug > 1:
        print("calc com")
    xm, ym, zm = 0.0, 0.0, 0.0
    M = 0.0
    # get mass of all elements
    elements = {}
    for el in pt.elements:
        elements[el.symbol] = el.mass

    for atom in self.atoms:
        m = elements[atom.element]
        if args.debug > 2:
            print(atom.element, m)
        xm += atom.coords.x * m
        ym += atom.coords.y * m
        zm += atom.coords.z * m
        M += m
        if args.debug > 2:
            print(xm, ym, zm, M)
    com_x = xm / M
    com_y = ym / M
    com_z = zm / M
    com = Coordinate([[com_x, com_y, com_z]])
    self.com = com
    if args.debug > 2:
        print(com_x, com_y, com_z)

    return com

def get_center(self):
    x_tot, y_tot, z_tot = 0.0, 0.0, 0.0
    for atom in self.atoms:
        x_tot += atom.coords.x
        y_tot += atom.coords.y
        z_tot += atom.coords.z
        if args.debug > 2:
            print(x_tot, y_tot, z_tot)
    c_x = x_tot / self.number_of_atoms
    c_y = y_tot / self.number_of_atoms
    c_z = z_tot / self.number_of_atoms
```

**Listing B.11:** MOF Guest Placer Molecule.py continued

---

```
c = Coordinate(c_x, c_y, c_z)
self.center = c
if args.debug > 2:
    print(c_x, c_y, c_z)
return c

def write_xyz(self, fileprefix):
    filename = fileprefix + '.xyz'
    lines = str(self.number_of_atoms) + '\n'
    if self.comment:
        lines += str(self.comment)
    lines += '\n'
    for atom in self.atoms:
        lines += '{:3}\t{:11.8f}\t{:11.8f}\t{:11.8f}\n'.format(atom.element,
                                                             atom.coords.x, atom.coords.y, atom.coords.z)
    with open(filename, 'w') as f:
        f.write(lines)

def write_pdb(self, fileprefix):
    filename = fileprefix + '.pdb'
    lines = 'CRYST1    0.000    0.000    0.000  90.00  90.00  90.00 P 1          1'
    lines += '\n'
    atomcount = 0
    for atom in self.atoms:
        atomcount += 1
        lines += '{atomcount:>7}{element:>3}{x:>8}{v:>4}{xcoord:>12}{ycoord:>8}
                {zcoord:>8}{p:>6}{p:>6}{element:>12}\n'.format("ATOM",
                                                            atomcount=atomcount, element=atom.element, x="X", v=1,
                                                            xcoord=f'{atom.coords.x:.3f}', ycoord=f'{atom.coords.y:.3f}',
                                                            zcoord=f'{atom.coords.z:.3f}', p=f'{0.00:.2f}')
    with open(filename, 'w') as f:
        f.write(lines)

def get_atomic_vectors(self):
    # defining the Atoms in the guest molecule as vectors
    # from the Center of Mass
    distance_dict = {}
    sorted_atom_list_by_com = []
    com = self.get_com()
    for atom in self.atoms:
        element = atom.element
        atom.get_com_vec(com)
        distance_dict[atom.dist_from_com] = atom
    # create dictionary entry for every distance
```

---

**Listing B.12:** MOF Guest Placer Molecule.py continued

---

```

        sorted_dist_dict = dict(sorted(distance_dict.items()))
        for key, value in sorted_dist_dict.items():
            sorted_atom_list_by_com.append(value)
        self.sorted_atom_list_by_com = sorted_atom_list_by_com
    def get_sorted_atom_list(self):
        self.get_atomic_vectors(self)

        return self.sorted_atom_list_by_com

class Atom:

    def __init__(self, number_of_atoms=None, element=None, coords=None, unit='angstrom'):
        self.number_of_atoms = number_of_atoms
        self.coords = coords
        self.unit = unit
        self.element = element
        self.vec_from_com = None
        self.dist_from_com = None

    def get_coords(self):
        return self.coords.x, self.coords.y, self.coords.z

    def get_com_vec(self, com):
        vec = self.coords - com
        com_distance = (vec.x**2 + vec.y**2 + vec.z**2)**0.5
        # magnitude of the vector
        self.vec_from_com = vec # gives the vector attribute to the Atom
        self.dist_from_com = com_distance
        # gives the distance from COM attribute to the Atom

```

---

**Listing B.13:** MOF Guest Placer Coordinate.py

---

```

from flags import *
import sympy as s
import math
class Coordinate(s.Matrix):
    __slots__ = ()

    def __init__(self, coords):
        super().__init__()
        self.x = coords[0][0]
        self.y = coords[0][1]
        self.z = coords[0][2]
        self.len = self.get_length()

```

---

**Listing B.14:** MOF Guest Placer Coordinate.py continued

---

```
def __sub__(self, second_coord):
    # Enable subtracting coordinate class Objects
    if type(second_coord) != Coordinate:
        raise TypeError("Not of type Coordinate.")
    new_x = self.x - second_coord.x
    new_y = self.y - second_coord.y
    new_z = self.z - second_coord.z
    subtracted_coord = Coordinate([[new_x, new_y, new_z]])

    return subtracted_coord

def __add__(self, second_coord): # Enable adding coordinate class Objects
    if type(second_coord) != Coordinate:
        raise TypeError("Not of type Coordinate.")
    new_x = self.x + second_coord.x
    new_y = self.y + second_coord.y
    new_z = self.z + second_coord.z
    added_coord = Coordinate([[new_x, new_y, new_z]])

    return added_coord

def get_length(self): # Length of vector
    len = math.sqrt(self.x**2 + self.y**2+ self.z**2)
    return len

