



PhD-FSTM-2024-002
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 31/01/2024 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN PHYSIQUE

by

Matej DITTE

Born on 25 July 1994 in Žiar nad Hronom (Slovakia)

MOLECULES IN ENVIRONMENTS: TOWARDS SYSTEMATIC QUANTUM EMBEDDING OF ELECTRONS AND DRUDE OSCILLATORS

Dissertation defence committee

Dr Alexandre Tkatchenko, dissertation supervisor
Professor, Université du Luxembourg

Dr *Massimiliano Esposito*, Chairman
Professor, Université du Luxembourg

Dr Rocco Martinazzo, Vice Chairman
Professor, University of Milan

Dr Lubos Mitas
Professor, North Carolina State University

Dr *Alex Thom*
Professor, University of Cambridge

Abstract

Biological, chemical and physical processes involve mechanisms that happen at different energetic, spatial and time scales. The interpretation of these phenomena requires the assistance of numerical calculations that help distinguish the origin and important properties observed in the experimental findings. Yet, in order to model such complexity, it is necessary to overcome the computational limits of even the most efficient high-performance computing facilities. For this purpose numerical methods based on the partitioning of the system of interest in different subspaces, each treated at a different level of theory and thus of accuracy are introduced. Such partitioning, also called *embedding*, allows to study large systems including for example solvation of large molecules and proteins.

In the most common of embedding approaches, the molecular fragment of interest is described via accurate *ab initio* quantum methods (QM), while the environment or the protein backbone is modelled via molecular mechanics (MM) with empirical Force-Fields (FF), or other coarse-grained methods, that by approximating the interactions, make the treatment of the entire system feasible. Unfortunately, these approximations often result in neglecting important intra and intermolecular interactions, especially those arising from polarization or dispersion effects. In this Thesis, we developed a new embedding method, in which molecular systems described through first-principle quantum methods are embedded in an accurate model environment, that explicitly includes important long-range dynamical correlation effects between the various sub-systems.

For the first time, we propose an explicitly quantum model environment, con-

structured via point charges and Coulomb interacting quantum Drude oscillators (QDO), *ie* charged quantum harmonic oscillators, designed to reproduce the electrostatics, polarization and dispersion interactions between the molecules of the environment and the electrons of an *ab initio* molecular sub-system in the non-covalent regime. We construct a single many-body Hamiltonian coupling the quantum electronic and drudonic degrees of freedom via Coulomb interactions and define a variational *ansatz* for the total system, which is integrated and optimized through quantum Monte Carlo (QMC) methods, which are able to capture the correlation effects within and between the two subsystems. In particular, we employ two of the most common QMC methods, namely variational (VMC) and diffusion Monte Carlo (DMC), which are implemented specifically to tackle these mixed systems of heterogeneous fermionic and distinguishable particles. This novel El-QDO embedding approach is first applied to describe dispersion interactions in a set of noble gas dimers, *ie* Ne, Ar, Kr and Xe. Afterwards, a more general QDO model for the water molecules is used to study the prototype hydrogen bond in the water dimer. Finally, the El-QDO method is applied to study the solvation end excitation energies of small molecules in a water environment, displaying remarkable agreement with accurate *ab initio* calculations. These first results prove the capability of the method to correctly describe the electrostatics, polarization, and dispersion effects within the QDO environment and between the QDO environment and the first-principle molecular sub-systems.

Finally, in the last application presented in this Thesis, we employ Density Functional Theory (DFT) together with an existing QDO-based dispersion method, *ie* the many-body dispersion (MBD) method, to study the thermochemistry of the diphenylmethyltelluronium cation interacting with the triphenylphosphine oxide in the ethylene dichloride solution, in order to interpret the experimental measurements of enthalpies obtained with Isotherm Titration Calorimetry (ITC). By comparing the results obtained with the MBD method with other pairwise dispersion methods, and accurate DMC and Coupled Cluster results, we again recognize the need for a correlated and computationally feasible embedding method like the El-QDO one, to describe molecular systems in complex environments including dynamical correlation effects that are crucial for the correct description of the supramolecular binding mechanisms in a solvent.

Preface

Acknowledgements

First of all, I would like to thank my supervisor, Prof Alexandre Tkatchenko, for giving me the opportunity to do the Ph.D. in his group, for his ideas, and for all the discussions we had during the last four years. I would also like to thank Dr. Matteo Barborini for all his help and especially for all the hours he spent by editing the text I wrote into a readable form; all the other members of the TCP group at the University of Luxembourg; and people who inspired me to study physics, namely: Dr. Matúš Dubecký, Ján Oravec, Martina Majzlíková, Ľubica Knoppová, and all the truly amazing lecturers at the Comenius University in Bratislava. Last but not least, I would like to thank my friends and family for all the support.

PhD publications

- [1] Molecules in Environments: Towards Systematic Quantum Embedding of Electrons and Drude Oscillators;
Matej Ditte, Matteo Barborini, Leonardo Medrano Sandonas, and Alexandre Tkatchenko;
Physical Review Letters **131**, 228001 (2023);
DOI: 10.1103/PhysRevLett.131.228001

- [2] Affinity of Telluronium Chalcogen Bond Donors for Lewis Bases in Solution: A Critical Experimental-Theoretical Joint study;
Loïc Gros Lambert, Yann Cornaton, Matej Ditte, Emmanuel Aubert, Patrick Pale, Alexandre Tkatchenko, Jean-Pierre Djukic, and Victor Mamane;
Chemistry - A European Journal e202302933 (2023);
DOI: 10.1002/chem.202302933
- [3] Quantum Drude Oscillators Coupled with Coulomb Potential as an Efficient Model for Bonded and Non-Covalent Interactions in Atomic Dimers;
Matej Ditte, Matteo Barborini, and Alexandre Tkatchenko;
in preparation
- [4] Electrons embedded in charged oscillators: A quantum embedding approach for molecular systems;
Matej Ditte, Matteo Barborini, and Alexandre Tkatchenko;
in preparation

Previous Publications

- [1] Toward Accurate Hydrogen Bonds by Scalable Quantum Monte Carlo;
Matúš Dubecký, Petr Jurečka, Lubos Mitas, Matej Ditte, and Roman Fanta;
Journal of Chemical Theory and Computation **15**, 3552 (2019);
DOI: 10.1021/acs.jctc.9b00096
- [2] Fractional Charge by Fixed-Node Diffusion Monte Carlo Method;
Matej Ditte and Matúš Dubecký;
Physical Review Letters **123**, 156402 (2019);
DOI: 10.1103/PhysRevLett.123.156402

Contents

List of Figures	xv
List of Tables	xix
List of Abbreviations	xxiii
List of Symbols and Notation	xxvii
1 Introduction	1
2 Electrons and Quantum Drude Oscillators	8
2.1 Schrödinger equation and the electronic Hamiltonian	9
2.2 Separation of scales	14
2.2.1 Force-Fields	19
2.3 Molecules in an environment and embedding methods	20
2.4 Quantum Drude Oscillators	23
2.4.1 Response properties of the QDO model	25
2.4.2 Applications of the QDO model	28
2.4.3 QDO model in the dipole approximation limit	28
2.5 El-QDO embedding	30
2.6 Damping of the Coulomb potential	32

2.7	Short-range repulsion	35
3	Computational methods	38
3.1	Variational Principle	40
3.2	Hartree-Fock theory	40
3.3	Density Functional Theory	44
3.4	Dispersion methods	47
3.4.1	D4 London-dispersion model	47
3.4.2	Tkatchenko-Scheffler method	48
3.4.3	Many-Body Dispersion method	49
3.5	Quantum Monte Carlo methods	50
3.5.1	Stochastic integration and the Metropolis-Hastings algorithm	51
3.5.2	Variational Monte Carlo	54
3.5.3	Wavefunction optimization methods	55
3.5.4	Diffusion Monte Carlo	56
3.6	Wavefunctions for many-electron and many-drudon systems	60
3.6.1	Electronic wavefunction	61
3.6.2	Drudonic wavefunction	64
3.6.3	Electron-drudon coupling	68
4	El-QDO applications	71
4.1	Noble gas dimers	72
4.2	Water dimers	76
4.3	Molecules in a water environment	80
4.3.1	Benzene in 4 water molecules	80
4.3.2	Ortho-benzyne in 4 and 30 water molecules	81
4.3.3	Benzene dimer in 50 water molecules	84

4.4	Computational details	85
4.4.1	Computational efficiency	85
4.4.2	Optimization of the wavefunctions	88
4.4.3	Time step $\delta\tau$ convergence	90
4.4.4	Effects of the damping functions	90
4.5	Conclusions	91
5	Binding mechanism of the QDO model	96
5.1	The homogeneous QDO dimer	97
5.2	Comparison between the homogeneous QDO dimer and the H ₂ molecule	98
5.3	Conclusions	104
6	Computational study of TeMePh₂⁺ interacting with OPh₃ in the DCE solvent	108
6.1	Computational details	111
6.2	Thermochemistry quantities	112
6.3	LED/DLPNO-CCSD(T) analysis	114
6.4	Thermochemistry results	117
6.4.1	Implicit solvation	119
6.4.2	Hybrid solvation	120
6.4.3	Vacuum and explicit solvation	121
6.5	Additional analysis of the pairwise vs many-body dispersion methods .	123
6.6	Conclusions	125
7	Summary and Outlook	128
	Appendices	156

A1	Parameterization of El-FF, QDO and El-QDO models, and supporting results	156
A2	Adducts of TeMePh ₂ ⁺ with OPPh ₃	160
A3	Additional results of TeMePh ₂ ⁺	163

List of Figures

1.1	Energy scale of the intramolecular versus intermolecular interactions . . .	2
2.1	Example of the potential energy surface	13
2.2	Intramolecular versus intermolecular interactions of the water dimer . .	15
2.3	Molecule in an environment	20
2.4	Schematic of the QDO model	25
2.5	QDO invariants	26
2.6	PES of the argon dimer	28
2.7	Schematic of the EI-QDO model	30
2.8	Mapping of a water cage onto QDOs	32
2.9	Damping functions	34
2.10	Uniform damping	35
3.1	Asymmetric potential of the drudons	65
4.1	Variants of the noble gas dimers	72
4.2	PES of the noble gas dimers without the short-range repulsion	73