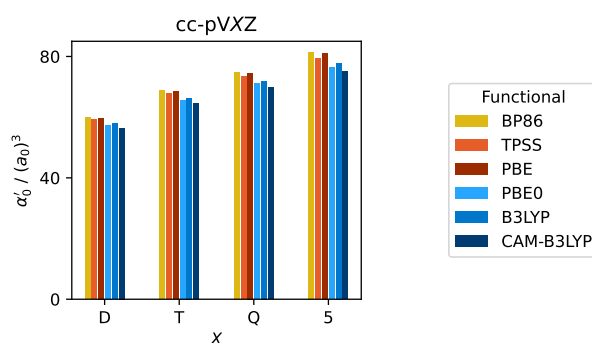
def normalize(self):
    norm = self/math.sqrt(u.x**2 + u.y**2+ u.z**2)
    self.x = norm[0,0]
    self.y = norm[0,1]
    self.z = norm[0,2]
    return norm
```

---

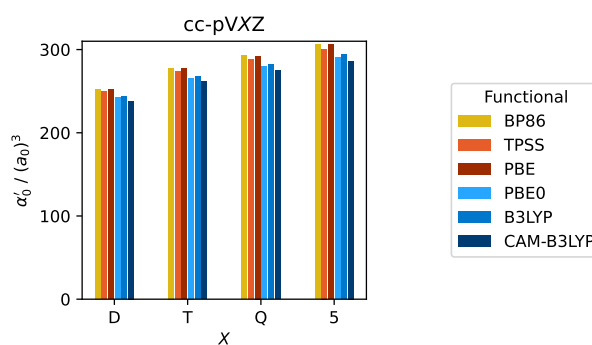
## B.2. Results

### B.2.1. DFT Functional Evaluation

In this section the static polarizability volumes that were calculated with CP-SCF in ORCA for the  $\text{Zn}^{2+}$ , IM, TET and DI fragments with different XC functionals and basis sets in double- $(\zeta = 2)$ , triple- $(\zeta = 3)$ , quadruple- $(\zeta = 4)$  and quintuple- $(\zeta = 5)$  quality are presented.

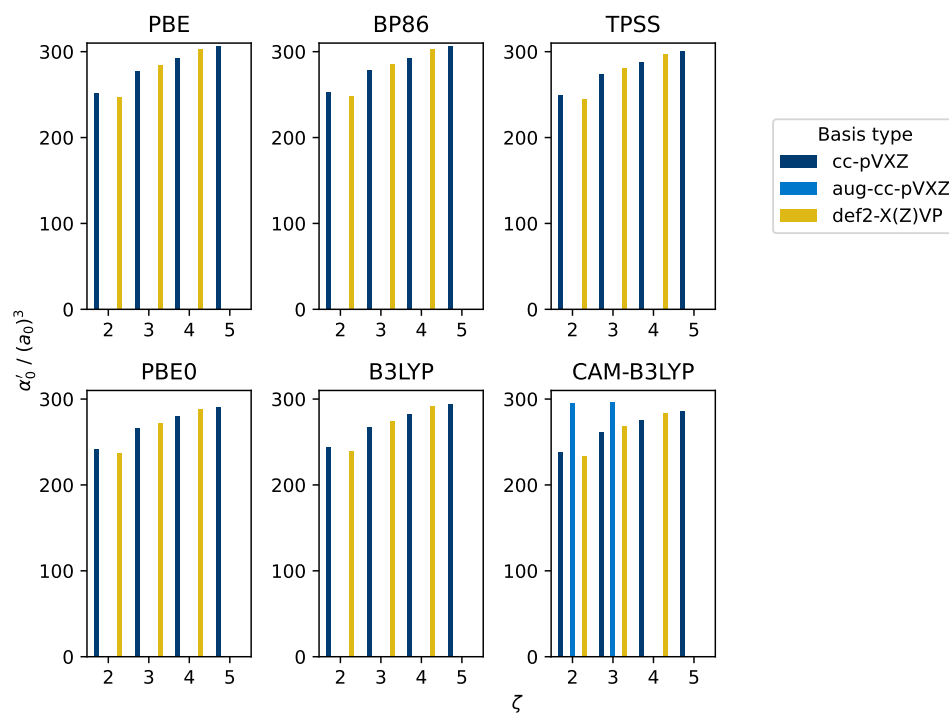


**Figure B.1.:** Static polarizability volume  $\alpha'_0$  calculated with CP-SCF, different XC functionals and different basis sets in double- $(\zeta = 2)$ , triple- $(\zeta = 3)$ , quadruple- $(\zeta = 4)$  and quintuple- $(\zeta = 5)$  quality for the IM fragment extracted from the experimental ZIF-8 SC XRD structure (model **A**).<sup>[366]</sup> Results for GGA-type functionals (BP86, TPSS and PBE) are presented in yellow and red tones, for hybrid functionals (PBE0, B3LYP and CAM-B3LYP) are shown in blue tones.<sup>‡</sup>



**Figure B.2.:** Static polarizability volume  $\alpha'_0$  calculated with CP-SCF, different XC functionals and different basis sets in double- $(\zeta = 2)$ , triple- $(\zeta = 3)$ , quadruple- $(\zeta = 4)$  and quintuple- $(\zeta = 5)$  quality for the TET fragment extracted from the experimental ZIF-8 SC XRD structure (model **A**).<sup>[366]</sup> Results for GGA-type functionals (BP86, TPSS and PBE) are presented in yellow and red tones, for hybrid functionals (PBE0, B3LYP and CAM-B3LYP) are shown in blue tones.<sup>‡</sup>

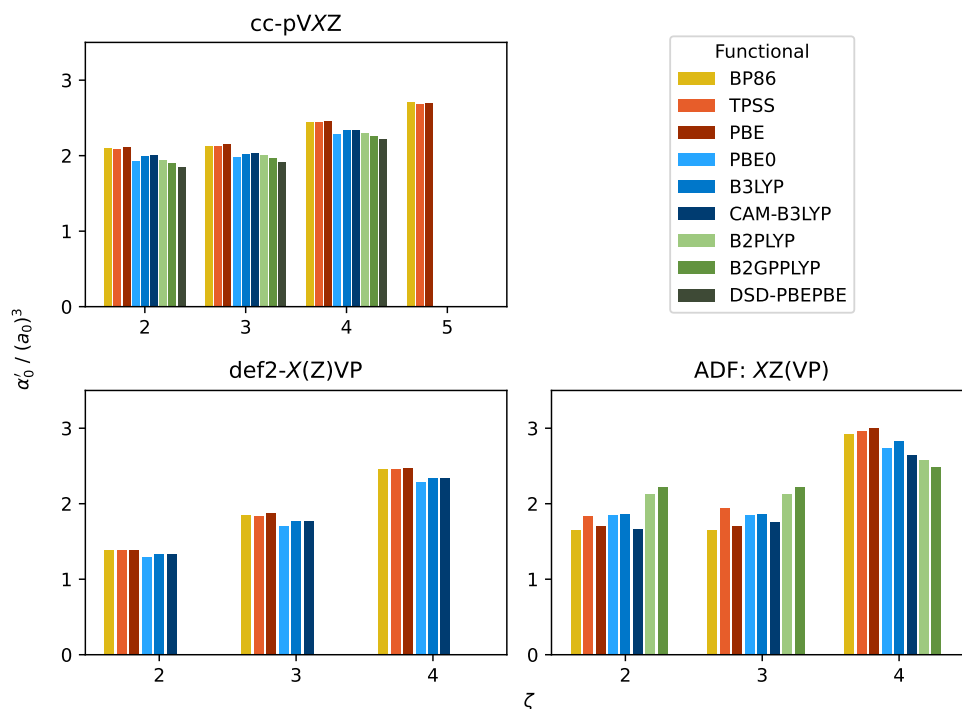
<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.



**Figure B.3.:** Static polarizability volume  $\alpha'_0$  calculated with CP-SCF, different XC functionals and different basis sets in double- ( $\zeta = 2$ ), triple- ( $\zeta = 3$ ), quadruple- ( $\zeta = 4$ ) and quintuple- ( $\zeta = 5$ ) quality for the TET fragment extracted from the experimental ZIF-8 SC XRD structure (model **A**)<sup>[366]‡</sup>

<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.





**Figure B.4.:** Static polarizability volume  $\alpha'_0$  calculated with CP-SCF, different XC functionals and different basis sets in double-, triple-, quadruple- and quintuple- $\zeta$  quality for the  $\text{Zn}^{2+}$  fragment. Results for GGA-type functionals (BP86, TPSS and PBE) are presented in yellow and red tones, for hybrid functionals (PBE0, B3LYP and CAM-B3LYP) are shown in blue tones and double-hybrid functionals (B2PLYP, B2GPPLYP and DSD-PBEPBE-D3BJ) are displayed in green tones.<sup>‡</sup>

<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.

### B.2.2. Program Comparison

In this section the static polarizabilities calculated with an CP-SCF ansatz and a TD-DFT ansatz in ORCA and TURBOMOLE, respectively, are presented for the different fragments and obtained with different Dunning- and Ahlrichs-type basis sets.

**Table B.1.:** Static polarizabilities in  $a_0^3$  calculated with ORCA and TURBOMOLE (TM) with CAM-B3LYP and different basis sets with and without diffuse functions for  $\text{Zn}^{2+}$ . x denotes erroneous calculations.<sup>‡</sup>

Basis set	ORCA	TM	Basis set	ORCA	TM
cc-pVDZ	2.00	-	def2-SVP	1.33	-
cc-pVTZ	2.04	2.04	def2-TZVP	1.77	1.77
cc-pVQZ	2.33	2.33	def2-QZVP	2.34	2.34
cc-pV5Z	x	2.55	-	-	-
aug-cc-pVDZ	2.59	-	def2-SVPD	-	-
aug-cc-pVTZ	2.57	2.57	def2-TZVPD	1.77	1.77
aug-cc-pVQZ	2.73	2.73	def2-QZVPD	2.34	2.34
aug-cc-pV5Z	x	2.75	-	-	-

**Table B.2.:** Static polarizabilities in  $a_0^3$  calculated with ORCA and TURBOMOLE with CAM-B3LYP and different basis sets for TET.<sup>‡||</sup>

Basis set	ORCA	TM	Basis set	ORCA	TM
cc-pVDZ	228.14	-	def2-SVP	233.11	-
cc-pVTZ	251.17	251.16	def2-TZVP	267.87	257.63
cc-pVQZ	264.25	264.22	def2-QZVP	283.40	272.41
cc-pV5Z	275.00	-	-	-	-
aug-cc-pVDZ	283.80	-	def2-SVPD	-	-
aug-cc-pVTZ	284.18	284.09	def2-TZVPD	-	278.20
aug-cc-pVQZ	-	-	def2-QZVPD	-	281.16

<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.

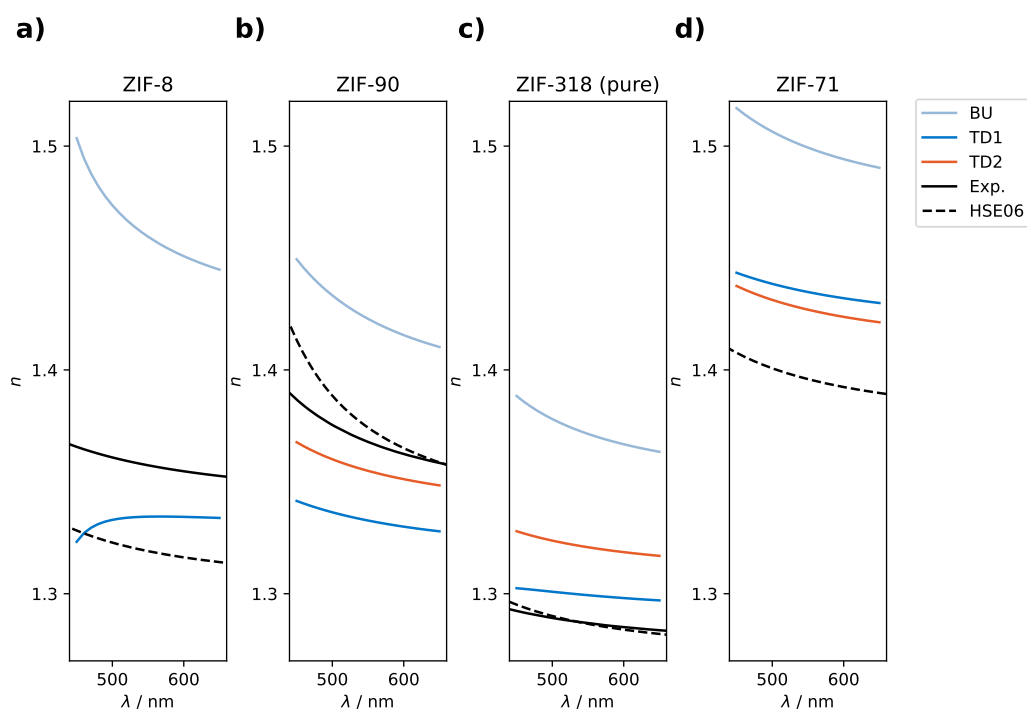
**Table B.3.:** Static polarizabilities in  $a_0^3$  calculated with ORCA and TURBOMOLE with CAM-B3LYP and different basis sets for DI.<sup>‡</sup><sup>||</sup>

Basis set	ORCA	TM	Basis set	ORCA	TM
cc-pVDZ	408.69	-	def2-SVP	-	-
cc-pVTZ	446.82	446.83	def2-TZVP	-	456.98
cc-pVQZ	467.77	-	def2-QZVP	-	-
cc-pV5Z	484.25	-	-	-	-
aug-cc-pVDZ	-	-	def2-SVPD	-	-
aug-cc-pVTZ	-	497.53	def2-TZVPD	-	488.88

<sup>‡</sup> Onno Strolka, Leibniz University Hannover, 2022.<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.

### B.2.3. Refractive Indices for Empty Zeolitic Imidazolate Frameworks

#### Model A



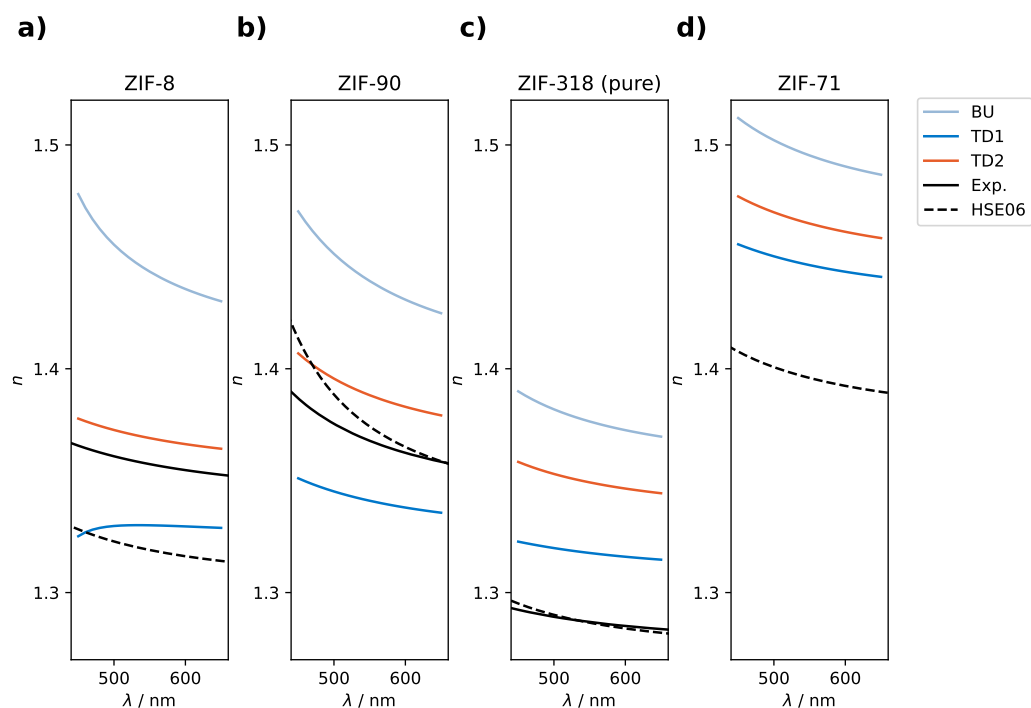
**Figure B.5.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental SC XRD structures. (model **A**)<sup>[366]</sup> Results are obtained for **a)** ZIF-8, **b)** ZIF-90, **c)** ZIF-318 with an extrapolated volume, **d)** ZIF-71. The results are compared to experimental and computational reference.<sup>[35, 129, 363]</sup>§\*

|| Pascal Czember, Leibniz University Hannover, 2022.

§ Daniel Bömke, Leibniz University Hannover, 2022.

\* Erona Shabani, Leibniz University Hannover, 2023.

## Model B

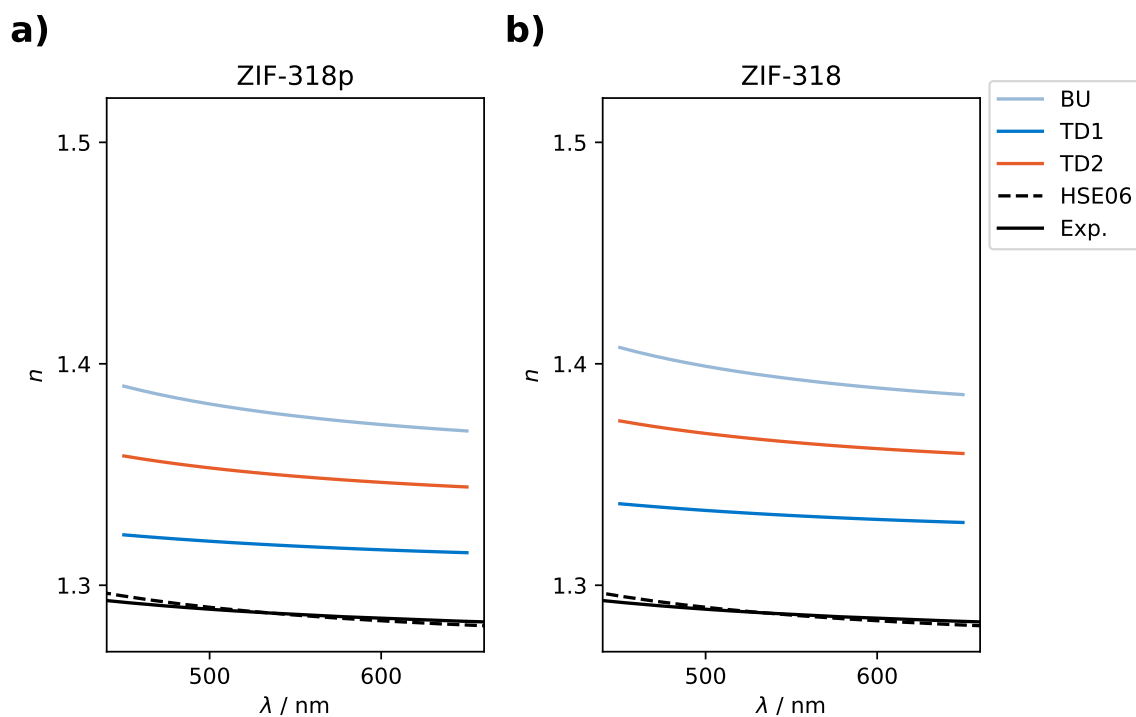


**Figure B.6.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental SC XRD structures and optimized with B97-3c/def2-mTZVP (model **B**).<sup>[366]</sup> Results are obtained for **a)** ZIF-8, **b)** ZIF-90, **c)** ZIF-318 with an extrapolated volume, **d)** ZIF-71. The results are compared to experimental and computational reference.<sup>[35, 129, 363]</sup>§♣

|| Pascal Czember, Leibniz University Hannover, 2022.

§ Daniel Bömke, Leibniz University Hannover, 2022.

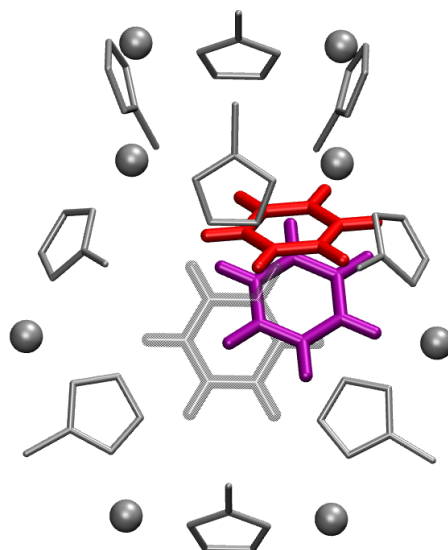
♣ Erona Shabani, Leibniz University Hannover, 2023.



**Figure B.7.:** Dispersion of refractive indices  $n$  calculated with TD-DFT in a CAM-B3LYP/aug-cc-pVTZ scheme for the fragments extracted from the experimental ZIF-318 SC XRD structure and optimized fragments with B97-3c/def2-mTZVP. (model **B**)<sup>[366]</sup> Results are obtained for ZIF-318 with different volumes: **a)** the extrapolated and **b)** the experimental volume (*cf.* Table 5.1.1). The results are compared to experimental and computational reference.<sup>[129, 363]\*</sup>

\* Erona Shabani, Leibniz University Hannover, 2023.

### B.2.4. Loading of Zeolitic Imidazolate Frameworks



**Figure B.8.:** Progression of a benzene guest molecule placed below a 4MR of the ZIF-8 pore. Red and purple molecules show the progression during the optimization, while the black molecule shows the ideal 6MR position.<sup>||</sup>

<sup>||</sup> Pascal Czember, Leibniz University Hannover, 2022.





# Bibliography

- (1) Kurtz, H. A.; Dudis, D. S., Quantum Mechanical Methods for Predicting Nonlinear Optical Properties In *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd: 1998, pp 241–279.
- (2) Lima, A. F.; Lalic, M. V. The usefulness of the first-principles calculations of optical properties of the materials and the type of information that can be accessed by them. *Optical Materials: X* **2022**, *15*, 100185.
- (3) Planck, M., Das Wesen des Lichts In *Das Wesen des Lichts: Vortrag Gehalten in der Hauptversammlung der Kaiser-Wilhelm-Gesellschaft am 28. Oktober 1919*, Planck, M., Ed.; Springer: Berlin, Heidelberg, 1920, pp 3–22.
- (4) Loudon, R., *The Quantum Theory of Light*; OUP Oxford: 2000; 454 pp.
- (5) Moran, A. M.; Kelley, A. M. Solvent effects on ground and excited electronic state structures of p-nitroaniline. *J. Chem. Phys.* **2001**, *115*, 912–924.
- (6) Sok, S.; Willow, S. Y.; Zahariev, F.; Gordon, M. S. Solvent-Induced Shift of the Lowest Singlet  $\pi \rightarrow \pi^*$  Charge-Transfer Excited State of p-Nitroaniline in Water: An Application of the TDDFT/EFP1 Method. *J. Phys. Chem. A* **2011**, *115*, 9801–9809.
- (7) Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. Chemical nature of the light emitter of the Aequorea green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13617–13622.
- (8) Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. Calculating Absorption Shifts for Retinal Proteins: Computational Challenges. *J. Phys. Chem. B* **2005**, *109*, 3606–3615.
- (9) Sigurdson, C. J.; Nilsson, K. P. R.; Hornemann, S.; Manco, G.; Polymenidou, M.; Schwarz, P.; Leclerc, M.; Hammarström, P.; Wüthrich, K.; Aguzzi, A. Prion strain discrimination using luminescent conjugated polymers. *Nat. Methods* **2007**, *4*, 1023–1030.
- (10) Altun, A.; Yokoyama, S.; Morokuma, K. Spectral Tuning in Visual Pigments: An ONIOM(QM:MM) Study on Bovine Rhodopsin and its Mutants. *J. Phys. Chem. B* **2008**, *112*, 6814–6827.
- (11) Magnusson, K.; Simon, R.; Sjölander, D.; Sigurdson, C. J.; Hammarström, P.; Nilsson, K. P. R. Multimodal fluorescence microscopy of prion strain specific PrP deposits stained by thiophene-based amyloid ligands. *Prion* **2014**, *8*, 319–329.

- (12) Morzan, U. N.; Alonso De Armiño, D. J.; Foglia, N. O.; Ramírez, F.; González Lebrero, M. C.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in Complex Environments from QM-MM Simulations. *Chem. Rev.* **2018**, *118*, 4071–4113.
- (13) Åslund, A.; Sigurdson, C. J.; Klingstedt, T.; Grathwohl, S.; Bolmont, T.; Dickstein, D. L.; Glimsdal, E.; Prokop, S.; Lindgren, M.; Konradsson, P.; Holtzman, D. M.; Hof, P. R.; Heppner, F. L.; Gandy, S.; Jucker, M.; Aguzzi, A.; Hammarström, P.; Nilsson, K. P. R. Novel pentameric thiophene derivatives for in vitro and in vivo optical imaging of a plethora of protein aggregates in cerebral amyloidoses. *ACS Chem. Biol.* **2009**, *4*, 673–684.
- (14) Klingstedt, T.; Shirani, H.; Mahler, J.; Wegenast-Braun, B. M.; Nyström, S.; Goedert, M.; Jucker, M.; Nilsson, K. P. R. Distinct spacing between anionic groups: An essential chemical determinant for achieving thiophene-based ligands to distinguish  $\beta$ -Amyloid or tau polymorphic aggregates. *Chem. Eur. J.* **2015**, *21*, 9072.
- (15) Sjöqvist, J.; Linares, M.; Lindgren, M.; Norman, P. Molecular dynamics effects on luminescence properties of oligothiophene derivatives: A molecular Mechanics-Response theory study based on the CHARMM force field and density functional theory. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17532–17542.
- (16) Sjöqvist, J.; Linares, M.; Mikkelsen, K. V.; Norman, P. QM/MM-MD simulations of conjugated polyelectrolytes: A study of luminescent conjugated oligothiophenes for use as biophysical probes. *J. Phys. Chem. A* **2014**, *118*, 3419–3428.
- (17) König, C.; Skånberg, R.; Hotz, I.; Ynnerman, A.; Norman, P.; Linares, M. Binding sites for luminescent amyloid biomarkers from non-biased molecular dynamics simulations. *Chem. Commun.* **2018**, *54*, 3030–3033.
- (18) Beerepoot, M. T. P.; Steindal, A. H.; Ruud, K.; Olsen, J. M. H.; Kongsted, J. Convergence of environment polarization effects in multiscale modeling of excitation energies. *Comput. Theor. Chem.* **2014**, *1040-1041*, 304–311.
- (19) Nørby, M. S.; Vahtras, O.; Norman, P.; Kongsted, J. Assessing frequency-dependent site polarisabilities in linear response polarisable embedding. *Mol. Phys.* **2017**, *115*, 39–47.
- (20) Rowsell, J. L. C.; Yaghi, O. M. Metal-organic frameworks: a new class of porous materials. *Microporous Mesoporous Mater.* **2004**, *73*, 3–14.
- (21) Batten, S. R.; Champness, N. R.; Chen, X.-M.; Garcia-Martinez, J.; Kitagawa, S.; Öhrström, L.; O’Keeffe, M.; Suh, M. P.; Reedijk, J. Terminology of metal-organic frameworks and coordination polymers (IUPAC Recommendations 2013). *Pure Appl. Chem.* **2013**, *85*, 1715–1724.
- (22) Yin, W.; Tao, C.-a.; Wang, F.; Huang, J.; Qu, T.; Wang, J. Tuning optical properties of MOF-based thin films by changing the ligands of MOFs. *Sci. China Mater.* **2018**, *61*, 391–400.

- (23) Syzgantseva, M. A.; Ireland, C. P.; Ebrahim, F. M.; Smit, B.; Syzgantseva, O. A. Metal Substitution as the Method of Modifying Electronic Structure of Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2019**, *141*, 6271–6278.
- (24) Hendrickx, K.; Vanpoucke, D. E. P.; Leus, K.; Lejaeghere, K.; Van Yperen-De Deyne, A.; Van Speybroeck, V.; Van Der Voort, P.; Hemelsoet, K. Understanding Intrinsic Light Absorption Properties of UiO-66 Frameworks: A Combined Theoretical and Experimental Study. *Inorg. Chem.* **2015**, *54*, 10701–10710.
- (25) Moosavi, S. M.; Nandy, A.; Jablonka, K. M.; Ongari, D.; Janet, J. P.; Boyd, P. G.; Lee, Y.; Smit, B.; Kulik, H. J. Understanding the diversity of the metal-organic framework ecosystem. *Nat. Commun.* **2020**, *11*, 4068.
- (26) Shekhah, O.; Liu, J.; Fischer, R. A.; Wöll, C. MOF thin films: existing and future applications. *Chem. Soc. Rev.* **2011**, *40*, 1081–1106.
- (27) Stassen, I.; Burtch, N.; Talin, A.; Falcaro, P.; Allendorf, M.; Ameloot, R. An updated roadmap for the integration of metal-organic frameworks with electronic devices and chemical sensors. *Chem. Soc. Rev.* **2017**, *46*, 3185–3241.
- (28) Olorunyomi, J.; Teng Geh, S.; A. Caruso, R.; M. Doherty, C. Metal-organic frameworks for chemical sensing devices. *Mater. Horiz.* **2021**, *8*, 2387–2419.
- (29) Gu, Z.-G.; Pfriem, A.; Hamsch, S.; Breitwieser, H.; Wohlgemuth, J.; Heinke, L.; Gliemann, H.; Wöll, C. Transparent films of metal-organic frameworks for optical applications. *Microporous Mesoporous Mater.* **2015**, *211*.
- (30) Nguyen, T. N.; Ebrahim, F. M.; Stylianou, K. C. Photoluminescent, upconversion luminescent and nonlinear optical metal-organic frameworks: From fundamental photophysics to potential applications. *Coord. Chem. Rev.* **2018**, *377*, 259–306.
- (31) Zheng, Y.; Sun, F.-Z.; Han, X.; Xu, J.; Bu, X.-H. Recent Progress in 2D Metal-Organic Frameworks for Optical Applications. *Adv. Opt. Mater.* **2020**, *8*, 2000110.
- (32) Hsu, K.-T.; Thanasekaran, P.; Hsu, T.-W.; Su, C.-H.; Chang, B.-C.; Liu, Y.-H.; Hung, C.-H.; Lu, K.-L. A nonlinear optical cadmium(II)-based metal-organic framework with chiral helical chains derived from an achiral bent dicarboxylate ligand. *CrystEngComm* **2021**, *23*, 824–830.
- (33) Li, H.; Zhang, L.; He, H.; Yang, Y.; Cui, Y.; Qian, G. Tunable nonlinear optical responses based on host-guest MOF hybrid materials. *Sci. China Mater.* **2021**, *64*, 698–705.
- (34) Huang, Y.; Tao, C.-a.; Chen, R.; Sheng, L.; Wang, J. Comparison of Fabrication Methods of Metal-Organic Framework Optical Thin Films. *Nanomater.* **2018**, *8*, 676.
- (35) Keppler, N. C.; Hindricks, K. D. J.; Behrens, P. Large refractive index changes in ZIF-8 thin films of optical quality. *RSC Adv.* **2022**, *12*, 5807–5815.
- (36) So, M. C.; Jin, S.; Son, H.-J.; Wiederrecht, G. P.; Farha, O. K.; Hupp, J. T. Layer-by-Layer Fabrication of Oriented Porous Thin Films Based on Porphyrin-Containing Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2013**, *135*, 15698–15701.

- (37) Lu, G.; Farha, O. K.; Kreno, L. E.; Schoenecker, P. M.; Walton, K. S.; Van Duyne, R. P.; Hupp, J. T. Fabrication of Metal-Organic Framework-Containing Silica-Colloidal Crystals for Vapor Sensing. *Adv. Mater.* **2011**, *23*, 4449–4452.
- (38) Yang, W. Direct calculation of electron density in density-functional theory. *Phys. Rev. Lett.* **1991**, *66*, 1438.
- (39) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (40) Singh, U. C.; Kollman, P. A. A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the CH<sub>3</sub>Cl + Cl<sup>-</sup> exchange reaction and gas phase protonation of polyethers. *J. Comp. Chem.* **1986**, *7*, 718–730.
- (41) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.
- (42) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.
- (43) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632–672.
- (44) Raghavachari, K.; Saha, A. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015**, *115*, 5643–5677.
- (45) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (46) Cramer, C. J.; Truhlar, D. G., Continuum Solvation Models: Classical and Quantum Mechanical Implementations In *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd: 1995, pp 1–72.
- (47) Senn, H. M.; Thiel, W. QM/MM studies of enzymes. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.
- (48) Senn, H. M.; Thiel, W., QM/MM Methods for Biological Systems In *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations*, Reiher, M., Ed.; Topics in Current Chemistry; Springer: Berlin, Heidelberg, 2007, pp 173–290.
- (49) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796.
- (50) Boulanger, E.; Harvey, J. N. QM/MM methods for free energies and photochemistry. *Curr. Opin. Struct. Biol.* **2018**, *49*, 72–76.

- (51) Lipparini, F.; Mennucci, B. Hybrid QM/classical models: Methodological advances and new applications. *Chem. Phys. Rev.* **2021**, *2*, 041303.
- (52) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. A tunable QM/MM approach to chemical reactivity, structure and physico-chemical properties prediction. *Theor. Chem. Acc.* **2007**, *118*, 219–240.
- (53) Stoyanov, S. R.; Gusarov, S.; Kuznicki, S. M.; Kovalenko, A. Theoretical Modeling of Zeolite Nanoparticle Surface Acidity for Heavy Oil Upgrading. *J. Phys. Chem. C* **2008**, *112*, 6794–6810.
- (54) Thompson, M. A.; Schenter, G. K. Excited states of the bacteriochlorophyll b dimer of Rhodospseudomonas viridis: A QM/MM study of the photosynthetic reaction center that includes MM polarization. *J. Phys. Chem.* **1995**, *99*, 6374–6386.
- (55) Daday, C.; Curutchet, C.; Sinicropi, A.; Mennucci, B.; Filippi, C. Chromophore-protein coupling beyond nonpolarizable models: Understanding absorption in green fluorescent protein. *J. Chem. Theory Comput.* **2015**, *11*, 4825–4839.
- (56) Bondanza, M.; Nottoli, M.; Cupellini, L.; Lipparini, F.; Mennucci, B. Polarizable embedding QM/MM: The future gold standard for complex (bio)systems? *Phys. Chem. Chem. Phys.* **2020**, *22*, 14433–14448.
- (57) Wesolowski, T. A.; Shedge, S.; Zhou, X. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* **2015**, *115*, 5891–5928.
- (58) Jacob, C. R.; Neugebauer, J. Subsystem density-functional theory. *WIREs Comput. Mol. Sci.* **2014**, *4*, 325–362.
- (59) Krishtal, A.; Sinha, D.; Genova, A.; Pavanello, M. Subsystem density-functional theory as an effective tool for modeling ground and excited states, their dynamics and many-body interactions. *J. Phys. Condens. Matter* **2015**, *27*, 183202.
- (60) Sun, Q.; Chan, G. K. L. Quantum embedding theories. *Acc. Chem. Res.* **2016**, *49*, 2705–2712.
- (61) Jacob, C. PyADF release v1.2, Zenodo, 2023.
- (62) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable Density Embedding: A New QM/QM/MM-Based Computational Strategy. *J. Phys. Chem. A* **2015**, *119*, 5344–5355.
- (63) Harshan, A. K.; Bronson, M. J.; Jensen, L. Local-field effects in linear response properties within a polarizable frozen density embedding method. *J. Chem. Theory Comput.* **2022**, *18*, 380–393.
- (64) Pal, P. P.; Liu, P.; Jensen, L. Polarizable frozen density embedding with external orthogonalization. *J. Chem. Theory Comput.* **2019**, *15*, 6588–6596.
- (65) Daday, C.; König, C.; Valsson, O.; Neugebauer, J.; Filippi, C. State-specific embedding potentials for excitation-energy calculations. *J. Chem. Theory Comput.* **2013**, *9*, 2355–2367.

- (66) Daday, C.; König, C.; Neugebauer, J.; Filippi, C. Wavefunction in density functional theory embedding for excited states: Which wavefunctions, which densities? *ChemPhysChem* **2014**, *15*, 3205–3217.
- (67) Olsen, J.; Jørgensen, P. Linear and nonlinear response functions for an exact state and for an MCSCF state. *J. Chem. Phys.* **1984**, *82*, 3235–3264.
- (68) Casida, M. E., Time-dependent density functional response theory for molecules In *Recent Advances in Density Functional Methods*; WORLD SCIENTIFIC: 1995, pp 155–192.
- (69) Norman, P. A perspective on nonresonant and resonant electronic response theory for time-dependent molecular properties. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20519–20535.
- (70) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent advances in wave function-based methods of molecular-property calculations. *Chem. Rev.* **2012**, *112*, 543–631.
- (71) Norman, P.; Ruud, K.; Saue, T., *Principles and Practices of Molecular Properties*; John Wiley & Sons Ltd.: 2018.
- (72) Gomes, A. S. P.; Jacob, C. R. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2012**, *108*, 222.
- (73) Defusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. Modeling solvent effects on electronic excited states. *J. Phys. Chem. Lett.* **2011**, *2*, 2184–2192.
- (74) Fujisawa, T.; Terazima, M.; Kimura, Y.; Maroncelli, M. Resonance Raman study of the solvation of P-nitroaniline in supercritical water. *Chem. Phys. Lett.* **2006**, *430*, 303–308.
- (75) Kosenkov, D.; Slipchenko, L. V. Solvent effects on the electronic transitions of P-nitroaniline: A QM/EFP study. *J. Phys. Chem. A* **2011**, *115*, 392–401.
- (76) Slipchenko, L. V. Solvation of the excited states of chromophores in polarizable environment: Orbital relaxation versus polarization. *J. Phys. Chem. A* **2010**, *114*, 8824–8830.
- (77) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.
- (78) Gustafsson, C.; Linares, M.; Norman, P. Quantum Mechanics/Molecular mechanics density functional theory simulations of the optical properties fingerprinting the ligand-binding of pentameric formyl thiophene acetic acid in Amyloid- $\beta$ (1-42). *J. Phys. Chem. A* **2020**, *124*, 875–888.

- (79) Gustafsson, C.; Shirani, H.; Leira, P.; Rehn, D. R.; Linares, M.; Nilsson, K. P. R.; Norman, P.; Lindgren, M. Deciphering the electronic transitions of thiophene-based donor-acceptor-donor pentameric ligands utilized for multimodal fluorescence microscopy of protein aggregates. *ChemPhysChem* **2021**, *22*, 323–335.
- (80) Collins, M. A.; Bettens, R. P. A. Energy-Based Molecular Fragmentation Methods. *Chem. Rev.* **2015**, *115*, 5607–5642.
- (81) Herbert, J. M. Fantasy versus reality in fragment-based quantum chemistry. *J. Chem. Phys.* **2019**, *151*, 170901.
- (82) Hellmers, J.; König, C. A unified and flexible formulation of molecular fragmentation schemes. *J. Chem. Phys.* **2021**, *155*, 164105.
- (83) Li, W.; Li, S.; Jiang, Y. Generalized Energy-Based Fragmentation Approach for Computing the Ground-State Energies and Properties of Large Molecules. *J. Phys. Chem. A* **2007**, *111*, 2193–2199.
- (84) Hua, W.; Fang, T.; Li, W.; Yu, J.-G.; Li, S. Geometry Optimizations and Vibrational Spectra of Large Molecules from a Generalized Energy-Based Fragmentation Approach. *J. Phys. Chem. A* **2008**, *112*, 10864–10872.
- (85) Hua, S.; Hua, W.; Li, S. An Efficient Implementation of the Generalized Energy-Based Fragmentation Approach for General Large Molecules. *J. Phys. Chem. A* **2010**, *114*, 8126–8134.
- (86) Richard, R. M.; Herbert, J. M. Many-Body Expansion with Overlapping Fragments: Analysis of Two Approaches. *J. Chem. Theory Comput.* **2013**, *9*, 1408–1416.
- (87) Wang, X.; Liu, J.; Zhang, J. Z. H.; He, X. Electrostatically Embedded Generalized Molecular Fractionation with Conjugate Caps Method for Full Quantum Mechanical Calculation of Protein Energy. *J. Phys. Chem. A* **2013**, *117*, 7149–7161.
- (88) Gadre, S. R.; Shirsat, R. N.; Limaye, A. C. Molecular Tailoring Approach for Simulation of Electrostatic Properties. *J. Phys. Chem.* **1994**, *98*, 9165–9169.
- (89) Netzloff, H. M.; Collins, M. A. Ab initio energies of nonconducting crystals by systematic fragmentation. *J. Chem. Phys.* **2007**, *127*, 134113.
- (90) Le, H.-A.; Tan, H.-J.; Ouyang, J. F.; Bettens, R. P. A. Combined Fragmentation Method: A Simple Method for Fragmentation of Large Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 469–478.