4.3	PES of the noble gas dimers with the short-range repulsion	74
4.4	Short-range repulsion of the noble gas dimers	75
4.5	QDO-based model of the water molecule	76
4.6	Variants of the water dimers	77
4.7	PES of the water dimers without the short-range repulsion	78
4.8	PES of the water dimers with the short-range repulsion	79
4.9	Short-range repulsion of the water dimer	79
4.10	Water cages for benzene composed of 4 water molecules	80
4.11	Water cages for ortho-benzyne composed of 4 and 30 water molecules .	82
4.12	Water cage for the T-shaped benzene dimer and its monomers composed of 50 water molecules	84
4.13	Optimization of El-QDO ortho-benzyne (S) in 30 QDOs	88
4.14	Time step $\delta\tau$ convergence of El-QDO DMC for ortho-benzyne (S) in 30 QDOs	89
4.15	Ar ₂ QDO-QDO with normal damping	91
4.16	Ar ₂ El-QDO with normal damping	92
4.17	Ar ₂ El-QDO with uniform damping	93
5.1	Comparison of the <i>ansätze</i> for the QDO dimer	97
5.2	PES of the QDO dimer as a function of the parametrization	98
5.3	One-body observables of the QDO dimer	99
5.4	Two-body observables of the QDO dimer	101
5.5	QDO dimer parametrized for the H ₂ molecule	104
6.1	Sigma holes of TeMePh ₂ ⁺	109

6.2	TeMePh ₂ ⁺ , OPPh ₃ and DCE molecules	110
6.3	1:1 adduct of TeMePh ₂ ⁺ with OPPh ₃ for σ_1 in DCE	111
6.4	LED/DLPNO-CCSD(T) analysis of [(Ph ₃ PO)·TeMePh ₂] ⁺	116
A2.1	1:1 adduct of TeMePh ₂ ⁺ with OPPh ₃ for σ_1	160
A2.2	1:1 adduct of TeMePh ₂ ⁺ with OPPh ₃ for σ_2	160
A2.3	1:1 adduct of TeMePh ₂ ⁺ with OPPh ₃ for σ_3	160
A2.4	1:2 adduct of TeMePh ₂ ⁺ with OPPh ₃ for $\sigma_{1,2}$	161
A2.5	1:2 adduct of TeMePh ₂ ⁺ with OPPh ₃ for $\sigma_{1,3}$	161
A2.6	1:2 adduct of TeMePh ₂ ⁺ with OPPh ₃ for $\sigma_{2,3}$	161
A2.7	1:3 adduct of TeMePh ₂ ⁺ with OPPh ₃ for $\sigma_{1,2,3}$	162
A2.8	1:1 adduct of TeMePh ₂ ⁺ with OPPh ₃ for σ_2 in DCE	162
A2.9	1:1 adduct of TeMePh ₂ ⁺ with OPPh ₃ for σ_3 in DCE	162

List of Tables

4.1	Solvation energies of benzene	80
4.2	Decomposition of the El-FF solvation energies of benzene	81
4.3	Solvation energies of ortho-benzyne	81
4.4	Singlet-triplet excitation energies of ortho-benzyne in a water environment	82
4.5	Decomposition of the El-FF solvation energies of ortho-benzyne	83
4.6	Solvation energies of the T-shaped benzene dimer and its monomers . . .	83
4.7	Decomposition of the El-FF singlet-triplet excitation energies of ortho-benzyne in a water environment	85
4.8	Decomposition of the El-FF solvation energies of the T-shaped benzene dimer and its monomers	86
4.9	El-QDO DMC/VMC efficiency	86
4.10	El-QDO optimization efficiency	87
4.11	σ of the El-QDO wavefunctions of ortho-benzyne (S)	87
6.1	Time step $\delta\tau$ convergence of the DMC calculations of $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$	115
6.2	Thermochemistry (PBE-D4/MBD@rsSCS) of the association of OPPh_3 with TeMePh_2^+ in the implicit and hybrid solvent	118

6.3	Thermochemistry (PBE- \emptyset /D4/TS/MBD@rsSCS/MBD-NL, intermediate basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent	122
6.4	PBE-D4 vs PBE-MBD-NL analysis of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum	123
6.5	Te-O distances of [(Ph ₃ PO)·TeMePh ₂] ⁺	125
A1.1	Parametrization of the El-FF embedding	156
A1.2	QDO parametrization	157
A1.3	Parametrization of the short-range repulsion	158
A1.4	ECP error of the excitation energies of ortho-benzyne	159
A3.1	Thermochemistry (PBE- \emptyset /D4/TS/MBD@rsSCS/MBD-NL, light basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent	164
A3.2	Thermochemistry (PBE0- \emptyset /D4/TS/MBD@rsSCS/MBD-NL, light basis set) of the association of OPPh ₃ molecules with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent	165
A3.3	Single point binding energies (PBE- \emptyset /D4/TS/MBD@rsSCS/, light basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent	166
A3.4	Single point binding energies (PBE- \emptyset /D4/TS/MBD@rsSCS/, intermediate basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent	167
A3.5	Single point binding energies (PBE0- \emptyset /D4/TS/MBD@rsSCS/, light basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent	168
A3.6	Single point binding energies (PBE- \emptyset /D4/TS/MBD@rsSCS/, light basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent at the PBE-D4-COSMO geometries	169

A3.7	Single point binding energies (PBE- \emptyset /D4/TS/MBD@rsSCS/, intermediate basis set) of the association of OPPh ₃ molecules with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent at the PBE-D4-COSMO geometries	170
A3.8	Single point binding energies (PBE- \emptyset /D4/TS/MBD@rsSCS/, tight basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent at the PBE-D4-COSMO geometries	171
A3.9	Single point binding energies (PBE0- \emptyset /D4/TS/MBD@rsSCS/, light basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent at the PBE-D4-COSMO geometries	172
A3.10	Single point binding energies (PBE0- \emptyset /D4/TS/MBD@rsSCS/, intermediate basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent at the PBE-D4-COSMO geometries	173
A3.11	Single point binding energies (PBE0- \emptyset /D4/TS/MBD@rsSCS/, tight basis set) of the association of OPPh ₃ with TeMePh ₂ ⁺ in the vacuum and partial explicit solvent at the PBE-D4-COSMO geometries	174

List of Abbreviations

- CC** Coupled Cluster
- CCSD(T)** Coupled Cluster with singles-doubles and perturbative triples
- CI** Configuration Interaction
- D4** Grimme's dispersion method
- DFT** Density Functional Theory
- DLA** Determinant Locality Approximation
- DLPNO-CCSD(T)** domain-based local pair-natural orbital Coupled Cluster with singles-doubles and perturbative triples
- DMC** Diffusion Monte Carlo
- Dsp** dispersion
- ECP** Effective Core Potential
- EI-FF** Electrons - Force-Field
- EI-QDO** Electrons - Quantum Drude Oscillators
- Ex** exchange
- F-Ex** full minus the exchange
- FF** Force-Field
- FN** Fixed-Node
- GGA** Generalized Gradient Approximation

GTO Gaussian-type orbitals

HF Hartree-Fock

ITC Isotherm Titration Calorimetry

KS Kohn-Sham

LA Locality Approximation

LED local energy decomposition

MBD many-body dispersion

MBD-NL non-local version of the many-body dispersion

MBD@rsSCS range-separated self-consistent screening version of the many-body dispersion

MD Molecular Dynamics

MM Molecular Mechanics

PBE Perdew-Burke-Ernzerhof density functional

PBE0 Perdew-Burke-Ernzerhof density functional with mixture of Hartree-Fock exchange

PES Potential Energy Surface

PIMC Path Integral Monte Carlo

QCH Quantum Chemistry

QDO Quantum Drude Oscillator

QM Quantum Mechanics

QM:MM Quantum Mechanics in Molecular Mechanics

QM:QM Quantum Mechanics in Quantum Mechanics

QMC Quantum Monte Carlo

SAPT Symmetry-Adapted Perturbation Theory

ST singlet-triplet

STO Slater-type orbitals

TDSE time-dependent Schrödinger equation

TISE time-independent Schrödinger equation

TS Tkatchenko-Scheffler dispersion method

UPBE0 PBE0 with unrestricted Kohn-Sham orbitals

vdW van der Waals

VMC Variational Monte Carlo

ZORA Zeroth-Order Regular Approximation

List of Symbols and Notation

e	absolute value of the charge of an electron / Euler's constant
m_e	mass of an electron
k_e	Coulomb constant
\mathbf{r}^e	position of an electron
\mathbf{r}^d	position of a drudon
\mathbf{R}^n	position of a nucleus
\mathbf{R}^O	position of a center of QDO
\mathbf{R}^p	position of a point charge
Z	charge of a nucleus
q	charge of a QDO
Q	charge of a point charge
μ	mass of a drudon
ω	frequency of a QDO
\hat{H}^e	electronic Hamiltonian
\hat{H}^C	QDO Hamiltonian with the Coulomb potential
\hat{H}^{dip}	QDO Hamiltonian with the dipole potential
$\hat{H}^{\text{El-QDO}}$	El-QDO Hamiltonian
Ψ_e	wavefunction of an electronic system
Ψ_d	wavefunction of a QDO system
$\Psi_{\text{El-QDO}}$	wavefunction of an El-QDO system
E	energy
H	Enthalpy
G	Gibbs free energy

Introduction

Biological, chemical and physical processes, such as the adhesion properties of liquids on solid surfaces [174], the catalytic reactions in proteins or membranes [82], the solvation properties of molecules in solutes at different temperatures and pressure [23, 76], usually involve mechanisms that happen at different energetic, spatial and time scales. The interpretation of these complex processes, captured through a variety of spectroscopic techniques [131], often requires the assistance of numerical calculations in order to interpret the various findings and distinguish the origin and time evolution of the phenomena.

Yet, in order to model such complexity, overcoming the computational limits of the most efficient high-performance computational (HPC) facilities, numerical methods based on the partitioning of the system of interest into different subspaces, each treated with a different level of theory and thus a different accuracy have to be introduced. Usually, the partitioning of the systems is made on the basis of the different intra and intermolecular interactions that characterize the binding energies and thus the dynamics of the various sub-systems (Fig. 1.1). In particular, the intramolecular interactions, which include strong covalent or ionic bonds, are characterized by the sharing of electrons and thus display higher binding energies and vibrational frequencies, while the intermolecular interaction, such as electrostatic interactions, polarization, dispersion, or hydrogen bonds are all characterized by lower binding energies and lower vibrational frequencies.

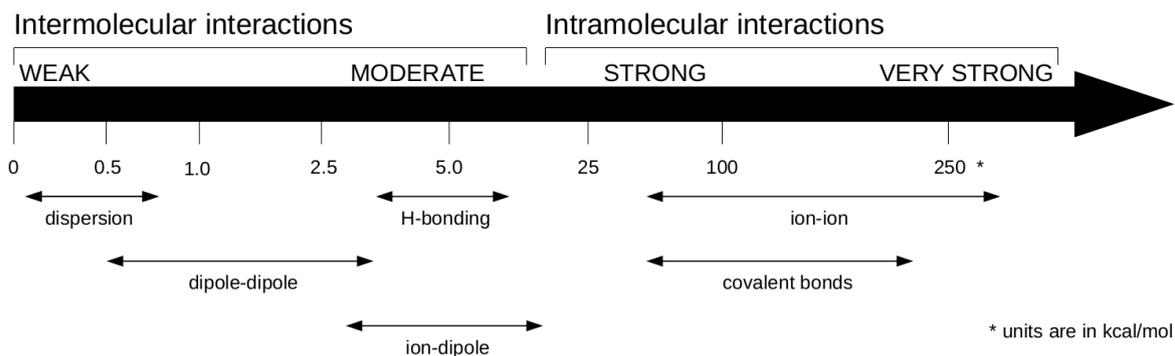


Figure 1.1: Energy scale of single inter- and intramolecular interactions between small molecules and atoms. The strength and order of the interactions can vary substantially in the case of larger and more complex systems.

This suggests that the systems can be partitioned into sub-systems based on the strength of their interactions, each treated with a different computational approach. These partitioning methods, usually referred to as *embedding approaches*, are based on a combination of different quantum mechanics (QM) methods, that explicitly tackle the description of the electronic Schrödinger equation, and molecular mechanics (MM) methods that approximate the interaction between atoms or molecular fragments via classical coarse-grained Force-Fields (FFs). While the QM methods, referred to as *ab initio*, try to give at least an approximate solution to the Schrödinger equation that describes the dynamics of the electrons as quantum particles, the classical FFs are built as approximations of the quantum environment, in which the intramolecular bonds are represented via harmonic springs that are tuned to reproduce the main vibrational motions of the molecules [199]. If the molecules present intrinsic dipoles, such as in the case of the water molecule, in FFs these are taken into account through classical point charges [92] that are essential to model the electrostatic intermolecular interactions. Furthermore, in order to recover intermolecular long-range dispersion, van der Waals (vdW) interactions or even hydrogen bonds, the FFs also include empirical vdW potentials [93] and in some cases explicit polarizable dipoles [200], that have the role of capturing the dynamical response of the environment.

Consequently, the standard embedding methods can be divided essentially into two groups. The first group is characterized by the joined usage of quantum mechanics and molecular mechanics (QM:MM) methods, that have been developed by the Nobel prize

winners, Karplus, Levitt, and Warshel amongst others [48]. These methods separate the systems into an environment, described via classical FF, and a target subsystem described at the quantum level, through *ab initio* methods usually Density Functional Theory (DFT) [177, 90]. These methods have been successfully applied to interpret or model a huge variety of phenomena, such as functional materials [101, 181], catalysis [212, 135, 134, 21], biochemical systems [61, 158, 3, 81, 7, 162], enzymes [11], proteins [124], DNA [151, 56] and also Raman spectroscopy [110, 140].

The second group is represented by the QM:QM methods, where the main fragment is embedded into the electron density of the environment, enabling charge transfer or charge delocalization between the two. In this group, we can include both Green's function methods for studying the spectroscopic and thermal quantities like dynamical mean-field theory [108] or self-energy embedding theory [111] and the wavefunction/density-based methods for studying ground state properties like density matrix embedding theory [105] or DFT embedding [203, 118, 122]. These QM:QM methods have been applied to a wide range of different systems and phenomena that include periodic graphene [96], catalysis [53, 159, 194, 83, 114, 34], spectroscopy [137, 85, 45, 46], defects [18], surfaces [113, 157], strongly correlated states in materials [121], perovskites [144, 206], complex oxides [138], nickelates [29] and variety of other real materials [139]. Here the two subsystems can be mutually polarized, usually in a self-consistent way.

While the more expensive QM:QM methods capture the mutual polarization of the main fragment and the environment more accurately than the QM:MM methods based on preparametrized FFs, the many-body dispersion effects are still approximated in both approaches. Some exceptions can be found in various perturbative approaches for QM:QM [183], but unfortunately these methods have a high computational cost.

For this reason, in this Thesis we propose a novel embedding method, named the QDO method, in which the main fragment described at the full electronic level is embedded into charged quantum harmonic oscillators, *ie* quantum Drude oscillators (QDOs) [197, 167, 88]. The QDO model, first applied by F. London to describe dispersion interactions between atoms and molecules [119], was derived from P. Drude's [39] classical model of absorption and refraction [40, 77] introduced in 1900. Each QDO consists of a single negatively charged quantum particle called drudon, oscillating around a positively charged classical center, bounded together via a quadratic potential well.

This coarse-grained system with only one quantum particle can mimic the response of the valence electrons of a real atom or molecule, for an appropriate parametrization deduced from first principles [88].

The QDO model is at the basis of the many-body dispersion (MBD) method [179, 9], in which the Coulomb interactions between oscillators are approximated by the first-order dipole-dipole interactions, leading to a quadratic Hamiltonian, which can be exactly diagonalized and used for the long-range many-body dispersion interactions in Quantum Chemistry methods such as DFT. On the other hand, many Coulomb interacting QDOs have been used as a coarse-grained model for atomic systems within the framework of diffusion Monte Carlo (DMC) and path integral Monte Carlo (PIMC) to study the dispersion interactions in noble gas dimers [88], solid [87] and fluid xenon [204], and applied to study the dynamics of liquid water [86, 166, 31, 89]. A different framework in which Coulomb interacting QDOs have been studied is that of the full Configuration Interaction (FCI) method, in which the oscillators are expanded in a basis set of Gaussian functions, and applied to prototype systems for dispersion interactions [152].

In this Thesis, we use for the first time the QDO model as an embedding environment for molecular systems described at the fully quantum level to study solvation properties of molecules in water, and to study how the environment affects molecular bindings and excitations. In the El-QDO method, we define a unique single many-body Hamiltonian describing the mixed system of electrons and the QDOs representing the environment [38]. The electronic and drudonic degrees of freedom are integrated on a collective variational *ansatz* within the framework of the quantum Monte Carlo (QMC) methods [52, 95, 14].

The QMC methods are an ensemble of stochastic methods used to integrate the time-independent Schrödinger equation over a chosen wavefunction approximating the system's ground state. In particular, in this Thesis, we use the variational (VMC) and the diffusion Monte Carlo (DMC) methods that we generalize to integrate the full El-QDO Hamiltonian on an *ansatz* introduced to recover all dominant dynamical correlation effects between the two types of quantum particles (electrons and drudons). Within the VMC framework, we use an energy minimization scheme to optimize the variational parameters of our approximate wavefunction. Afterwards, in order to im-

prove the estimation of the energy, we apply the DMC projection algorithm that is able to select the ground state component of a system from an approximate initial trial wavefunction.

Our novel procedure is applied to study the interaction energies of a set of four noble gas dimers (Ne_2 , Ar_2 , Kr_2 , Xe_2), computing the potential energy surfaces along the bond length and comparing the results with accurate Coupled Cluster (CC) references and with the corresponding symmetry-adapted perturbation theory (SAPT) decomposed energies [161]. Afterwards, we compute the binding energies of the water dimer in two different structural configurations, reconstructing the binding curves and again compare them with accurate references, in order to demonstrate the generalization of our approach beyond the dispersion-dominated noble gases dimers, to stronger interactions like the hydrogen bond. Finally, we apply the El-QDO approach to calculate the solvation energies of small molecules (benzene, benzene dimer, ortho-benzyne) in an environment composed of up to 50 water molecules, and the singlet-triplet adiabatic excitation energy of ortho-benzyne in a water environment. For these systems, we obtained excellent agreement with the reference *ab initio* calculations, while the comparison with a standard QM:MM method demonstrates the ability of our procedure to more accurately include correlation effects between the environment and electronic subsystem, that go beyond the approximate pairwise description of the dispersion of traditional FFs. We conclude that our El-QDO embedding is able to describe electrostatics, polarization and dispersion between the main fragment and the environment at the many-body level, for a cost comparable with calculations in the vacuum, as discussed in the text. The El-QDO embedding could replace the standard QM:MM approaches for highly correlated computational methods, which are necessary to capture important correlation effects between the environment and the electronic subsystem.

After verifying the efficiency of the El-QDO approach to model long-range interactions, we proceed to discuss the necessary implementations to describe the bonds at all possible interatomic distances. We proceed to construct the ‘nearly exact’ *ansatz* for the solution of the QDO model, with the goal of going beyond its limitations and tackling a wider range of interactions. By going beyond the model’s original purpose and understanding its behavior also at bond length around 0.5 \AA , we propose a generalized approach to explore its further extension, including the possibility of qualitatively

describing charge transfer effects, and paving the way towards a new generation of universal quantum Force-Fields [100], which is crucial for the generalization of the El-QDO embedding method to model a variety of different environments.

Finally, in the last application presented in this Thesis, we employ Density Functional Theory (DFT) together with the many-body dispersion (MBD) method, to study the thermochemistry of the diphenylmethyltelluronium cation interacting with the triphenylphosphine oxide in the ethylene dichloride solution, in order to interpret the experimental measurements of enthalpies obtained with Isotherm Titration Calorimetry (ITC). In particular, we analyze the role of the solvent and dispersion interactions in the description of the thermochemical quantities, namely enthalpies and Gibbs free energies, by comparing the results obtained with the MBD method to other pairwise dispersion methods, and accurate DMC and CC results. Through such comparison, we are able to reveal important correlation effects between the molecular conformers and between the conformers and the environment, again motivating the need for a correlated yet computationally feasible embedding method such as the El-QDO one, to describe molecular systems in large explicit molecular solvents.

The Thesis is structured as follows: in Chap. 2 we give a general introduction to embedding approaches and describe the El-QDO Hamiltonian, focusing on the details of the QDO model; in Chap. 3 we introduce the various computational methods used in this Thesis, with particular attention to the QMC methods, their generalizations and the construction of the El-QDO trial wavefunction; in Chap. 4 we discuss the main results of the El-QDO model, and in Chap. 5 we discuss the possible directions to further generalize the model; and finally, in Chap. 6 we discuss the computational results regarding the study of the diphenylmethyltelluronium cation interacting with the triphenylphosphine oxide in the ethylene dichloride solution.

Electrons and Quantum Drude Oscillators

Parts of Secs. 2.5, 2.6 and 2.7 in this Chapter have been published in this or a similar form in:

M. Ditte, M. Barborini, L. M. Sandonas, A. Tkatchenko. “Molecules in Environments: Toward Systematic Quantum Embedding of Electrons and Drude Oscillators” *Phys. Rev. Lett.* **131**, 228001 (2023)

and will be published in:

M. Ditte, M. Barborini, A. Tkatchenko. “Electrons embedded in charged oscillators: A quantum embedding approach for molecular systems” In preparation.

and in:

M. Ditte, M. Barborini, A. Tkatchenko. “Quantum Drude Oscillators Coupled with Coulomb Potential as an Efficient Model for Bonded and Non-Covalent Interactions in Atomic Dimers” In preparation.

and have been produced in collaboration with the above authors.