- (91) Huang, L.; Massa, L.; Karle, J. Kernel energy method illustrated with peptides. *Int. J. Quantum Chem.* **2005**, *103*, 808–817.
- (92) Mayhall, N. J.; Raghavachari, K. Many-Overlapping-Body (MOB) Expansion: A Generalized Many Body Expansion for Non-disjoint Monomers in Molecular Fragmentation Calculations of Covalent Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 2669–2675.

- (93) Richard, R. M.; Herbert, J. M. A generalized many-body expansion and a unified view of fragment-based methods in electronic structure theory. *J. Chem. Phys.* **2012**, *137*, 064113.
- (94) Collins, M. A.; Cvitkovic, M. W.; Bettens, R. P. A. The Combined Fragmentation and Systematic Molecular Fragmentation Methods. *Acc. Chem. Res.* **2014**, *47*, 2776–2785.
- (95) Hirata, S. Fast electron-correlation methods for molecular crystals: An application to the  $\alpha$ ,  $\beta 1$ , and  $\beta 2$  modifications of solid formic acid. *The Journal of Chemical Physics* **2008**, *129*, 204104.
- (96) Park, K. S.; Ni, Z.; Côté, A. P.; Choi, J. Y.; Huang, R.; Uribe-Romo, F. J.; Chae, H. K.; O’Keeffe, M.; Yaghi, O. M. Exceptional chemical and thermal stability of zeolitic imidazolate frameworks. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10186–10191.
- (97) Chen, B.; Yang, Z.; Zhu, Y.; Xia, Y. Zeolitic imidazolate framework materials: recent progress in synthesis and applications. *J. Mater. Chem. A* **2014**, *2*, 16811–16831.
- (98) Vandezande, W.; Janssen, K. P. F.; Delpoort, F.; Ameloot, R.; De Vos, D. E.; Lamertyn, J.; Roeffaers, M. B. J. Parts per Million Detection of Alcohol Vapors via Metal Organic Framework Functionalized Surface Plasmon Resonance Sensors. *Anal. Chem.* **2017**, *89*, 4480–4487.
- (99) Kim, K.-J.; Lu, P.; Culp, J. T.; Ohodnicki, P. R. Metal–Organic Framework Thin Film Coated Optical Fiber Sensors: A Novel Waveguide-Based Chemical Sensing Platform. *ACS Sens.* **2018**, *3*, 386–394.
- (100) Demessence, A.; Boissière, C.; Grosso, D.; Horcajada, P.; Serre, C.; Férey, G.; A. Soler-Illia, G. J. A.; Sanchez, C. Adsorption properties in high optical quality nanoZIF-8 thin films with tunable thickness. *J. Mater. Chem.* **2010**, *20*, 7676–7681.
- (101) Dalstein, O.; Ceratti, D. R.; Boissière, C.; Grosso, D.; Cattoni, A.; Faustini, M. Nanoimprinted, Submicrometric, MOF-Based 2D Photonic Structures: Toward Easy Selective Vapors Sensing by a Smartphone Camera. *Adv. Funct. Mater.* **2016**, *26*, 81–90.
- (102) Schweinefuß, M.; Springer, S.; A. Baburin, I.; Hikov, T.; Huber, K.; Leoni, S.; Wiebcke, M. Zeolitic imidazolate framework-71 nanocrystals and a novel SOD-type polymorph: solution mediated phase transformations, phase selection via coordination modulation and a density functional theory derived energy landscape. *Dalton Trans.* **2014**, *43*, 3528–3536.
- (103) Morris, W.; Doonan, C. J.; Furukawa, H.; Banerjee, R.; Yaghi, O. M. Crystals as Molecules: Postsynthesis Covalent Functionalization of Zeolitic Imidazolate Frameworks. *J. Am. Chem. Soc.* **2008**, *130*, 12626–12627.



- (104) Mondal, S. S.; Hovestadt, M.; Dey, S.; Paula, C.; Glomb, S.; Kelling, A.; Schilde, U.; Janiak, C.; Hartmann, M.; Holdt, H.-J. Synthesis of a partially fluorinated ZIF-8 analog for ethane/ethene separation. *CrystEngComm* **2017**, *19*, 5882–5891.
- (105) Tu, M.; Reinsch, H.; Rodríguez-Hermida, S.; Verbeke, R.; Stassin, T.; Egger, W.; Dickmann, M.; Dieu, B.; Hofkens, J.; Vankelecom, I. F. J.; Stock, N.; Ameloot, R. Reversible Optical Writing and Data Storage in an Anthracene-Loaded Metal-Organic Framework. *Angew. Chem. Int. Ed.* **2019**, *58*, 2423–2427.
- (106) Kedenburg, S.; Vieweg, M.; Gissibl, T.; Giessen, H. Linear refractive index and absorption measurements of nonlinear optical liquids in the visible and near-infrared spectral region. *Opt. Mater. Express, OME* **2012**, *2*, 1588–1611.
- (107) Hobday, C. L.; Bennett, T. D.; Fairen-Jimenez, D.; Graham, A. J.; Morrison, C. A.; Allan, D. R.; Düren, T.; Moggach, S. A. Tuning the Swing Effect by Chemical Functionalization of Zeolitic Imidazolate Frameworks. *J. Am. Chem. Soc.* **2018**, *140*, 382–387.
- (108) Bux, H.; Chmelik, C.; van Baten, J. M.; Krishna, R.; Caro, J. Novel MOF-Membrane for Molecular Sieving Predicted by IR-Diffusion Studies and Molecular Modeling. *Adv. Mater.* **2010**, *22*, 4741–4743.
- (109) Kwon, H. T.; Jeong, H.-K. In Situ Synthesis of Thin Zeolitic-Imidazolate Framework ZIF-8 Membranes Exhibiting Exceptionally High Propylene/Propane Separation. *J. Am. Chem. Soc.* **2013**, *135*, 10763–10768.
- (110) Pan, Y.; Li, T.; Lestari, G.; Lai, Z. Effective separation of propylene/propane binary mixtures by ZIF-8 membranes. *J. Membr. Sci.* **2012**, *390–391*, 93–98.
- (111) Kwon, H. T.; Jeong, H.-K. Highly propylene-selective supported zeolite - imidazolate framework (ZIF-8) membranes synthesized by rapid microwave-assisted seeding and secondary growth. *Chem. Comm.* **2013**, *49*, 3854–3856.
- (112) Bux, H.; Liang, F.; Li, Y.; Cravillon, J.; Wiebcke, M.; Caro, J. Zeolitic Imidazolate Framework Membrane with Molecular Sieving Properties by Microwave-Assisted Solvothermal Synthesis. *J. Am. Chem. Soc.* **2009**, *131*, 16000–16001.
- (113) Redel, E.; Wang, Z.; Walheim, S.; Liu, J.; Gliemann, H.; Wöll, C. On the dielectric and optical properties of surface-anchored metal-organic frameworks: A study on epitaxially grown thin films. *Appl. Phys. Lett.* **2013**, *103*, 091903.
- (114) Wu, X.; Niknam Shahrak, M.; Yuan, B.; Deng, S. Synthesis and characterization of zeolitic imidazolate framework ZIF-7 for CO<sub>2</sub> and CH<sub>4</sub> separation. *Microporous Mesoporous Mater.* **2014**, *190*, 189–196.
- (115) Coudert, F.-X.; Fuchs, A. H. Computational characterization and prediction of metal-organic framework properties. *Coordination Chemistry Reviews* **2016**, *307*, 211–236.

- (116) Morris, W.; Leung, B.; Furukawa, H.; Yaghi, O. K.; He, N.; Hayashi, H.; Houndonougbo, Y.; Asta, M.; Laird, B. B.; Yaghi, O. M. A Combined Experimental-Computational Investigation of Carbon Dioxide Capture in a Series of Isoreticular Zeolitic Imidazolate Frameworks. *J. Am. Chem. Soc.* **2010**, *132*, 11006–11008.
- (117) Colón, Y. J.; Q. Snurr, R. High-throughput computational screening of metal-organic frameworks. *Chem. Soc. Rev.* **2014**, *43*, 5735–5749.
- (118) Verploegh, R. J.; Kulkarni, A.; Boulfelfel, S. E.; Haydak, J. C.; Tang, D.; Sholl, D. S. Screening Diffusion of Small Molecules in Flexible Zeolitic Imidazolate Frameworks Using a DFT-Parameterized Force Field. *J. Phys. Chem. C* **2019**, *123*, 9153–9167.
- (119) Wu, X.-P.; Choudhuri, I.; Truhlar, D. G. Computational Studies of Photocatalysis with Metal–Organic Frameworks. *ENERGY Environ. Mater.* **2019**, *2*, 251–263.
- (120) Yang, Q.; Liu, D.; Zhong, C.; Li, J.-R. Development of Computational Methodologies for Metal–Organic Frameworks and Their Application in Gas Separations. *Chem. Rev.* **2013**, *113*, 8261–8323.
- (121) Valenzano, L.; Civalleri, B.; Chavan, S.; Bordiga, S.; Nilsen, M. H.; Jakobsen, S.; Lillerud, K. P.; Lamberti, C. Disclosing the Complex Structure of UiO-66 Metal Organic Framework: A Synergic Combination of Experiment and Theory. *Chem. Mater.* **2011**, *23*, 1700–1718.
- (122) Chizallet, C.; Bats, N. External Surface of Zeolite Imidazolate Frameworks Viewed Ab Initio: Multifunctionality at the Organic-Inorganic Interface. *J. Phys. Chem. Lett.* **2010**, *1*, 349–353.
- (123) Chizallet, C.; Bats, N. External Surface of Zeolite Imidazolate Frameworks Viewed Ab Initio: Multifunctionality at the Organic-Inorganic Interface. *J. Phys. Chem. Lett.* **2010**, *1*, 349–353.
- (124) Choudhuri, I.; Ye, J.; Truhlar, D. G. Computational quantum chemistry of metal-organic frameworks. *Chemical Physics Reviews* **2023**, *4*, 031304.
- (125) Valenzano, L.; Civalleri, B.; Chavan, S.; Bordiga, S.; Nilsen, M. H.; Jakobsen, S.; Lillerud, K. P.; Lamberti, C. Disclosing the Complex Structure of UiO-66 Metal Organic Framework: A Synergic Combination of Experiment and Theory. *Chem. Mater.* **2011**, *23*, 1700–1718.
- (126) Ni, B.; Sun, W.; Kang, J.; Zhang, Y. Understanding the Linear and Second-Order Nonlinear Optical Properties of UiO-66-Derived Metal–Organic Frameworks: A Comprehensive DFT Study. *J. Phys. Chem. C* **2020**, *124*, 11595–11608.
- (127) Treger, M.; Hannebauer, A.; Schaate, A.; L. Budde, J.; Behrens, P.; M. Schneider, A. Tuning the optical properties of the metal-organic framework UiO-66 via ligand functionalization. *Phys. Chem. Chem. Phys.* **2023**, *25*, 6333–6341.
- (128) Treger, M.; Hannebauer, A.; Behrens, P.; M. Schneider, A. Development of high refractive index UiO-66 framework derivatives via ligand halogenation. *Phys. Chem. Chem. Phys.* **2023**, *25*, 15391–15399.

- (129) Treger, M.; König, C.; Behrens, P.; M. Schneider, A. Fragment-based approach for the efficient calculation of the refractive index of metal-organic frameworks. *Phys. Chem. Chem. Phys.* **2023**, *25*, 19013–19023.
- (130) Lorenz, L. Ueber die Refraktionsconstante. *Ann. Phys.* **1880**, *247*, 70–103.
- (131) Lorentz, H. A. Ueber die Beziehung zwischen der Fortpflanzungsgeschwindigkeit des Lichtes und der Körperdichte. *Ann. Phys.* **1880**, *245*, 641–665.
- (132) Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **1927**, *389*, 457–484.
- (133) Slater, J. C. The theory of complex spectra. *Phys. Rev.* **1929**, *34*, 1293–1322.
- (134) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion. *Math. Proc. Camb. Phil. Soc.* **1928**, *24*, 111–132.
- (135) Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Physik* **1930**, *61*, 126–148.
- (136) Wigner, E.; Seitz, F. On the Constitution of Metallic Sodium. *Phys. Rev.* **1933**, *43*, 804–810.
- (137) Boys, S. F.; Egerton, A. C. Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system. *Proc. Math. Phys. Eng. Sci.* **1950**, *200*, 542–554.
- (138) Coester, F. Bound states of a many-particle system. *Nuc. Phys.* **1958**, *7*, 421–424.
- (139) Lührmann, K. H.; Kümmel, H. Equations for linked clusters and Brueckner-Bethe theory. *Nucl. Phys. A* **1972**, *194*, 225–236.
- (140) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
- (141) Møller, Chr.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (142) Burke, K. Perspective on density functional theory. *J. Chem. Phys.* **2012**, *136*, 150901.
- (143) Thomas, L. H. The calculation of atomic fields. *Proc. Cambridge Phil. Soc.* **1927**, *23*, 542–548.
- (144) Fermi, E. Statistical method to determine some properties of atoms. *Rend. Accad. Naz.* **1927**, *6*, 5.
- (145) Fermi, E. Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente. *Z. Physik* **1928**, *48*, 73–79.
- (146) Sheldon, J. W. Use of the Statistical Field Approximation in Molecular Physics. *Phys. Rev.* **1955**, *99*, 1291–1301.

- (147) Teller, E. On the Stability of Molecules in the Thomas-Fermi Theory. *Rev. Mod. Phys.* **1962**, *34*, 627-631.
- (148) Dirac, P. a. M. Note on Exchange Phenomena in the Thomas Atom. *Math. Proc. Camb.* **1930**, *26*, 376-385.
- (149) Von Weizsäcker, C. F. Zur Theorie der Kernmassen. *Z. Physik* **1935**, 431.
- (150) Balázs, N. L. Formation of Stable Molecules within the Statistical Theory of Atoms. *Phys. Rev.* **1967**, *156*, 42-47.
- (151) Lieb, E. H.; Simon, B. Thomas-Fermi Theory Revisited. *Phys. Rev. Lett.* **1973**, *31*, 681-683.
- (152) Lieb, E. H.; Simon, B. The Thomas-Fermi theory of atoms, molecules and solids. *Adv. Math.* **1977**, *23*, 22-116.
- (153) Chan, G. K.-L.; Cohen, A. J.; Handy, N. C. Thomas-Fermi-Dirac-von Weizsäcker models in finite systems. *J. Chem. Phys.* **2001**, *114*, 631-638.
- (154) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864-B871.
- (155) Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 6062-6065.
- (156) Parr, R. G.; Yang, W., *Density-functional theory of atoms and molecules*; Oxford University Press, USA: 1994.