In this Chapter we discuss the main aspects of the electronic structure theory, with a focus on the separation of spatial and energetic scales for intermolecular interactions used for the construction of embedding methods. We introduce the quantum Drude oscillator (QDO) model for long-range interactions and the novel El-QDO embedding framework, in which the electronic system is embedded in a bath of QDOs. We describe the Hamiltonians of all the systems, the parameters of the model and technicalities like the damping of the Coulomb potential needed to avoid unphysical behaviors.

The Schrödinger equation, the Born-Oppenheimer approximation and the electronic Hamiltonian are defined in Sec. 2.1; in Sec. 2.2 we show the separation of scales for intermolecular interactions; in Sec. 2.3 we discuss the concept of the embedding methods; in Sec. 2.4 we introduce the QDO model for long-range interactions; in Sec. 2.5 we couple the QDO model with the electronic system, in order to develop the El-QDO embedding framework, in which the QDOs play the role of the environment; and finally in Secs. 2.6 and 2.7 we discuss the damping functions used to cure some possible divergences of the models, and the ways to introduce the short-range repulsion into the models, which is missing due to the missing exchange interactions.

2.1 Schrödinger equation and the electronic Hamiltonian

The time evolution of a non-relativistic quantum molecular system, containing N_n nuclei defined by charges and masses $\{Z_i, M_i\}_{i=1}^{N_n}$, and N_e electrons, is given by the time-dependent Schrödinger equation (TDSE)[155]

$$i \frac{\partial}{\partial t} \Psi(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t) = \hat{H}^{\text{mol}} \Psi(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t), \quad (2.1)$$

where the wavefunction Ψ contains the complete information about the system's state (equivalent to positions and momenta in classical mechanics). From here on and onward we use the Hartree atomic units ($\hbar = e = m_e = k_e = 1$) and define the $3N_e$ dimensional vector of the positions of the electrons and the $3N_n$ dimensional vector of the positions of the nuclei as $\bar{\mathbf{r}}^e$ and $\bar{\mathbf{R}}^n$ respectively. \hat{H}^{mol} is the Hamiltonian of the molecular

system, composed of 2 terms

$$\hat{H}^{\text{mol}} = \hat{T} + \hat{V}, \quad (2.2)$$

where the first one corresponds to the operator of the kinetic energy of the nuclei and the electrons

$$\hat{T} = \hat{T}_n + \hat{T}_e = - \sum_{a=1}^{N_n} \frac{1}{2M_a} \nabla_{\mathbf{R}_a}^2 - \sum_{i=1}^{N_e} \frac{1}{2} \nabla_{\mathbf{r}_i^e}^2, \quad (2.3)$$

and the second one to the Coulomb potential energy between all pairs of charged particles (nuclei-electrons, electrons-electrons and nuclei-nuclei)

$$\hat{V} = \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} = - \sum_{a=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_a}{|\mathbf{r}_i^e - \mathbf{R}_a^n|} + \sum_{i < j=1}^{N_e} \frac{1}{|\mathbf{r}_i^e - \mathbf{r}_j^e|} + \sum_{a < b=1}^{N_n} \frac{Z_a Z_b}{|\mathbf{R}_a^n - \mathbf{R}_b^n|}. \quad (2.4)$$

An initial state $\Psi(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t_0)$ can be thus evolved in time using the differential equation 2.1 with open boundary conditions (representing isolated molecular structures), and an expectation value of an observable, corresponding to an operator \hat{O} , can be computed as the integral

$$O(t) = \int \int d\bar{\mathbf{r}}^e d\bar{\mathbf{R}}^n \Psi^*(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t) \hat{O} \Psi(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t) = \langle \Psi(t) | \hat{O} | \Psi(t) \rangle, \quad (2.5)$$

where we introduced the so-called bra-ket notation and assumed the normalization condition $\int \int d\bar{\mathbf{r}}^e d\bar{\mathbf{R}}^n \Psi^*(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t) \Psi(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t) = \langle \Psi(t) | \Psi(t) \rangle = 1$.

The time integration of the TDSE is done by the separation of variables, which leads to

$$\Psi(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n, t) = \sum_j A_j e^{-iE_j^{\text{mol}} t} \Psi_j(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n), \quad (2.6)$$

where the coefficients A_j depend on the initial condition and E_j^{mol} together with $\Psi_j(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n)$ are the set of eigenvalues and corresponding eigenfunctions of the time-independent Schrödinger equation (TISE)

$$\hat{H}^{\text{mol}} \Psi_j(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n) = E_j^{\text{mol}} \Psi_j(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n). \quad (2.7)$$

Thus the complicated differential equation 2.1 is reduced to finding the eigenfunctions, corresponding to the stationary states of the TISE that form an orthonormal basis ($\langle \Psi_i | \Psi_j \rangle = \delta_{ij}$), with eigenvalues, that correspond to the energy levels of the continuous differential operator \hat{H}^{mol} . Despite this simplification, the problem 2.7 remains

analytically unsolvable and can only be tackled numerically in an approximate way. For this reason, additional simplifications are necessary, out of which the most common is the Born-Oppenheimer approximation [19].

The basic idea behind the Born-Oppenheimer approximation is to separate the nuclear and electronic degrees of freedom, inspired by the fact that the mass of a proton or a neutron is ≈ 2000 larger than the mass of an electron. For this reason, the motion of the nuclei is much slower than the motion of electrons and consequently, the latter can be considered as being adiabatic with respect to the first one. Adiabaticity in this context indicates that the electrons are instantaneously relaxed after each change of the positions of the nuclei. Without loss of generality, we can treat the nuclear coordinates $\bar{\mathbf{R}}^n$ as fixed parameters and write the time-independent Schrödinger equation for the electrons as

$$\hat{H}^e \psi_j(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = E_j^e(\bar{\mathbf{R}}^n) \psi_j(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n), \quad (2.8)$$

where the energies $E_j^e(\bar{\mathbf{R}}^n)$ now depend on the fixed position of the nuclei and the electronic Hamiltonian, which depends explicitly only on the electronic coordinates, only includes the terms

$$\hat{H}^e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}. \quad (2.9)$$

The eigenfunctions from Eq. 2.8 can be chosen to form an orthonormal basis ($\langle \psi_i(\bar{\mathbf{R}}^n) | \psi_j(\bar{\mathbf{R}}^n) \rangle = \delta_{ij}$) for all positions of the nuclei $\bar{\mathbf{R}}^n$ and thus the total wavefunction can be written as a linear combination

$$\Psi_j(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n) = \sum_k C_{jk}(\bar{\mathbf{R}}^n) \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n), \quad (2.10)$$

where the expansion coefficients $C_{jk}(\bar{\mathbf{R}}^n)$ are the wavefunctions of the j th roto-vibrational state of the nuclei with electrons in the electronic state k . This linear expansion can be substituted for $\Psi_j(\bar{\mathbf{r}}^e, \bar{\mathbf{R}}^n)$ in the total TISE 2.7

$$\left[\hat{T}_n + \hat{V}_{nn} + \hat{H}^e \right] \sum_k C_{jk}(\bar{\mathbf{R}}^n) \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = E_j^{\text{mol}} \sum_k C_{jk}(\bar{\mathbf{R}}^n) \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n). \quad (2.11)$$

The nuclear wavefunction $C_{jk}(\bar{\mathbf{R}}^n)$ can be now obtained by multiplying Eq. 2.11 by $\psi_l(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)$ from the left and integrating over the electronic coordinates. After some simple algebra, using the orthonormality condition, one obtains

$$\sum_k \langle \psi_l | \hat{T}_n | \psi_k \rangle C_{jk}(\bar{\mathbf{R}}^n) + \left[E_l(\bar{\mathbf{R}}^n) + \hat{V}_{nn} - E_j^{\text{mol}} \right] C_{jl}(\bar{\mathbf{R}}^n) = 0 \quad \forall l, \quad (2.12)$$

which coupled with Eq. 2.8 is equivalent to the original TISE of the total molecular Hamiltonian \hat{H}^{mol} . In order to introduce the Born-Oppenheimer approximation we have to have a closer look at the first term in Eq. 2.12, which contains the Laplacian applied to the product of the two functions

$$\begin{aligned} \hat{T}_n C_{jk}(\bar{\mathbf{R}}^n) \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = & - \sum_{a=1}^{N_n} \frac{1}{2M_a} \{ C_{jk}(\bar{\mathbf{R}}^n) \nabla_{\mathbf{R}_a}^2 \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) + \\ & + 2 [\nabla_{\mathbf{R}_a} C_{jk}(\bar{\mathbf{R}}^n)] \cdot [\nabla_{\mathbf{R}_a} \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)] + [\nabla_{\mathbf{R}_a}^2 C_{jk}(\bar{\mathbf{R}}^n)] \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) \}. \end{aligned} \quad (2.13)$$

In the Born-Oppenheimer approximation, the first two terms are assumed to be negligible since the electronic wavefunction varies slowly with respect to the nuclear coordinates $\bar{\mathbf{R}}^n$ and the Schrödinger equations for the electronic and nuclear degrees of freedom remain coupled only via the electronic energy levels $E_k^e(\bar{\mathbf{R}}^n)$ so that

$$\hat{H}^e \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = E_k^e(\bar{\mathbf{R}}^n) \psi_k(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) \quad (2.14)$$

and

$$\left[\hat{T}_n + \hat{V}_{nn} + E_k^e(\bar{\mathbf{R}}^n) \right] C_{jk}(\bar{\mathbf{R}}^n) = E_{jk}^{\text{mol}} C_{jk}(\bar{\mathbf{R}}^n). \quad (2.15)$$

Another simplification comes from the observation, that the electronic energy levels are often well separated. For example in the case of the hydrogen atom, the difference between the electronic energy of the lowest electronic state E_0 , *ie* the ground state, and the energy of the first excited state E_1 is 10.2 eV, which is equivalent to 118366 K. This energy difference is thus far from thermal fluctuations at normal conditions and the system can be considered to be always in the electronic ground state. An important exception to these conditions, for which the Born-Oppenheimer approximation is not valid, are phenomena related to optics, where a small subset of the excited states is also accessed during the dynamics. Yet, when the Born-Oppenheimer approximation is valid, it is also often possible to neglect the quantum nuclear effects, so that the nuclei $\bar{\mathbf{R}}^n$ can be considered as point charges and evolved according to classical Newtonian dynamics, where the internuclear potential is defined by the function $\hat{V}_{nn}(\bar{\mathbf{R}}^n) + E_0^e(\bar{\mathbf{R}}^n)$, which takes the name of potential energy surface (PES).