- (157) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133-A1138.
- (158) Ceperley, D. M.; Alder, B. J. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.* **1980**, *45*, 566-569.
- (159) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200-1211.
- (160) Perdew, J. P.; Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **1992**, *45*, 13244-13249.
- (161) Jensen, F., *Introduction to Computational Chemistry*; John Wiley & Sons Inc: 2017.
- (162) Engel, E.; Dreizler, R. M. Density functional theory. *Theor. Math. Phys.* **2011**, 11-56.
- (163) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098-3100.
- (164) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.

- (165) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (166) Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- (167) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem. B* **1994**, *98*, 11623–11627.
- (168) Cramer, C. J., *Essentials of computational chemistry: theories and models*; John Wiley & Sons: 2013.
- (169) Savin, A.; Flad, H.-J. Density functionals for the Yukawa electron-electron interaction. *Int. J. Quantum Chem.* **1995**, *56*, 327–332.
- (170) Adamson, R. D.; Dombroski, J. P.; Gill, P. M. W. Chemistry without Coulomb tails. *Chem. Phys. Lett.* **1996**, *254*, 329–336.
- (171) Gill, P. M. W.; Adamson, R. D.; Pople, J. A. Coulomb-attenuated exchange energy density functionals. *Mol. Phys.* **1996**, *88*, 1005–1009.
- (172) Jaramillo, J.; Scuseria, G. E.; Ernzerhof, M. Local hybrid functionals. *J. Chem. Phys.* **2003**, *118*, 1068–1073.
- (173) Pernal, K.; Gidopoulos, N. I.; Pastorczak, E., Chapter Eleven - Excitation Energies of Molecules from Ensemble Density Functional Theory: Multiconfiguration Approaches In *Advances in Quantum Chemistry*, Hoggan, P. E., Ozdogan, T., Eds.; Electron Correlation in Molecules - ab initio Beyond Gaussian Quantum Chemistry, Vol. 73; Academic Press: 2016, pp 199–229.
- (174) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (175) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.
- (176) Sure, R.; Grimme, S. Corrected small basis set Hartree-Fock method for large systems. *J. Comp. Chem.* **2013**, *34*, 1672–1685.
- (177) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.* **2015**, *143*, 054107.
- (178) Brandenburg, J. G.; Bannwarth, C.; Hansen, A.; Grimme, S. B97-3c: A revised low-cost variant of the B97-D density functional method. *J. Chem. Phys.* **2018**, *148*, 064104.
- (179) Becke, A. D. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.
- (180) Schatz, G. C.; Ratner, M. A., *Quantum Mechanics in Chemistry*; Courier Corporation: 2002.

- (181) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* **2005**, *105*, 4009-4037.
- (182) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997-1000.
- (183) Kramers, H. A.; Heisenberg, W. Über die Streuung von Strahlung durch Atome. *Z. Physik* **1925**, *31*, 681-708.
- (184) Autschbach, J.; Ziegler, T. Calculating molecular electric and magnetic properties from time-dependent density functional response theory. *J. Chem. Phys.* **2002**, *116*, 891-896.
- (185) Condon, E. U. Theories of Optical Rotatory Power. *Rev. Mod. Phys.* **1937**, *9*, 432-457.
- (186) Condon, E. U.; Shortley, G. H., *The Theory of Atomic Spectra*; Cambridge University Press: 1951.
- (187) Casida, M. E., Time-dependent density functional response theory for molecules In *Recent advances in density functional methods*, 1995, pp 155-192.
- (188) Thompson, M. A. QM/MMpol: A Consistent Model for Solute/Solvent Polarization. Application to the Aqueous Solvation and Spectroscopy of Formaldehyde, Acetaldehyde, and Acetone. *J. Phys. Chem.* **1996**, *100*, 14492-14507.
- (189) Sulpizi, M.; Röhrig, U. F.; Hutter, J.; Rothlisberger, U. Optical properties of molecules in solution via hybrid TDDFT/MM simulations. *Int. J. Quantum Chem.* **2005**, *101*, 671-682.
- (190) König, C.; Neugebauer, J. Quantum chemical description of absorption properties and excited-state processes in photosynthetic systems. *ChemPhysChem* **2012**, *13*, 386-425.
- (191) Rossano-Tapia, M.; Brown, A. Quantum Mechanical/Molecular mechanical studies of photophysical properties of fluorescent proteins. *WIREs Comput. Mol. Sci.* **2022**, *12*, e1557.
- (192) Jones, L. O.; Mosquera, M. A.; Schatz, G. C.; Ratner, M. A. Embedding Methods for Quantum Chemistry: Applications from Materials to Life Sciences. *J. Am. Chem. Soc.* **2020**, *142*, 3281-3295.
- (193) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited states in solution through polarizable embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721-3734.
- (194) Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050-8053.
- (195) Senatore, G.; Subbaswamy, K. R. Density dependence of the dielectric constant of rare-gas crystals. *Phys. Rev. B* **1986**, *34*, 5754-5757.
- (196) Johnson, M. D.; Subbaswamy, K. R.; Senatore, G. Hyperpolarizabilities of alkali halide crystals using the local-density approximation. *Phys. Rev. B* **1987**, *36*, 9202-9211.

- (197) Cortona, P. Self-consistently determined properties of solids without band-structure calculations. *Phys. Rev. B* **1991**, *44*, 8454–8458.
- (198) Goez, A.; Neugebauer, J., Embedding Methods in Quantum Chemistry In *Frontiers of Quantum Chemistry*, Wójcik, M. J., Nakatsuji, H., Kirtman, B., Ozaki, Y., Eds.; Springer: Singapore, 2018, pp 139–179.
- (199) Wesolowski, T. A., One-Electron Equations for Embedded Electron Density: Challenge for Theory and Practical Payoffs in Multi-Level Modelling of Complex Polyatomic Systems In *Computational Chemistry: Reviews of Current Trends*; Computational Chemistry: Reviews of Current Trends Volume 10, Vol. Volume 10; WORLD SCIENTIFIC: 2006, pp 1–82.
- (200) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.
- (201) Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. Density Differences in Embedding Theory with External Orbital Orthogonality. *J. Phys. Chem. A* **2014**, *118*, 9182–9200.
- (202) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.
- (203) Kaminski, J. W.; Gusarov, S.; Wesolowski, T. A.; Kovalenko, A. Modeling solvatochromic shifts using the orbital-free embedding potential at statistically mechanically averaged solvent density. *J. Phys. Chem. A* **2010**, *114*, 6082–6096.
- (204) Wesolowski, T. A.; Weber, J. Kohn - Sham equations with constrained electron density: An iterative evaluation of the ground-state electron density of interacting molecules. *Chem. Phys. Lett.* **1996**, *248*, 71–76.
- (205) Chulhai, D. V.; Jensen, L. Frozen density embedding with external orthogonality in delocalized covalent systems. *J. Chem. Theory Comput.* **2015**, *11*, 3080–3088.
- (206) Fux, S.; Kiewisch, K.; Jacob, C. R.; Neugebauer, J.; Reiher, M. Analysis of electron density distributions from subsystem density functional theory applied to coordination bonds. *Chem. Phys. Lett.* **2008**, *461*, 353–359.
- (207) Schlüns, D.; Klahr, K.; Mück-Lichtenfeld, C.; Visscher, L.; Neugebauer, J. Subsystem-DFT potential-energy curves for weakly interacting systems. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14323–14341.
- (208) Lembarki, A.; Chermette, H. Obtaining a gradient-corrected kinetic-energy functional from the Perdew-Wang exchange functional. *Phys. Rev. A* **1994**, *50*, 5328–5331.
- (209) Fux, S.; Jacob, C. R.; Neugebauer, J.; Visscher, L.; Reiher, M. Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. *J. Chem. Phys.* **2010**, *132*.

- (210) Jacob, C. R.; Visscher, L., Towards the Description of Covalent Bonds in Subsystem Density-Functional Theory In *Recent Progress in Orbital-free Density Functional Theory*; Recent Advances in Computational Chemistry, Vol. 6; WORLD SCIENTIFIC: 2012, pp 297-322.
- (211) Wesolowski, T. A. Hydrogen-Bonding-Induced Shifts of the Excitation Energies in Nucleic Acid Bases: An Interplay between Electrostatic and Electron Density Overlap Effects. *J. Am. Chem. Soc.* **2004**, *126*, 11444-11445.
- (212) Scholz, L.; Neugebauer, J. Protein Response Effects on Cofactor Excitation Energies from First Principles: Augmenting Subsystem Time-Dependent Density-Functional Theory with Many-Body Expansion Techniques. *J. Chem. Theory Comput.* **2021**, *17*, 6105-6121.
- (213) Neugebauer, J. Chromophore-specific theoretical spectroscopy: From subsystem density functional theory to mode-specific vibrational spectroscopy. *Phys. Rep.* **2010**, *489*, 1-87.
- (214) König, C.; Schlüter, N.; Neugebauer, J. Direct determination of exciton couplings from subsystem time-dependent density-functional theory within the Tamm-Dancoff approximation. *J. Chem. Phys.* **2013**, *138*, 34104.
- (215) Neugebauer, J.; Curutchet, C.; Muñoz-Losa, A.; Mennucci, B.; Muñoz-Losa, A.; Mennucci, B.; Muñoz-Losa, A.; Mennucci, B. A subsystem TDDFT approach for solvent screening effects on excitation energy transfer couplings. *J. Chem. Theory Comput.* **2010**, *6*, 1843-1851.
- (216) Neugebauer, J. Couplings between electronic transitions in a subsystem formulation of time-dependent density functional theory. *J. Chem. Phys.* **2007**, *126*, 134116.
- (217) Neugebauer, J. On the calculation of general response properties in subsystem density functional theory. *J. Chem. Phys.* **2009**, *131*, 084104.
- (218) Höfener, S.; Visscher, L. Wave Function Frozen-Density Embedding: Coupled Excitations. *J. Chem. Theory Comput.* **2016**, *12*, 549-557.
- (219) Tölle, J.; Böckers, M.; Neugebauer, J. Exact subsystem time-dependent density-functional theory. *J. Chem. Phys.* **2019**, *150*, 181101.
- (220) Tölle, J.; Böckers, M.; Niemeyer, N.; Neugebauer, J. Inter-subsystem charge-transfer excitations in exact subsystem time-dependent density-functional theory. *J. Chem. Phys.* **2019**, *151*, 174109.
- (221) Tölle, J.; Neugebauer, J. The Seamless Connection of Local and Collective Excited States in Subsystem Time-Dependent Density Functional Theory. *J. Phys. Chem. Lett.* **2022**, *13*, 1003-1018.
- (222) Olsen, J. M. H.; Kongsted, J., Molecular properties through polarizable embedding In *Adv. Quantum Chem.* Elsevier: 2011; Vol. 61, pp 107-143.
- (223) Olsen, M. H. Development of Quantum Chemical Methods towards Rationalization and Optimal Design of Photoactive Proteins, 2012.



- (224) Helgaker, T.; Jorgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory*; John Wiley & Sons: 2013; 949 pp.
- (225) Stone, A., *The Theory of Intermolecular Forces*; Oxford University Press: 2013.
- (226) Ahlrichs, R. Convergence properties of the intermolecular force series (1/R-expansion). *Theor. Chim. Acta* **1976**, *41*, 7-15.
- (227) Gordon, M. S.; Slipchenko, L.; Li, H.; Jensen, J. H. Chapter 10 the effective fragment potential: A general method for predicting intermolecular interactions. *Annu. Rep. Comput. Chem.* **2007**, *3*, 177-193.
- (228) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for Large Systems, with Applications to Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 46-53.
- (229) Wen, S.; Nanda, K.; Huang, Y.; O. Beran, G. J. Practical quantum mechanics-based fragment methods for predicting molecular crystal properties. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7578-7590.
- (230) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. Molecular tailoring approach for geometry optimization of large molecules: Energy evaluation and parallelization strategies. *J. Chem. Phys.* **2006**, *125*, 104109.
- (231) Rahalkar, A. P.; Ganesh, V.; Gadre, S. R. Enabling ab initio Hessian and frequency calculations of large molecules. *J. Chem. Phys.* **2008**, *129*, 234101.
- (232) Rahalkar, A. P.; Katouda, M.; Gadre, S. R.; Nagase, S. Molecular tailoring approach in conjunction with MP2 and Ri-MP2 codes: A comparison with fragment molecular orbital method. *J. Comp. Chem.* **2010**, *31*, 2405-2418.
- (233) Yeole, S. D.; Gadre, S. R. On the applicability of fragmentation methods to conjugated  $\pi$  systems within density functional framework. *J. Chem. Phys.* **2010**, *132*, 094102.
- (234) Suárez, E.; Díaz, N.; Suárez, D. Thermochemical Fragment Energy Method for Biomolecules: Application to a Collagen Model Peptide. *J. Chem. Theory Comput.* **2009**, *5*, 1667-1679.
- (235) Zhang, D. W.; Zhang, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *J. Chem. Phys.* **2003**, *119*, 3599-3605.
- (236) Hellmers, J.; Hedegård, E. D.; König, C. Fragmentation-Based Decomposition of a Metalloenzyme-Substrate Interaction: A Case Study for a Lytic Polysaccharide Monooxygenase. *J. Phys. Chem. B* **2022**, *126*, 5400-5412.
- (237) Hua, D.; Leverentz, H. R.; Amin, E. A.; Truhlar, D. G. Assessment and Validation of the Electrostatically Embedded Many-Body Expansion for Metal-Ligand Bonding. *J. Chem. Theory Comput.* **2011**, *7*, 251-255.
- (238) Yuan, D.; Li, Y.; Li, W.; Li, S. Structures and properties of large supramolecular coordination complexes predicted with the generalized energy-based fragmentation method. *Phys. Chem. Chem. Phys.* **2018**, *20*, 28894-28902.

- (239) Kurbanov, E. K.; Leverentz, H. R.; Truhlar, D. G.; Amin, E. A. Analysis of the Errors in the Electrostatically Embedded Many-Body Expansion of the Energy and the Correlation Energy for Zn and Cd Coordination Complexes with Five and Six Ligands and Use of the Analysis to Develop a Generally Successful Fragmentation Strategy. *J. Chem. Theory Comput.* **2013**, *9*, 2617–2628.
- (240) Chemla, D. S.; Oudar, J. L.; Jerphagnon, J. Origin of the second-order optical susceptibilities of crystalline substituted benzene. *Phys. Rev. B* **1975**, *12*, 4534–4546.
- (241) Lin, C.; Wu, K. Theoretical studies on the nonlinear optical susceptibilities of 3-methoxy-4-hydroxy-benzaldehyde crystal. *Chem. Phys. Lett.* **2000**, *321*, 83–88.
- (242) Wu, K.; Chen, C. Theoretical studies for novel non-linear optical crystals. *J. Cryst. Growth* **1996**, *166*, 533–536.
- (243) Wu, K.; Snijders, J. G.; Lin, C. Reinvestigation of Hydrogen Bond Effects on the Polarizability and Hyperpolarizability of Urea Molecular Clusters. *J. Phys. Chem. B* **2002**, *106*, 8954–8958.
- (244) Shannon, R. D.; Fischer, R. X. Empirical electronic polarizabilities in oxides, hydroxides, oxyfluorides, and oxychlorides. *Phys. Rev. B* **2006**, *73*, 235111.
- (245) Shannon, R. D.; Fischer, R. X. Empirical electronic polarizabilities of ions for the prediction and interpretation of refractive indices: Oxides and oxysalts. *Am. Min.* **2016**, *101*, 2288–2300.
- (246) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. Derivative studies in hartree-fock and møller-plesset theories. *Int. J. Quantum Chem.* **1979**, *16*, 225–241.
- (247) Liang, W.; Zhao, Y.; Head-Gordon, M. An efficient approach for self-consistent-field energy and energy second derivatives in the atomic-orbital basis. *J. Chem. Phys.* **2005**, *123*, 194106.
- (248) Handy, N. C.; Tozer, D. J.; Laming, G. J.; Murray, C. W.; Amos, R. D. Analytic Second Derivatives of the Potential Energy Surface. *Isr. J. Chem.* **1993**, *33*, 331–344.
- (249) Johnson, B. G.; Fisch, M. J. An implementation of analytic second derivatives of the gradient-corrected density functional energy. *J. Chem. Phys.* **1994**, *100*, 7429–7442.
- (250) Ang, S. T.; Pal, A.; Manzhos, S. Comparison of optical absorption spectra of organic molecules and aggregates computed from real frequency dependent polarizability to TD-DFT and the dipole approximation. *J. Chem. Phys.* **2018**, *149*, 044114.
- (251) Pal, A.; Arabnejad, S.; Yamashita, K.; Manzhos, S. Influence of the aggregate state on band structure and optical properties of C60 computed with different methods. *J. Chem. Phys.* **2018**, *148*, 204301.

- (252) Afzal, M. A. F.; Cheng, C.; Hachmann, J. Combining first-principles and data modeling for the accurate prediction of the refractive index of organic polymers. *J. Chem. Phys.* **2018**, *148*, 241712.
- (253) Afzal, M. A. F.; Hachmann, J. Benchmarking DFT approaches for the calculation of polarizability inputs for refractive index predictions in organic polymers. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4452–4460.
- (254) Park, S. S.; Lee, S.; Bae, J. Y.; Hagelberg, F. Refractive indices of liquid-forming organic compounds by density functional theory. *Chem. Phys. Lett.* **2011**, *511*, 466–470.
- (255) Haichuan, C.; Guozhu, J.; Liang, Z.; Yinshu, H. Density functional theory calculation of refractive indices of the highly polar compounds. *Phys. Chem. Liq.* **2015**, *53*, 435–442.
- (256) Maekawa, S.; Moorthi, K. Polarizabilities from Long-Range Corrected DFT Calculations. *J. Chem. Eng. Data* **2014**, *59*, 3160–3166.
- (257) Maekawa, S.; Moorthi, K.; Shigeta, Y. Refractive indices of organo-metallic and -metalloid compounds: A long-range corrected DFT study. *J. Comp. Chem.* **2016**, *37*, 2759–2769.
- (258) Ksianzou, V.; Velagapudi, R. K.; Grimm, B.; Schrader, S. Polarization-dependent optical characterization of poly(phenylquinoxaline) thin films. *J. Appl. Phys.* **2006**, *100*, 063106.
- (259) Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. Theoretical and Numerical Comparison of Quantum- and Classical Embedding Models for Optical Spectra. *J. Phys. Chem. A* **2023**, *127*, 5689–5703.
- (260) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (261) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (262) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.
- (263) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent molecular orbital methods. XII. Further extensions of Gaussian—Type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (264) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- (265) Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

- (266) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: A different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.
- (267) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (268) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (269) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (270) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (271) Páll, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E. In *Int. Conf. Exascale appl. Softw.* Springer: 2014, pp 3–27.
- (272) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D., et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.
- (273) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (274) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
- (275) Lindahl, E.; Hess, B.; Van Der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *Mol. Model. Annu.* **2001**, *7*, 306–317.
- (276) Berendsen, H. J. C.; vander Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
- (277) Lindahl, E.; Abraham, M. J.; Hess, B.; van der Spoel, D. GROMACS 2019.3 source code, Zenodo, 2019.
- (278) Feller, S. E.; MacKerell, A. D. An improved empirical potential energy function for molecular simulations of phospholipids. *J. Phys. Chem. B* **2000**, *104*, 7510–7515.
- (279) Klauda, J. B.; Brooks, B. R.; MacKerell, A. D.; Venable, R. M.; Pastor, R. W. An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and a DPPC bilayer. *J. Phys. Chem. B* **2005**, *109*, 5300–5311.
- (280) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

- (281) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 14101.
- (282) Berendsen, H. J. C.; van Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A. R. H. J.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.
- (283) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.
- (284) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463-1472.
- (285) Aidas, K. et al. The Dalton quantum chemistry program system. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269-284.
- (286) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007-1023.
- (287) Olsen, J. M. H.; List, N. H.; Steinmann, C. PEAS (Polarizable Embedding Assistant Script).
- (288) Fdez. Galván, I. et al. OpenMolcas: From source code to insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925-5964.
- (289) Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: The LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494-4500.
- (290) Almlöf, J.; Taylor, P. R. General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. *J. Chem. Phys.* **1987**, *86*, 4070-4077.
- (291) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796-6806.
- (292) Olsen, J. M. H.; List, N. H.; Kristensen, K.; Kongsted, J. Accuracy of protein embedding potentials: An analysis in terms of electrostatic potentials. *J. Chem. Theory Comput.* **2015**, *11*, 1832-1842.
- (293) Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theor. Chim. Acta* **1995**, *90*, 87-114.
- (294) Olsen, J. M. H.; List, N. H.; Steinmann, C.; Steindal, A. H.; Nørby, M. S.; Reinholdt, P. PElib: The Polarizable Embedding library, Zenodo, 2020.
- (295) Marefat Khah, A.; Reinholdt, P.; Olsen, J. M. H.; Kongsted, J.; Hättig, C. Avoiding electron spill-out in QM/MM calculations on excited states with simple pseudopotentials. *J. Chem. Theory Comput.* **2020**, *16*, 1373-1381.

- (296) Jacob, C. R.; Beyhan, S. M.; Bulo, R. E.; Gomes, A. S. P.; Götz, A. W.; Kiewisch, K.; Sikkema, J.; Visscher, L. PyADF - A scripting framework for multiscale quantum chemistry. *J. Comput. Chem.* **2011**, *32*, 2328–2338.
- (297) Franchini, M.; Philipsen, P. H. T.; Visscher, L. The Becke Fuzzy Cells Integration Scheme in the Amsterdam Density Functional Program Suite. *J. Comp. Chem.* **2013**, *34*, 1819–1827.
- (298) Te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.
- (299) Perdew, J. P. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (300) Kiewisch, K.; Eickerling, G.; Reiher, M.; Neugebauer, J. Topological analysis of electron densities from Kohn-Sham and subsystem density functional theory. *J. Chem. Phys.* **2008**, *128*, 44114.
- (301) Schaftenaar, G.; Noordik, J. Molden: a pre- and post-processing program for molecular and electronic structures. *J. Comput. - Aided Mol.* **2000**, *14*, 123–134.
- (302) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.
- (303) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108.
- (304) Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Comput. Mol. Sci.* **2022**, *12*, e1606.
- (305) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Electronic structure calculations on workstation computers: The program system turbomole. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (306) TURBOMOLE V7.5.1, University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 2021.
- (307) Balasubramani, S. G. et al. TURBOMOLE: Modular program suite for *ab initio* quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020**, *152*, 184107.
- (308) Neese, F. The SHARK integral generation and digestion system. *J. Comp. Chem.* **2023**, *44*, 381–396.
- (309) Valeev, E.; Abbott, A.; Seewald, P.; Lewis, D.; Calvin, J.; Dullea, J.; Kawashima, E.; Peng, C.; Nishimra, K.; Herbst, M. F.; al., e. libint: 2.7.2. **2022**.
- (310) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. Recent developments in libxc — A comprehensive library of functionals for density functional theory. *SoftwareX* **2018**, *7*, 1–5.

- (311) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (312) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comp. Chem.* **2011**, *32*, 1456–1465.
- (313) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials (Chem. Phys. Letters 240 (1995) 283-290). *Chem. Phys.* **1995**, *242*, 652–660.
- (314) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (315) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A ‘chain-of-spheres’ algorithm for the Hartree–Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109.
- (316) Stoychev, G. L.; Auer, A. A.; Neese, F. Efficient and Accurate Prediction of Nuclear Magnetic Resonance Shielding Tensors with Double-Hybrid Density Functional Theory. *J. Chem. Theory Comput.* **2018**, *14*, 4756–4771.