For the scope of this Thesis, which does not investigate the system's dynamics, we need only the time-independent Schrödinger equation for the electronic degrees of

freedom, where the positions of the nuclei $\bar{\mathbf{R}}^n$ are considered to be fixed parameters

$$\hat{H}^e \Psi_i(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = E_i(\bar{\mathbf{R}}^n) \Psi_i(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n), \quad (2.16)$$

with the redefined electronic Hamiltonian

$$\begin{aligned} \hat{H}^e &= \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} = \\ &= - \sum_{i=1}^{N_e} \frac{1}{2} \nabla_{\mathbf{r}_i^e}^2 - \sum_{i=1}^{N_e} \sum_{a=1}^{N_n} \frac{Z_a}{|\mathbf{r}_i^e - \mathbf{R}_a^n|} + \sum_{i<j=1}^{N_e} \frac{1}{|\mathbf{r}_i^e - \mathbf{r}_j^e|} + \sum_{a<b=1}^{N_n} \frac{Z_a Z_b}{|\mathbf{R}_a^n - \mathbf{R}_b^n|}, \end{aligned} \quad (2.17)$$

where we included also the Coulomb potential energy between the pairs of nuclei \hat{V}_{nn} (just a constant for given positions $\bar{\mathbf{R}}^n$), so that the energies $E_i(\bar{\mathbf{R}}^n)$ can be directly considered the PES of the system.

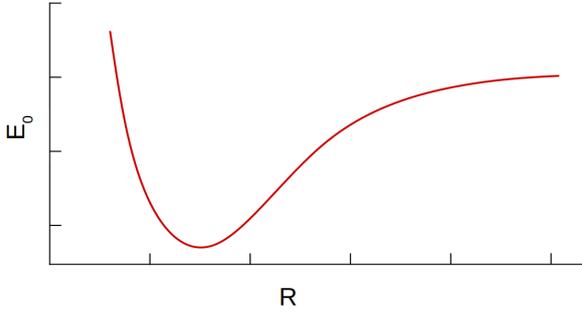


Figure 2.1: An example of the potential energy surface dependent on one coordinate (e.g. distance between two atoms in the case of a dimer composed of two atoms).

Our focus, in almost all the cases, except one specific application to excited states, is to find the electronic ground state energy $E_0(\bar{\mathbf{R}}^n)$ for a given fixed positions of the nuclei $\bar{\mathbf{R}}^n$. An example of the PES in a one-dimensional case, e.g. for a molecule composed of two atoms, where the coordinate is the distance between them, is shown in Fig. 2.1. It has normally a short-range repulsive region, a minimum corresponding to the optimal geometry at 0 K and a tail converging to the dissociation limit, as the separation

of the nuclei approaches infinity. We are not interested in dynamics, and thus we do not solve the Newton equation for the nuclei. In this Thesis, we use the PES only to obtain the optimal structures of molecules and study their binding energies and dissociation curves, where the ground state binding energy of a complex composed of the subsystems X and Y is defined as

$$\Delta E_0 = E_0^{XY} - E_0^X - E_0^Y, \quad (2.18)$$

with E_0^{XY} being the energy of the complex and E_0^X and E_0^Y being the energies of the isolated subsystems.

As discussed above, finding the electronic ground state energy of a molecule is a still highly nontrivial task and sophisticated numerical methods are needed. This will be further discussed in the Thesis. One of the main difficulties, responsible for the combinatorial complexity of the problem, is the condition for the wavefunction to be antisymmetric under the exchange of two fermions. This condition coming from Pauli exclusion principles imposes, that when two electrons of the same spin are exchanged in the wavefunction, then its value has to change the sign

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_a, \dots, \mathbf{r}_b, \dots, \mathbf{r}_{N_e}) = -\Psi(\mathbf{r}_1, \dots, \mathbf{r}_b, \dots, \mathbf{r}_a, \dots, \mathbf{r}_{N_e}) \quad \text{for } \sigma_a = \sigma_b, \quad (2.19)$$

while when the exchange is done for two electrons of different spin the sign must be preserved

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_a, \dots, \mathbf{r}_b, \dots, \mathbf{r}_{N_e}) = \Psi(\mathbf{r}_1, \dots, \mathbf{r}_b, \dots, \mathbf{r}_a, \dots, \mathbf{r}_{N_e}) \quad \text{for } \sigma_a \neq \sigma_b, \quad (2.20)$$

where σ denotes the eigenvalue of the \hat{S}_z operator of the electrons that can take values $1/2$ or $-1/2$ (or up and down), and we assume that the spin degrees of freedom are integrated out.

2.2 Separation of scales

The main focus of quantum chemists shifted during the last decade from the description of intramolecular interactions within individual molecules towards more challenging intermolecular interactions between molecules. The intermolecular interactions are in general challenging for two reasons: the total size of the system is larger and the accuracy needed to correctly describe phenomena that happen at the intermolecular scales is almost always higher.

An example of the difference between intra and intermolecular interactions is shown in Fig. 2.2 for the water dimer in its equilibrium geometry. Here we can distinguish between intramolecular O-H covalent bonds characterized by shared electrons between the atoms, and intermolecular H - - O non-covalent bond (more precisely so-called hydrogen bond in this case) between the two molecules, which does not involve shared electrons. The energy needed to break the O-H covalent bond of the

water molecule ($E_{H_2O} - E_{OH} - E_H$) is around 22 times higher than the energy needed to break the water dimer into two separated water molecules ($E_{(H_2O)_2} - 2E_{H_2O}$).

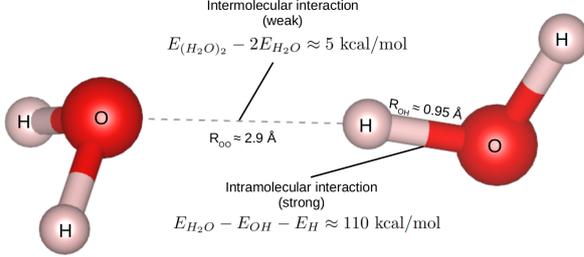


Figure 2.2: The difference between the intramolecular and intermolecular interactions on an example of the water dimer.

It means that if the goal is to calculate the two binding energies with an error up to 5%, one needs a 22 times more accurate method for the H - - O bond when compared to the O-H one. The length of the covalent O-H bonds (R_{OH}) within the water molecules is around three times shorter than the H - - O non-covalent bond between the two molecules (here measured by the R_{OO} distance). We thus observe a clear separation of both, the energetic and the spatial scale.

This separation of scales allows us to understand the behavior of the PES in limiting situations, when the distances between subsystems are large, and thus the interactions can be studied using the perturbation theory. The following is a derivation of the leading order intermolecular interactions and it can be found in more detailed form for example in Ref. 173. If we assume two interacting, but well-separated molecules A and B, the total Hamiltonian can be written as $\mathcal{H} = \mathcal{H}_A + \mathcal{H}_B + \mathcal{H}' = \mathcal{H}_0 + \mathcal{H}'$, where \mathcal{H}_A is the Hamiltonian for molecule A, \mathcal{H}_B is the Hamiltonian for molecule B, and all the cross Coulomb interactions between the pairs of charged particles from A and B are in Hamiltonian \mathcal{H}' , and it is assumed to be a small perturbation of the Hamiltonian \mathcal{H}_0 of two non-interacting molecules A and B. In this situation, we can formally write the Schrödinger equation for the two subsystems as $\mathcal{H}_A |m\rangle = W_m^A |m\rangle$ and $\mathcal{H}_B |n\rangle = W_n^B |n\rangle$, where we used the bra-ket notation with $|m\rangle$ being the m th state of the system A, and $|n\rangle$ being the n th state of the system B. The wavefunction of the total system AB, in the case of no interactions between A and B, and with the distance being sufficiently large so that the exchange interactions coming from the antisymmetry of the total wavefunction are negligible, can be partitioned as $\Psi_m^A \Psi_n^B = |mn\rangle$, giving the corresponding Schrödinger equation

$$\mathcal{H}_0 |mn\rangle = (\mathcal{H}_A + \mathcal{H}_B) |mn\rangle = (W_m^A + W_n^B) |mn\rangle = W_{mn}^0 |mn\rangle, \quad (2.21)$$

where W_{mn}^0 is the unperturbed energy with system A being in the state m and system

B being in the state n . Finally, if we assume non-degenerate ground states of the two subsystems (when the total spin of the system is zero), we can approximate the ground state energy of the total system W_{00} using the second-order perturbation theory as:

$$\begin{aligned} W_{00} &= W_{00}^0 + W_{00}^{(1)} + W_{00}^{(2)} = \\ &= W_{00}^0 + \langle 00 | \mathcal{H}' | 00 \rangle - \sum_{m+n \neq 0} \frac{\langle 00 | \mathcal{H}' | mn \rangle \langle mn | \mathcal{H}' | 00 \rangle}{W_{mn}^0 - W_{00}^0}. \end{aligned} \quad (2.22)$$

The first term in Eq. 2.22 is just the sum of the ground state energies of the two non-interacting systems A and B, and thus it does not depend on their distance.

The term $W_{00}^{(1)} = \langle 00 | \mathcal{H}' | 00 \rangle$, coming from the first-order perturbation, is the electrostatic interaction between the two subsystems, and it can be written as

$$\begin{aligned} U_{es} = \langle 00 | \mathcal{H}' | 00 \rangle &= Tq^A q^B + T_\alpha (q^A \langle \hat{\mu}_\alpha^B \rangle_0 - q^B \langle \hat{\mu}_\alpha^A \rangle_0) + \\ &+ T_{\alpha\beta} \left(\frac{1}{3} q^A \langle \hat{\Theta}_{\alpha\beta}^B \rangle_0 - \langle \hat{\mu}_\alpha^A \rangle_0 \langle \hat{\mu}_\alpha^B \rangle_0 + \frac{1}{3} q^B \langle \hat{\Theta}_{\alpha\beta}^A \rangle_0 \right) + \dots, \end{aligned} \quad (2.23)$$

where we used the Einstein summation convention (e.g. $x_\alpha y_\alpha = \sum_{i=1}^3 x_i y_i$) and the multipolar expansion of the cross Coulomb interactions between the charged particles from A and B, where the potential at point \mathbf{R}_B due to the molecule A can be expanded as

$$V^A(\mathbf{R}_B) = Tq^A - T_\alpha \hat{\mu}_\alpha^A + \frac{1}{3} T_{\alpha\beta} \hat{\Theta}_{\alpha\beta}^A - \dots + \frac{(-1)^n}{(2n-1)!!} T_{\alpha\beta\dots\nu}^{(n)} \xi_{\alpha\beta\dots\nu}^{A(n)} + \dots, \quad (2.24)$$

with the interaction tensor being

$$T_{\alpha\beta\dots\nu}^{(n)} = \nabla_\alpha \nabla_\beta \dots \nabla_\nu \frac{1}{R}, \quad (2.25)$$

and the operators q^A , $\hat{\mu}_\alpha^A$, $\hat{\Theta}_{\alpha\beta}^A$, etc. correspond to the zeroth-order moment (total charge of the molecule) $q^A = \sum_a q_a$, the first-order moment (dipole moment) $\hat{\mu}_\alpha^A = \sum_a q_a r_\alpha$, the second-order moment (quadrupole moment) $\frac{2}{3} \hat{\Theta}_{\alpha\beta}^A = \sum_a q_a r_\alpha r_\beta$ and so on. The electrostatic term has thus a very simple, classical interpretation: it is the interaction between the multipole moments of the isolated molecules. The condition for U_{es} being nonzero is that at least one of the multipole moments of both systems, A and B (not necessarily the moment of the same order) is nonzero. For example in the

case of two molecules that both have only nonzero dipole moment, the electrostatic contribution to the interaction energy is equal to the dipole coupling

$$U_{\mu\mu} = \frac{R^2 \boldsymbol{\mu}^A \cdot \boldsymbol{\mu}^B - 3(\boldsymbol{\mu}^A \cdot \mathbf{R})(\boldsymbol{\mu}^B \cdot \mathbf{R})}{R^5}. \quad (2.26)$$