- (317) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.
- (318) Stoychev, G. L.; Auer, A. A.; Neese, F. Automatic Generation of Auxiliary Basis Sets. *J. Chem. Theory Comput.* **2017**, *13*, 554–562.
- (319) Fletcher, R., *Practical methods of optimization*; John Wiley & Sons: New York, NY, USA, 1980; Vol. 1.
- (320) Klawohn, S.; Bahmann, H.; Kaupp, M. Implementation of Molecular Gradients for Local Hybrid Density Functionals Using Seminumerical Integration Techniques. *J. Chem. Theory Comput.* **2016**, *12*, 4254–4262.
- (321) Von Arnim, M.; Ahlrichs, R. Performance of parallel TURBOMOLE for density functional calculations. *J. Comp. Chem.* **1998**, *19*, 1746–1757.
- (322) Bauernschmitt, R.; Häser, M.; Treutler, O.; Ahlrichs, R. Calculation of excitation energies within time-dependent density functional theory using auxiliary basis set expansions. *Chem. Phys. Lett.* **1997**, *264*, 573–578.
- (323) Bauernschmitt, R.; Ahlrichs, R. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* **1996**, *256*, 454–464.
- (324) Furche, F.; Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* **2002**, *117*, 7433–7447.
- (325) Rappoport, D.; Furche, F. Analytical time-dependent density functional derivative methods within the RI-J approximation, an approach to excited states of large molecules. *J. Chem. Phys.* **2005**, *122*, 064105.

- (326) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
- (327) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (328) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (329) Karton, A.; Tarnopolsky, A.; Lamère, J.-F.; Schatz, G. C.; Martin, J. M. L. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.
- (330) Kozuch, S.; Martin, J. M. L. Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J. Comp. Chem.* **2013**, *34*, 2327–2344.
- (331) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *J. Chem. Phys.* **2005**, *123*, 064107.
- (332) Balabanov, N. B.; Peterson, K. A. Basis set limit electronic excitation energies, ionization potentials, and electron affinities for the 3d transition metal atoms: Coupled cluster and multireference methods. *J. Chem. Phys.* **2006**, *125*, 074110.
- (333) Peterson, K. A.; Puzzarini, C. Systematically convergent basis sets for transition metals. II. Pseudopotential-based correlation consistent basis sets for the group 11 (Cu, Ag, Au) and 12 (Zn, Cd, Hg) elements. *Theor. Chem. Acc.* **2005**, *114*, 283–296.
- (334) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (335) Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820.
- (336) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010**, *133*, 134105.
- (337) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin trans.* **1993**, *0*, 799–805.
- (338) Van Lenthe, E.; Baerends, E. J. Optimized Slater-type basis sets for the elements 1–118. *J. Comput. Chem.* **2003**, *24*, 1142–1156.



- (339) Chong, D. P.; Van Lenthe, E.; Van Gisbergen, S.; Baerends, E. J. Even-tempered slater-type orbitals revisited: From hydrogen to krypton. *J. Comp. Chem.* **2004**, *25*, 1030-1036.
- (340) Chong, D. Augmenting basis set for time-dependent density functional theory calculation of excitation energies: Slater-type orbitals for hydrogen to krypton. *Mol. Phys.* **2005**, *103*, 749-761.
- (341) Zeiss, G.; Scott, W.; Suzuki, N.; Chong, D.; Langhoff, S. Finite-field calculations of molecular polarizabilities using field-induced polarization functions: second- and third-order perturbation correlation corrections to the coupled Hartree-Fock polarizability of H<sub>2</sub>O. *Mol. Phys.* **1979**, *37*, 1543-1572.
- (342) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Phys. Rev. Lett.* **2008**, *100*, 136406.
- (343) Paier, J.; Marsman, M.; Hummer, K.; Kresse, G.; Gerber, I. C.; Ángyán, J. G. Screened hybrid density functionals applied to solids. *J. Chem. Phys.* **2006**, *124*, 154709.
- (344) Jansen, M.; Thi Minh, N. N.; Hedegård, E. D.; König, C., Quantum-derived embedding schemes for local excitations In *Chemical Modelling: Volume 17*, 2022.
- (345) Casida, M. E.; Wesolowski, T. A. Generalization of the Kohn-Sham equations with constrained electron density formalism and its time-dependent response theory formulation. *Int. J. Quantum Chem.* **2004**, *96*, 577-588.
- (346) Jacob, C. R.; Neugebauer, J.; Jensen, L.; Visscher, L. Comparison of frozen-density embedding and discrete reaction field solvent models for molecular properties. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2349-2359.
- (347) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353-5362.
- (348) Applequist, J.; Carl, J. R.; Fung, K. K. An Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *J. Am. Chem. Soc.* **1972**, *94*, 2952-2960.
- (349) List, N. H.; Jensen, H. J. A.; Kongsted, J. Local electric fields and molecular properties in heterogeneous environments through polarizable embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 10070-10080.
- (350) Fradelos, G.; Wesolowski, T. A. Importance of the Intermolecular Pauli Repulsion in Embedding Calculations for Molecular Properties: The Case of Excitation Energies for a Chromophore in Hydrogen-Bonded Environments. *J. Phys. Chem. A* **2011**, *115*, 10018-10026.

- (351) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.* **2002**, *116*, 6941-6947.
- (352) Jacob, C. R.; Beyhan, S. M.; Visscher, L. Exact functional derivative of the nonadditive kinetic-energy bifunctional in the long-distance limit. *J. Chem. Phys.* **2007**, *126*.
- (353) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Correction to "Polarizable Density Embedding: A New QM/QM/MM-based Computational Strategy". *J. Phys. Chem. A* **2015**, *119*, 6928.
- (354) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable density embedding: A solution to the electron spill-out problem in multiscale modeling. *J. Phys. Chem. Lett.* **2017**, *8*, 5949-5958.
- (355) Reinholdt, P.; Jørgensen, F. K.; Kongsted, J.; Olsen, J. M. H. Polarizable density embedding for large biomolecular systems. *J. Chem. Theory Comput.* **2020**, *16*, 5999-6006.
- (356) Hršak, D.; Olsen, J. M. H.; Kongsted, J. Polarizable density embedding coupled cluster method. *J. Chem. Theory Comput.* **2018**, *14*, 1351-1360.
- (357) Plötner, J.; Tozer, D. J.; Dreuw, A. Dependence of Excited State Potential Energy Surfaces on the Spatial Overlap of the Kohn-Sham Orbitals and the Amount of Nonlocal Hartree-Fock Exchange in Time-Dependent Density Functional Theory. *J. Chem. Theory Comput.* **2010**, *6*, 2315-2324.
- (358) Oviedo, M. B.; Ilawe, N. V.; Wong, B. M. Polarizabilities of  $\pi$ -Conjugated Chains Revisited: Improved Results from Broken-Symmetry Range-Separated DFT and New CCSD(T) Benchmarks. *J. Chem. Theory Comput.* **2016**, *12*, 3593-3602.
- (359) Thomsen, C. L.; Thøgersen, J.; Keiding, S. R. Ultrafast Charge-Transfer Dynamics: Studies of p-Nitroaniline in Water and Dioxane. *J. Phys. Chem. A* **1998**, *102*, 1062-1067.
- (360) Kovalenko, S. A.; Schanz, R.; Farztdinov, V. M.; Hennig, H.; Ernsting, N. P. Femtosecond relaxation of photoexcited para-nitroaniline: solvation, charge transfer, internal conversion and cooling. *Chem. Phys. Lett.* **2000**, *323*, 312-322.
- (361) Eriksen, J. J.; Sauer, S. P. A.; Mikkelsen, K. V.; Christiansen, O.; Jørgen, H.; Jensen, A.; Kongsted, J.; Jensen, H. J. A.; Kongsted, J. Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline. *Mol. Phys.* **2013**, *111*, 1235-1248.
- (362) Steinmann, C.; Olsen, J. M. H.; Kongsted, J. Nuclear Magnetic Shielding Constants from Quantum Mechanical/Molecular Mechanical Calculations Using Polarizable Embedding: Role of the Embedding Potential. *J. Chem. Theory Comput.* **2014**, *10*, 981-988.

- (363) Keppler, N. Frequency-dependent refractive indices for ZIF-90 and ZIF-318 under argon via ellipsometry. Institute of Inorganic Chemistry, Leibniz University Hannover, Callinstr. 9, 30167 Hannover, Germany., 2022.
- (364) Treger, M. Frequency-dependent refractive indices for ZIF-8 loaded with toluene obtained via density-functional theory with periodic boundary conditions and a HSE06 exchange–correlation functional, Institute of Inorganic Chemistry, Leibniz University Hannover, Callinstr. 9, 30167 Hannover, Germany., 2022.
- (365) Karagiari, O.; Lalonde, M. B.; Bury, W.; Sarjeant, A. A.; Farha, O. K.; Hupp, J. T. Opening ZIF-8: A Catalytically Active Zeolitic Imidazolate Framework of Sodalite Topology with Unsubstituted Linkers. *J. Am. Chem. Soc.* **2012**, *134*, 18790–18796.
- (366) Morris, W.; Stevens, C. J.; Taylor, R. E.; Dybowski, C.; Yaghi, O. M.; Garcia-Garibay, M. A. NMR and X-ray Study Revealing the Rigidity of Zeolitic Imidazolate Frameworks. *J. Phys. Chem. C* **2012**, *116*, 13307–13312.
- (367) Hait, D.; Head-Gordon, M. How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry. *Phys. Chem. Chem. Phys.* **2018**, *20*, 19800–19810.
- (368) Champagne, B.; Perpète, E. A.; van Gisbergen, S. J. A.; Baerends, E.-J.; Snijders, J. G.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. Assessment of conventional density functional schemes for computing the polarizabilities and hyperpolarizabilities of conjugated oligomers: An ab initio investigation of polyacetylene chains. *J. Chem. Phys.* **1998**, *109*, 10489–10498.
- (369) McDowell, S. A. C.; Amos, R. D.; Handy, N. C. Molecular polarisabilities - a comparison of density functional theory with standard ab initio methods. *Chem. Phys. Lett.* **1995**, *235*, 1–4.
- (370) Van Gisbergen, S. J. A.; Osinga, V. P.; Gritsenko, O. V.; van Leeuwen, R.; Snijders, J. G.; Baerends, E. J. Improved density functional theory results for frequency-dependent polarizabilities, by the use of an exchange-correlation potential with correct asymptotic behavior. *J. Chem. Phys.* **1996**, *105*, 3142–3151.
- (371) Champagne, B.; Bishop, D. M., Calculations of Nonlinear Optical Properties for the Solid State In *Advances in Chemical Physics*; John Wiley & Sons, Ltd: 2003, pp 41–92.
- (372) Champagne, B.; Bulat, F. A.; Yang, W.; Bonness, S.; Kirtman, B. Density functional theory investigation of the polarizability and second hyperpolarizability of polydiacetylene and polybutatriene chains: Treatment of exact exchange and role of correlation. *J. Chem. Phys.* **2006**, *125*, 194114.
- (373) Champagne, B.; Perpète, E. A.; Jacquemin, D.; van Gisbergen, S. J. A.; Baerends, E.-J.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. Assessment of Conventional Density Functional Schemes for Computing the Dipole Moment and (Hyper)polarizabilities of Push-Pull  $\pi$ -Conjugated Systems. *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

- (374) Champagne, B.; Perpète, E. A. Bond length alternation effects on the static electronic polarizability and second hyperpolarizability of polyacetylene chains. *Int. J. Quantum Chem.* **1999**, *75*, 441-447.
- (375) Gough, K. M. Theoretical analysis of molecular polarizabilities and polarizability derivatives in hydrocarbons. *J. Chem. Phys.* **1989**, *91*, 2424-2432.
- (376) Werner, H.-J.; Meyer, W. PNO-CI and PNO-CEPA studies of electron correlation effects. *Mol. Phys.* **1976**, *31*, 855-872.
- (377) Amos, R. An accurate ab initio study of the multipole moments and polarizabilities of methane. *Mol. Phys.* **1979**, *38*, 33-45.
- (378) Mardirossian, N.; Head-Gordon, M. Characterizing and Understanding the Remarkably Slow Basis Set Convergence of Several Minnesota Density Functionals for Intermolecular Interaction Energies. *J. Chem. Theory Comput.* **2013**, *9*, 4453-4461.
- (379) Hait, D.; Head-Gordon, M. How Accurate Is Density Functional Theory at Predicting Dipole Moments? An Assessment Using a New Database of 200 Benchmark Values. *J. Chem. Theory Comput.* **2018**, *14*, 1969-1981.
- (380) Spitznagel, G. W.; Clark, T.; Chandrasekhar, J.; Schleyer, P. V. R. Stabilization of methyl anions by first-row substituents. The superiority of diffuse function-augmented basis sets for anion calculations. *J. Comp. Chem.* **1982**, *3*, 363-371.