The second-order contribution in Eq. 2.22 can be split into three terms as $W_{00}^{(2)} = U_{\text{ind}}^A + U_{\text{ind}}^B + U_{\text{disp}}$, where

$$U_{\text{ind}}^A = - \sum_{m \neq 0} \frac{\langle 00 | \mathcal{H}' | m0 \rangle \langle m0 | \mathcal{H}' | 00 \rangle}{W_m^A - W_0^A}, \quad (2.27)$$

$$U_{\text{ind}}^B = - \sum_{n \neq 0} \frac{\langle 00 | \mathcal{H}' | 0n \rangle \langle 0n | \mathcal{H}' | 00 \rangle}{W_n^B - W_0^B}, \quad (2.28)$$

and

$$U_{\text{disp}} = - \sum_{m \neq 0, n \neq 0} \frac{\langle 00 | \mathcal{H}' | mn \rangle \langle mn | \mathcal{H}' | 00 \rangle}{W_m^A + W_n^B - W_0^A - W_0^B}, \quad (2.29)$$

where the first two contributions are called induction (or polarization) and the third one is called dispersion.

The induction contribution U_{ind}^B (and similar for U_{ind}^A) can be written in a more understandable form as

$$U_{\text{ind}}^B = -\frac{1}{2} F_\alpha^A \alpha_{\alpha\alpha'}^B F_{\alpha'}^A - \frac{1}{3} F_\alpha^A A_{\alpha,\alpha'\beta'}^B F_{\alpha'\beta'}^A - \frac{1}{6} F_{\alpha\beta}^A C_{\alpha\beta,\alpha'\beta'}^B F_{\alpha'\beta'}^A - \dots, \quad (2.30)$$

where we introduced a notation for the electric field F (and its gradients) generated by the charge distribution of the molecule A and multipole polarizabilities $\alpha_{\alpha\alpha'}^B$, $A_{\alpha,\alpha'\beta'}^B$, $C_{\alpha\beta,\alpha'\beta'}^B$, etc. of the molecule B. For example, the dipole-dipole polarizability tensor of molecule B, which is a measure of the dipole moment of the molecule B induced by the constant electric field of the molecule A can be calculated as

$$\alpha_{\alpha\alpha'}^B = \sum_{n \neq 0} \frac{\langle 0 | \hat{\mu}_\alpha^B | n \rangle_B \langle n | \hat{\mu}_{\alpha'}^B | 0 \rangle_B + \langle 0 | \hat{\mu}_{\alpha'}^B | n \rangle_B \langle n | \hat{\mu}_\alpha^B | 0 \rangle_B}{W_n^B - W_0^B}. \quad (2.31)$$

The condition for a nonzero induction contribution to the total interaction energy is thus at least one nonzero multipole moment of at least one of the isolated molecules. It is possible to obtain also non-linear polarizability effects and hyperpolarizabilities,

when higher-order perturbation theory is used. The induction effects can be, up to some extent, mimicked via classical polarizable multipoles.

The last contribution U^{disp} is the dispersion term, which does not have any classical interpretation. If we consider two electrically neutral molecules and only the dipole-dipole coupling from the expansion of the Coulomb potential, the dispersion reduces to

$$U_{\text{disp}}^{(6)} = -T_{\alpha\beta}T_{\gamma\delta} \sum_{m \neq 0, n \neq 0} \frac{\langle 0 | \hat{\mu}_{\alpha}^A | m \rangle \langle 0 | \hat{\mu}_{\beta}^B | n \rangle \langle m | \hat{\mu}_{\gamma}^A | 0 \rangle \langle n | \hat{\mu}_{\delta}^B | 0 \rangle}{W_{m0}^A + W_{n0}^B} \quad (2.32)$$

labeled as $U_{\text{disp}}^{(6)}$, since it can be shown using the Casimir-Polder integral that

$$U_{\text{disp}}^{(6)} = -\frac{1}{R^6} \frac{3}{\pi} \int_0^{\infty} \alpha^A(i\omega) \alpha^B(i\omega) d\omega = -\frac{C_6}{R^6}, \quad (2.33)$$

ie it scales at the inverse of the 6th power of the intermolecular distance. The integral involves the dynamic polarizabilities $\alpha^A(\omega)$, which as a function of the real frequency ω measure the response of the molecule to the presence of a sinusoidally time-dependent external electric field. The dynamic polarizabilities in the imaginary frequencies $\alpha^A(i\omega)$ have no clear physical interpretation and are just a result of a mathematical trick in order to obtain the expression 2.33. It can be shown that when higher terms from the expansion of the Coulomb potential are included, one obtains

$$U_{\text{disp}} = - \sum_{n=6,8,10,\dots} \frac{C_n}{R^n}, \quad (2.34)$$

with for example $C_8 = C_{\mu,\Theta}^{AB} + C_{\Theta,\mu}^{AB}$, where for instance $C_{\mu,\Theta}^{AB} = \frac{15}{2\pi} \int_0^{\infty} \alpha^A(i\omega) \alpha^B(i\omega) d\omega$. The dispersion interactions can be understood as interactions between spontaneously induced multipole moments of the interacting molecules due to their quantum nature. The effects are non-additive and one can obtain also additional terms from higher-order perturbation theory, including the many-body dispersion with the leading order three-body term called Axilrod-Teller-Muto (ATM) [173].

The combination of dispersion interactions between 'instantaneously induced dipoles', induction interactions between permanent and induced dipoles and the electrostatic interactions between permanent dipoles is referred to as van der Waals (vdW) interactions.

Another important limit is when the two subsystems get closer to each other. The key role in this situation is played by the exchange interactions, coming from the antisymmetry of the total wavefunction and by the repulsion of the positively charged nuclei. One can show, that the contribution containing the exchange and the repulsion can be up to some extent approximated as the exponential function [173]

$$U_{\text{er}} = Ke^{-b(R-R_0)}, \quad (2.35)$$

where R is the intermolecular distance and K, b and R_0 are parameters depending on the specific system [199].

2.2.1 Force-Fields

The two limits described above are widely exploited to construct parametrized multidimensional functions used to approximate the potential energy surface $E_0(\bar{\mathbf{R}}^n)$ of large molecular compounds, without the necessity of explicitly tackling the Schrödinger equation. These approximate functions, *ie* Force-Fields (FFs), have physically inspired functional forms, and their parameters are fitted using a set of quantum mechanical calculations, often on smaller subsystems of the total system of interest [199]. Usually, the available FFs distinguish between covalently bonded and nonbonded pairs of nuclei, where each of them is treated differently due to the separation of scales. The easiest approximation used for the bonded pairs is a quadratic functional form $U_{\text{bonded}}(\mathbf{R}_i^n, \mathbf{R}_j^n) = A_{ij}(R_{ij} - B_{ij})^2$, where the parameters are fitted for the specific pairs of atoms. The nonbonded part, used also in many of the existing embedding methods discussed in Sec. 2.3, is often written as a sum of electrostatics, polarization and dispersion

$$U_{\text{nonbonded}} = U_{\text{elstat}} + U_{\text{pol}} + U_{\text{disp}}, \quad (2.36)$$

where U_{elstat} is the Coulomb potential between parametrized fractional charges of individual atoms, originating in the electrostatics coming from the first-order perturbation theory; U_{pol} , corresponding to the polarization coming from the second-order perturbation theory, can be obtained for example by using classical polarizable dipoles and FFs with non-zero polarization term are referred to as polarizable FFs; and finally the dispersion U_{disp} , which is almost always approximated using the pairwise $\propto -\frac{1}{R^6}$ dependency discussed before in this section. The short-range repulsion between the

nonbonded pairs is often included in the Lennard-Jones functional form of the U_{disp} term, as will be shown in Sec. 2.3.

Another application of the known scaling laws of the long-range intermolecular interactions are dispersion methods for low-cost computational methods of Quantum Chemistry, in which the dispersion contributions are completely missing or are included just partially [176]. Some of these methods, used for calculations in the Thesis, are discussed in Sec. 3.4.

2.3 Molecules in an environment and embedding methods

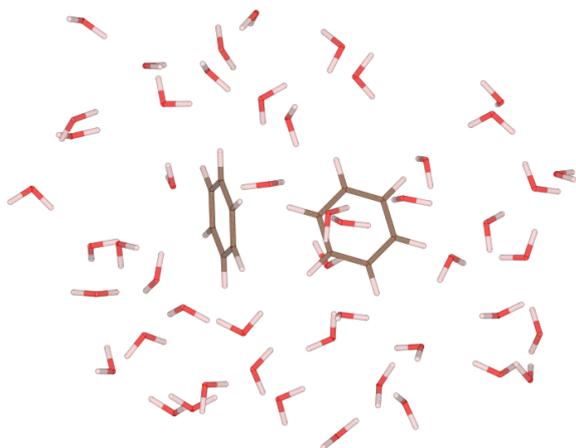


Figure 2.3: The benzene dimer in an environment composed of 50 water molecules.

One important category of systems of interest in Quantum Chemistry, considered also in this Thesis, are molecules in an environment, most often in a water solvent. These are especially important for the study of biological complexes [125], for drug development [33], and also for direct comparison of the computational methods to experiments, where the measurements are often not accessible in the vacuum [60]. An example of such a system is shown in Fig. 2.3, which contains the benzene dimer in an environment composed of 50 water molecules. The effects of the environment can often influence the prop-

erties (e.g. structure, binding/excitation energies) of the solvated molecules and this is, due to the typical distances between the two subsystems (*ie* the solute and the solvent), a consequence of the intermolecular interactions: electrostatics, polarization and dispersion.

This separation of scales between the molecules of interest and the environment is exploited in the so-called embedding methods [90], where the total system is split into the main fragment and the environment, each of them treated at a different level of theory. The embedding methods are usually divided into two groups: QM:MM (quantum mechanics in molecular mechanics), where the main fragment is described at the fully quantum level and the environment is approximated using classical Force-Fields (briefly described in Sec. 2.2); and QM:QM (quantum mechanics in quantum mechanics), where both subsystems are treated at the quantum level, but the environment is usually treated using a cheaper numerical method. The aim of the embedding methods is to reduce the computational cost, retaining the ability to model how the environment modifies the properties of the main fragment.

In the QM:MM methods, the degrees of freedom of the potential energy surface corresponding to the interactions between the nuclei of the environment are completely described via the ‘bonded’ and ‘nonbonded’ terms of a Force-Field [92, 93]. On the other hand, the interactions between the main electronic fragment and the FF environment are described only via the ‘nonbonded’ terms $U_{\text{elstat}} + U_{\text{pol}} + U_{\text{disp}}$. The first two are included directly in the electronic Hamiltonian of the main fragment and thus can polarize the electronic structure. In the case of polarizable FFs [63, 61] that include a polarization term, for example, via classical polarizable dipoles, the polarization of the environment in the presence of the main electronic fragment is obtained in a self-consistent way. The dispersion contribution between the main fragment and the environment is almost always approximated using a pairwise function with the long-range $\propto -\frac{1}{R^6}$ scaling law, dependent only on the distances between the nuclei of the two subsystems. The explicit many-body dispersion effects are thus not included in the QM:MM embedding because the electronic structure is not affected by the external U_{disp} function [38]. The interactions within the main fragment are governed by the full electronic Hamiltonian, the solution of which in theory gives the exact contribution to the PES for the nuclear degrees of freedom of the QM region.

To give the reader an example of the QM:MM embedding, we show a method based on the unpolarizable TIP3P model of the water molecule with fractional charges $q_O = -0.834$ and $q_H = 0.417$ [92]. These fractional charges are included in the QM:MM Hamiltonian $\hat{H}^{\text{QM:MM}}$, which is composed of the standard electronic Hamiltonian \hat{H}^e

for the main fragment from Eq. 2.17 and of extra Coulomb potentials between all the cross charged particles from the two subsystems

$$\begin{aligned} \hat{H}^{\text{QM:MM}} = \hat{H}^e + \sum_{i=1}^{N_e} \sum_{j=1}^{N_H} \frac{q_H}{|\mathbf{r}_i^e - \mathbf{R}_j^H|} + \sum_{i=1}^{N_e} \sum_{j=1}^{N_O} \frac{q_O}{|\mathbf{r}_i^e - \mathbf{R}_j^{Ox}|} + \\ + \sum_{i=1}^{N_n} \sum_{j=1}^{N_H} \frac{Z_i q_H}{|\mathbf{R}_i^n - \mathbf{R}_j^H|} + \sum_{i=1}^{N_n} \sum_{j=1}^{N_O} \frac{Z_i q_O}{|\mathbf{R}_i^n - \mathbf{R}_j^{Ox}|}, \end{aligned} \quad (2.37)$$

where \mathbf{R}_j^H and \mathbf{R}_j^{Ox} are fixed positions of the hydrogens and oxygens of the water molecules in the environment and N_H with N_O are respectively the number of hydrogen and oxygen atoms in the environment. The dispersion and the short-range repulsion contributions are included via the pairwise Lennard-Jones potential of the form

$$E_{\text{disp}}(R_{ij}) = \frac{A_{ij}}{R_{ij}^{12}} - \frac{C_{ij}}{R_{ij}^6}, \quad (2.38)$$

where R_{ij} are the distances between the cross pairs of nuclei from the two subsystems, and the parameters are obtained via combination rules $A_{ij} = \sqrt{A_i A_j}$ and $C_{ij} = \sqrt{C_i C_j}$ for distinct atoms, being $A_i = 4\epsilon_i \sigma_i^{12}$ and $C_i = 4\epsilon_i \sigma_i^6$, where ϵ and σ are from the more famous form of the Eq 2.38

$$E_{\text{LJ}}(R) = 4\epsilon \left[\left(\frac{\sigma}{R} \right)^{12} - \left(\frac{\sigma}{R} \right)^6 \right]. \quad (2.39)$$

The parametrization of Eq. 2.38 used in this work is taken from Ref. 93 (Tab A1.1 in Appendix A1) and was obtained to describe the dispersion interactions between the water molecules and solvated benzene, studied in this Thesis.

The point charges of the TIP3P model of water cause a polarization of the electronic system, but do not allow the environment to be polarized and the model includes only the external pairwise dispersion between the electronic system and the environment and within the environment. Although the first problem can be partially avoided using polarizable FFs in the QM:MM [61], the second problem persists. We will use the QM:MM method described above (later labeled as El-FF for consistency reasons) in order to show the importance of the many-body dispersion and mutual polarization effects in our novel El-QDO method.

In the second class of the embedding methods, the QM:QM category, in which all fragments are treated quantum mechanically, the main goal is to solve the Schrödinger

equation for the environment using a cheap computational method, to obtain an approximated charge density of the environment used in the QM Hamiltonian of the main fragment, so that the electrons and nuclei of the main fragment can interact with the charge density of the environment via the Coulomb potential. This approach describes with higher accuracy the polarization of the electronic structure of the main fragment if compared to the QM:MM approaches. The mutual polarization of the main fragment and the environment can be achieved by a self-consistent loop, where in the second step the charge density of the main fragment is used for the QM Hamiltonian of the environment. These iterative steps can be repeated until the equilibrium is reached. The dispersion interactions are often included via the classical pairwise terms coming from the FFs, already mentioned in the QM:MM part. Exceptions are more sophisticated and thus more expensive QM:QM methods, where the dispersion contribution between the two fragments is calculated in a perturbative way [183].

In order to overcome the disadvantages of the existing embedding methods coming from the various approximations of polarization and dispersion, we develop a new embedding approach, where the environment is represented via a quantum model for the long-range interactions without using perturbative approximations. The advantage of our novel approach when compared to the standard QM:MM and QM:QM methods are the presence of explicit electrostatics, mutual polarization and many-body dispersion, which are included in a non self-consistent and non-perturbative way; straightforward parametrization, which comes from the response properties of isolated molecules; and computational cost comparable to calculations in the vacuum. The method is based on the quantum Drude oscillator (QDO) model, which is introduced in the following section.

2.4 Quantum Drude Oscillators

The quantum Drude oscillator (QDO) is a model for long-range interactions, constructed from an approximation of the fluctuation-dissipation theorem [109], which defines the correlation energy of a system in terms of the electron density response function. In this model, the response properties of the valence electrons of a real electronic atom/molecule are mapped onto a single quantum particle, mimicking the

quantum fluctuations of the real charge density (but not the total charge density itself). The interactions between the fluctuating charge densities give rise to the effects behind intermolecular interactions, namely polarization and dispersion [88].

Each QDO consists of two particles that correspond to a classical *center* of charge $+q$ and to the *drudon*, a distinguishable quantum particle of charge $-q$ and mass μ . The center and the drudon interact via a quadratic potential

$$v_i(\mathbf{r}_i^d) = \frac{1}{2}\mu_i\omega_i^2|\mathbf{r}_i^d - \mathbf{R}_i^O|^2, \quad (2.40)$$

where \mathbf{R}_i^O is the fixed position of the center of the QDO, \mathbf{r}_i^d is the position of the quantum drudon and ω_i determines the slope of the quadratic well. A system of N_d QDOs is thus fully defined by a set of parameters $\{\mathbf{R}_i^O, q_i, \mu_i, \omega_i\}_{i=1}^{N_d}$, and if needed, in order to include the electrostatic effects via the static dipole moment of isolated polar molecules (e.g. water molecule) into the QDO model, additional N_p point charges, defined by positions, charges and indices of the parental QDO $\{\mathbf{R}_i^p, Q_i, p_i\}_{i=1}^{N_p}$, have to be introduced.

The Hamiltonian of a system of N_d interacting QDOs and N_p point charges has the form

$$\begin{aligned} \hat{H}^d = & \sum_{i=1}^{N_d} \hat{h}_i^d(\mathbf{r}_i^d) + \sum_{i<j=1}^{N_d} \frac{q_i q_j}{|\mathbf{R}_i^O - \mathbf{R}_j^O|} + \sum_{i<j=1}^{N_d} \frac{q_i q_j}{|\mathbf{r}_i^d - \mathbf{r}_j^d|} + \\ & + \sum_{i=1}^{N_d} \sum_{\substack{j=1 \\ p_j \neq i}}^{N_p} \frac{q_i Q_j}{|\mathbf{R}_i^O - \mathbf{R}_j^p|} + \sum_{\substack{i<j=1 \\ p_i \neq p_j}}^{N_p} \frac{Q_i Q_j}{|\mathbf{R}_i^p - \mathbf{R}_j^p|}, \end{aligned} \quad (2.41)$$

with one-body operator

$$\hat{h}_i^d(\mathbf{r}_i^d) = -\frac{1}{2\mu_i}\nabla_{\mathbf{r}_i^d}^2 + \frac{1}{2}\mu_i\omega_i^2|\mathbf{r}_i^d - \mathbf{R}_i^O|^2 - \sum_{\substack{j=1 \\ j \neq i}}^{N_d} \frac{q_i q_j}{|\mathbf{r}_i^d - \mathbf{R}_j^O|} - \sum_{\substack{j=1 \\ p_j \neq i}}^{N_p} \frac{q_i Q_j}{|\mathbf{r}_i^d - \mathbf{R}_j^p|}. \quad (2.42)$$

Each drudon interacts via the Coulomb potential with all the other charged particles in the system, except for its center, where it feels the quadratic attraction instead, and except for the point charges belonging to its parental QDO, where the interaction is omitted (thus we need to define the indices of the parental QDO mentioned above).

A schematic of the Hamiltonian 2.41 for two QDOs and one point charge is shown in Fig. 2.4.