- (381) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Correlation Energy, with Applications to the Calculation of Accurate Second-Order Møller-Plesset Perturbation Theory Energies for Large Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 1342-1348.
- (382) Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. Evaluation of the Electrostatically Embedded Many-Body Expansion and the Electrostatically Embedded Many-Body Expansion of the Correlation Energy by Application to Low-Lying Water Hexamers. *J. Chem. Theory Comput.* **2008**, *4*, 33-41.
- (383) Antony, J.; Grimme, S. Fully ab initio protein-ligand interaction energies with dispersion corrected density functional theory. *J. Comput. Chem.* **2012**, *33*, 1730-1739.
- (384) Isegawa, M.; Wang, B.; Truhlar, D. G. Electrostatically Embedded Molecular Tailoring Approach and Validation for Peptides. *J. Chem. Theory Comput.* **2013**, *9*, 1381-1393.
- (385) Klamt, A. Calculation of UV/Vis Spectra in Solution. *J. Phys. Chem.* **1996**, *100*, 3349-3353.
- (386) Scalmani, G.; Frisch, M. J.; Mennucci, B.; Tomasi, J.; Cammi, R.; Barone, V. Geometries and properties of excited states in the gas phase and in solution: Theory and application of a time-dependent density functional theory polarizable continuum model. *J. Chem. Phys.* **2006**, *124*, 094107.

- 
- (387) Ryder, M. R.; Zeng, Z.; Titov, K.; Sun, Y.; Mahdi, E. M.; Flyagina, I.; Bennett, T. D.; Civalleri, B.; Kelley, C. S.; Frogley, M. D.; Cinque, G.; Tan, J.-C. Dielectric Properties of Zeolitic Imidazolate Frameworks in the Broad-Band Infrared Regime. *J. Phys. Chem. Lett.* **2018**, *9*, 2678–2684.
- (388) Warmbier, R.; Quandt, A.; Seifert, G. Dielectric Properties of Selected Metal–Organic Frameworks. *J. Phys. Chem. C* **2014**, *118*, 11799–11805.
- (389) Moutzouris, K.; Papamichael, M.; Betsis, S. C.; Stavrakas, I.; Hloupis, G.; Triantis, D. Refractive, dispersive and thermo-optic properties of twelve organic solvents in the visible and near-infrared. *Appl. Phys. B* **2014**, *116*, 617–622.



# Acknowledgements

First and foremost, I would like to express my sincere gratitude to Prof. Dr. Carolin König for her guidance and support throughout my time as a doctoral student. Her mentorship, patience, and dedication, especially during the challenging times of the COVID-19 pandemic, have been invaluable. It is thanks to her that I made my aspiration of participating in the ESQC a reality and had the opportunity to engage in great collaborations. Beyond quantum chemistry she fuelled my personal development and also enabled me to take on the role of a system administrator. I thank her for embarking on this journey with me, which has been marked by a fair share of difficulties, but has been even more abundant in reward.

I extend my appreciation to Prof. Dr. Jörg-August Becker for being my second referee. Moreover, I would like to thank him for his efforts in the lecture on intermolecular interactions, that has been extremely helpful. I would like to thank Prof. Dr. Annika Bande, who agreed to be my third examiner.

I am grateful to the Deutsche Forschungsgemeinschaft (DFG) for their financial support through the Emmy Noether Young Group Leader Programme (project KO 5423/1-1) and funding within the Cluster of Excellence PhoenixD (EXC 2122) both awarded to Prof. Dr. Carolin König. The LUIS scientific computing cluster and the cluster team have been instrumental in my research from day one. I address special thanks to Gizo and Gabriel for their extensive support. I thank Yvonne and Kerstin for their organizational support.

The fruitful collaboration with the Behrens group has been valuable in my work on the dispersion of refractive indices of ZIFs. I convey my thanks to Marvin Treger for the initiation of the collaboration, insightful discussions and provision of reference data. I thank Nils Keppler and Lukas Steinbach for supplying experimental reference data, and Andreas M. Schneider for supervision and productive discussions. I also acknowledge my internship students, Onno Strolka, Pascal Czember, and Daniel Bömke, for their contributions.

I am thankful to Christoph Jacob and Kevin Focke for their support in the PyADF framework and the implementation of the DensityEvaluator module, that were indispensable in the project on local embedding schemes.

My sincere gratitude goes to Toby for supervising the system administration project as well as for the enjoyable board game nights we had with Kevin.

I extend my appreciation to everyone who proofread this thesis, with special thanks to Janine, Anja, Erona, Nghia, Jonas, and Chris. I am particularly grateful to Anja for providing the excellent LaTeX template that saved me from many headaches.

I can't forget to acknowledge my outstanding HIWI KIDS, Pascal and Erona, for being wonderful to work with and for the fun times we've had. "Two and a half girls", definitely the winners of hearts in the beer pong tournament.

PASCAL, thank you for being both a visionary and the man of action, as well as for the memorable board game nights, concert and short trip adventures.

ERONA, you are my *shpirt*, taking understanding each other and taking care to a whole new level, and being the best travel, life, and diving partner one could wish for.

NGHIA, our friendship has been an exchange of so many things - from hobbies to cultural phenomena and delicious foods. The time when it was just the two of us in the office, you made having just one colleague incredibly worthwhile. Your calming presence is amazing to me, however, I must admit a hint of pride for introducing the card game that made us see you really mad for the first time. ;) I'll miss your sassy comments and *Schadenfreude*, that come out once you get to know someone.

JANINE, you've been more than a friend; you've been my counsellor and voice of reason during challenging times (and there have been some). I hope I've been able to reciprocate at least some of the extensive support you've given me. Your positivity and thoughtfulness have really made a significant impact on my life.

Many thanks to both of you, the *AK König originals*, for your understanding of my impulsiveness and mood swings, and for being a rock during the challenging times, particularly the harsh times of the pandemic. Our shared adventures, conversations, and memories hold a special place in my heart. I've learnt a lot from you. I'm glad I could also bring all those games a bit closer to you, even if bringing my kind of music closer didn't work out as well. :P Your presence has made time fly by a lot quicker, and I'll miss having you around every day. We've built the coolest small working group. :)

Moreover, I am grateful for the wonderful friendships I've made during my time here, in Münster and at summer/winter schools, you guys know who you are. :) A big shout-out goes to Sam, Elisa and Sofi who have been around forever, keep me sane and are there whenever I need them.

I'm deeply grateful for the unwavering support and love from my family throughout this journey and I want to thank them for being there for me every step of the way.

NUNIS, thank you for being the constant in my life. We share so many things and I won't list them all here now. You always bring laughter to our times together, even if I can't possibly ever watch all those reels you've sent me. Thank you for your perseverant support and patience with me and most of all for believing in me.



## Ausbildung

---

- 10/2020–  
heute      **Promotion in Chemie an der Leibniz-Universität Hannover**
- *Promotionsthema:* Subsystem-Based Methods for Global and Local Optical Properties
- Sommerschulen:
- *Molecular Response Summer School:* October 2021, Stockholm, Schweden
  - *European Summer School of Quantum Chemistry:* September 2022, Palermo, Italien
- Konferenzen:
- *Quantum Embedding Methods:* Mai 2021, Münster, Deutschland
  - *57th Symposium on Theoretical Chemistry:* September 2021, Würzburg, Deutschland
  - *Leibniz-University Chemistry Symposium:* September 2021, Hannover, Deutschland
  - *PhoenixD Retreat (geplant):* September 2023, Schneverdingen, Deutschland
  - *59th Symposium on Theoretical Chemistry (geplant):* September 2023, Zürich, Schweiz
- 02/2020–  
09/2020      **Promotion in Chemie an der Christian–Albrechts Universität Kiel**
- *Promotionsthema:* Subsystem-Based Methods for Global and Local Optical Properties
- 10/2018–  
02/2020      **Bachelorstudium in Physik an der Westfälische Wilhelms-Universität Münster**
- 10/2015–  
12/2017      **Master of Science in Chemie an der Westfälische Wilhelms-Universität Münster**
- *Masterarbeit:* NMR calculations for Co-factors of a photosynthetic reaction center
- 08/2016–  
02/2017      **Auslandsaufenthalt (Master) an der University of Melbourne, Australien**
- In der Gruppe von Prof. Dr. Lars Goerigk
- 10/2011–  
10/2015      **Bachelor of Science in Chemie an der Westfälische Wilhelms-Universität Münster**
- *Bachelorarbeit:* Untersuchung zur Modellierung analoger Schlussfolgerungsprozesse in der Reaktionsplanung
- 08/2002–  
07/2011      **Allgemeine Hochschulreife am Mariengymnasium Papenburg**
- 08/1998–  
06/2002      **Grundschule Lehe**

## Berufserfahrung

---

- 10/2020–  
08/2023      **Wissenschaftliche Mitarbeiterin an der Leibniz-Universität Hannover**
- 02/2020–  
09/2020      **Wissenschaftliche Mitarbeiterin an der Christian–Albrechts Universität Kiel**
- 10/2016–  
02/2017      **Studentische Hilfskraft an der Westfälische Wilhelms-Universität Münster**
- 10/2015–  
12/2017      **Studentische Hilfskraft an der Westfälische Wilhelms-Universität Münster**

## Fähigkeiten

---

- Sprachen:**
- Deutsch (Muttersprache)
  - Englisch (Verhandlungssicher)
  - Französisch (Fortgeschritten)
  - Spanisch (Grundkenntnisse)
- IT:**
- $\LaTeX$ (sehr gut)
  - Python (sehr gut)
  - Bash (sehr gut)
  - Ansible (sehr gut)
  - Git (gut)
  - Java (gut)
  - C++(grundlegend)
  - Dalton (sehr gut)
  - PyADF (sehr gut)
  - Turbomole (sehr gut)
  - AMS (sehr gut)
  - ORCA (sehr gut)
  - Gaussian (gut)
- Zertifikate:**
- Projektmanagement (2022)
  - Agiles Arbeiten (2022)
  - SCRUM (2022)
  - Teambildung und -führung (2022)
  - Unternehmensmanagement (2022)
  - Karriereentwicklung (2022)
  - CMake (2022)
  - Speed-Reading (2021)
  - Wissenschaftliches Schreiben (2020)
  - High-Performance Computing (2018)

## Publikationen

---

- 2017 Xu, J.; Zhang, B.; Jansen, M.; Goerigk, L.; Wong, W. W. H.; Ritchie, C. *Angew. Chem. Int. Ed.* **2017**, *56*, 13882.
- 2022 Jansen, M.; Thi Minh, N. N.; Hedegaard, E. D.; König, C., Quantum-derived embedding schemes for local excitations In *Chemical Modelling: Volume 17*, **2022**.
- 2023 Jansen, M.; Reinholdt, P.; Hedegård, E. D.; König, C. Theoretical and Numerical Comparison of Quantum- and Classical Embedding Models for Optical Spectra. *J. Phys. Chem. A* **2023**, *127*, 56895703.

## Präsentationen

---

- 07/2021 Dissecting the polarization contribution in explicit embedding models for local optical spectra. Leibniz-University Chemistry Symposium  
Hannover, Deutschland, **2021**
- 09/2021 Dissecting the polarization contribution in explicit embedding models for local optical spectra. 57th Symposium on Theoretical Chemistry  
Würzburg, Deutschland, September **2021**
- 09/2022 Dissecting the polarization contribution in explicit embedding models for local optical spectra. European Summer School of Quantum Chemistry  
Palermo, Italien, **2022**
- 09/2023 Theoretical Molecular Assembly Kit for the Calculation of Refractive Indices of Metal-Organic Frameworks. PhoenixD Retreat, *In Vorbereitung*.  
Schneverdingen, Deutschland, **2023**
- 09/2023 Theoretical Molecular Assembly Kit for the Calculation of Refractive Indices of Metal-Organic Frameworks. 59th Symposium on Theoretical Chemistry, *In Vorbereitung*.  
Zürich, Schweiz, **2023**