The Hamiltonian of a single QDO without the presence of additional point charges is one of the very few examples, for which the exact solution is known. The wavefunctions in spherical coordinates r, θ, ϕ labeled by quantum numbers k, l, m can be written as

$$\psi_{klm}(r, \theta, \phi) = N_{kl} r^l e^{-\nu r^2} L_k^{(l+\frac{l}{2})} (2\nu r^2) Y_{lm}(\theta, \phi), \quad (2.43)$$

where $N_{kl} = \sqrt{\sqrt{\frac{2\nu^3}{\pi} \frac{2^{k+2l+3} k! l!}{(2k+2l+1)!!}}$ is a normalization factor, $\nu = \frac{\mu\omega}{2}$, $L_k^{(l+\frac{l}{2})}$ are generalized

Laguerre polynomials and Y_{lm} are spherical harmonic functions. The corresponding energy for the state with quantum numbers k, l, m is

$$E_{k,l,m} = \omega \left(2k + l + \frac{3}{2} \right). \quad (2.44)$$

In the case of the ground state, this simplifies to a Gaussian function

$$\psi_0(r) = \left(\frac{\sqrt{\mu\omega}}{\pi} \right)^{\frac{3}{4}} e^{-\sqrt{\mu\omega} r^2}, \quad (2.45)$$

with the ground state energy $E_0 = \frac{3}{2}\omega$. The solution of a system of two or more interacting QDOs with the Coulomb potential is analytically unsolvable and thus numerical methods are needed.

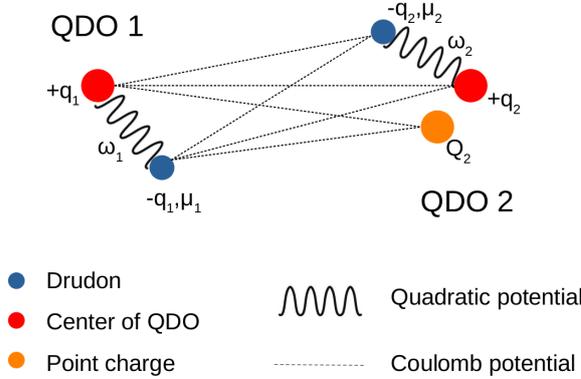


Figure 2.4: A schematic of the QDO model described by the Hamiltonian in Eq. 2.41 in the case of two interacting QDOs, where QDO 2 has one additional point charge.

2.4.1 Response properties of the QDO model

It can be shown from the second-order perturbation theory, as done in Ref. 88, that the model possesses multipole polarizabilities measuring the response of the QDO to the presence of an external electric field, which can be expressed in the SI units as

$$\alpha_l = \left(\frac{q^2}{\mu\omega^2} \right) \left[\frac{(2l-1)!!}{l} \right] \left(\frac{\hbar}{2\mu\omega} \right)^{l-1}, \quad (2.46)$$

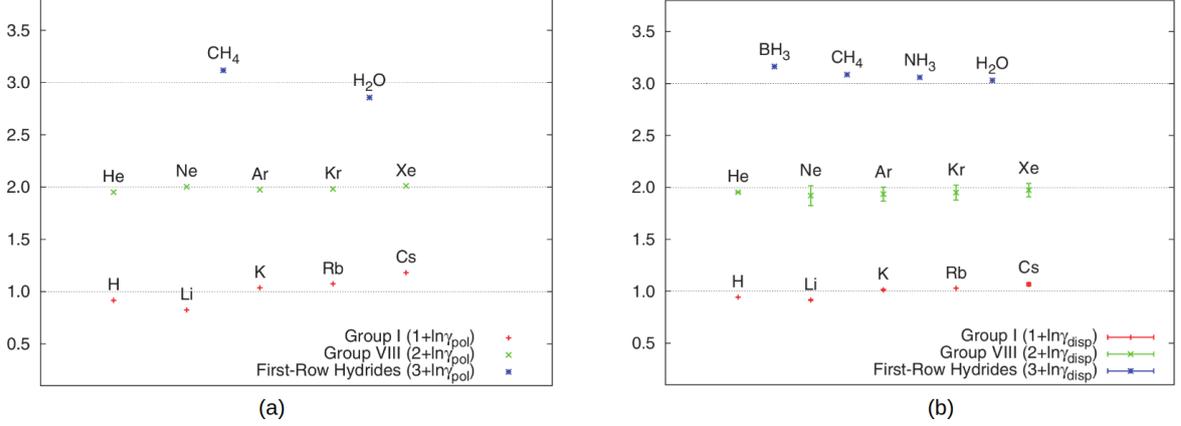


Figure 2.5: (a) The polarizability invariant γ_{pol} for alkali-metal and noble gas atoms and small molecules; and (b) the dispersion invariant γ_{disp} for alkali-metal, noble gas and small molecule dimers. The figures are taken from Ref. 88.

with α_1 being the dipole polarizability, α_2 being the quadrupole polarizability, α_3 being the octupole polarizability, etc. In general, for nonsymmetric molecules, the polarizability has to be treated as a tensor, but it is always diagonal in the case of the spherically symmetric QDO model. The model possesses also multipole hyperpolarizabilities, out of which the first nonzero one is dipole-dipole-quadrupole, again due to the spherical symmetry of the model. Interestingly, the classical Drude model [40] contains the same dipole polarizability α_1 , which corresponds to Eq. 2.46, with $\alpha_l = 0$ for $l > 1$. As can be seen from Eq. 2.46, the polarizabilities of the QDO model are not independent and can be all expressed via α_1 as a consequence of the simplicity of the wavefunctions from Eq. 2.43. After a very basic algebra, one can obtain an invariant for the first three coefficients [88]

$$\gamma_{\text{pol}} = \frac{2\sqrt{5}\alpha_2}{3\sqrt{\alpha_1\alpha_3}} = 1, \quad (2.47)$$

which is independent of the parameters q, μ, ω and thus it holds for every QDO. Surprisingly, the invariant in Eq. 2.47 holds also for simple real electronic systems, as shown in Fig. 2.5a. This implies that when the parameterization of the QDO model is done using the leading order polarization coefficients, the higher order coefficients are reasonably approximated, without the necessity of increasing the number of parameters of the model.

From the second-order perturbation theory, it can be also shown that the model can describe dispersion interactions [88]. For simplicity here we discuss only the two-body interactions arising from the second-order perturbation

$$E_{\text{disp}} = - \sum_{n=6,8,10,\dots} \frac{C_n^{AB}}{R_{AB}^n}, \quad (2.48)$$

but it is possible to obtain also many-body terms of all orders from higher-order perturbation theory. The dispersion coefficients C_i , previously introduced in Sec. 2.2 for electronic systems, can be expressed analytically as functions of the parameterization [88]

$$C_6^{AB} = \frac{3}{2} \alpha_1^A \alpha_1^B \frac{\hbar \omega_A \omega_B}{(\omega_A + \omega_B)}, \quad (2.49)$$

$$C_8^{AB} = \frac{15}{2} \left[\alpha_1^A \alpha_2^B \frac{\hbar \omega_A \omega_B}{(\omega_A + 2\omega_B)} + \alpha_2^A \alpha_1^B \frac{\hbar \omega_A \omega_B}{(2\omega_A + \omega_B)} \right], \quad (2.50)$$

$$C_{10}^{AB} = \left[21 \alpha_1^A \alpha_3^B \frac{\hbar \omega_A \omega_B}{(\omega_A + 3\omega_B)} + 21 \alpha_3^A \alpha_1^B \frac{\hbar \omega_A \omega_B}{(3\omega_A + \omega_B)} + 70 \alpha_2^A \alpha_2^B \frac{\hbar \omega_A \omega_B}{(2\omega_A + 2\omega_B)} \right], \quad (2.51)$$

again with invariants due to the simplicity of the model, for example [88]

$$\gamma_{\text{disp}} = \frac{7}{2\sqrt{10}} \frac{C_8}{\sqrt{C_6 C_{10}}} \quad (2.52)$$

in the case of $A=B$ with $C_i^{AA} = C_i$. This invariant also holds quite well for simple real systems as shown in Fig. 2.5b. A similar invariant quantity can be obtained also for the case $A \neq B$ and the three-body Axilrod-Teller-Muto coefficients, again with good agreement for simple real atoms/molecules [88].

The QDO parameters considered in this work, reported in Appendix A1 in Tab. A1.2, are taken from Refs. 88, 31. They were obtained by reproducing the leading order polarizabilities (α_i) and dispersion coefficients (C_i). This parametrization is not unique and we have chosen a specific one with the relations $\omega = \frac{1}{\hbar} \frac{4C_6}{3\alpha_1^2}$, $\mu = \frac{\hbar}{\omega} \frac{5C_6}{C_8}$ and $q = \sqrt{\mu\omega^2\alpha_1}$ [88].

It is important to underline here that drudons are considered to be distinguishable spin-less particles, and thus the repulsive energy between QDOs that comes mainly from the exchange energy is missing in the model. This problem is further discussed in Sec. 2.7.

2.4.2 Applications of the QDO model

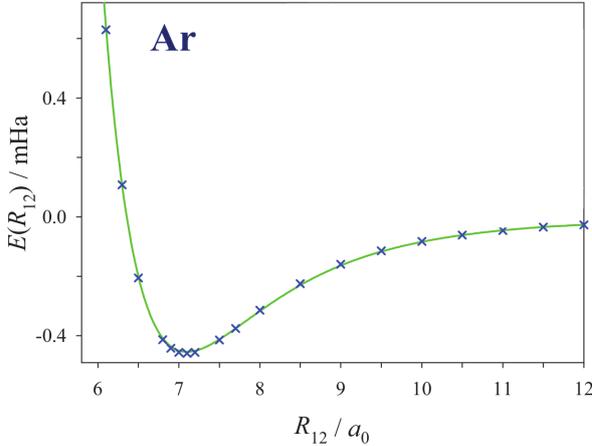


Figure 2.6: The dissociation curve of the argon dimer. The full line is a reference empirical potential energy surface and the crosses are obtained using the QDO model via diffusion Monte Carlo calculations with an external potential for the short-range repulsion, that is missing in the model. The figure is taken from Ref. 88.

universal pairwise interatomic van der Waals potentials [100].

The QDO model has been successfully applied to noble gas systems and water clusters using full Configuration Interaction, diffusion Monte Carlo and path integral Monte Carlo computation methods [88, 87, 86, 166, 31, 89, 152]. One of these applications is shown in Fig. 2.6 taken from Ref. 88, where the authors reproduce the binding curve of the argon dimer using the QDO model, containing only two quantum particles instead of the 36 electrons in the fully electronic case.

The model was recently used also in the derivation of a quantum scaling law between the size of the system represented via the van der Waals radius and the polarizability of the system, with $\alpha \propto R_{\text{vdW}}^7$ [47] and for the construction of uni-

2.4.3 QDO model in the dipole approximation limit

The Hamiltonian from Eq. 2.41 can be approximated using the dipole potential instead of the Coulomb potential when the distances between the QDOs are large enough. In the specific case of two interacting QDOs with equal parametrization, this approximate dipole Hamiltonian has the form

$$\hat{H}^{\text{dip}} = \sum_{i=1}^2 \left[-\frac{1}{2\mu} \nabla_{\mathbf{r}_i^d}^2 + \frac{1}{2} \mu \omega^2 |\mathbf{r}_i^d - \mathbf{R}_i^O|^2 \right] + V^{\text{dip}}(\mathbf{r}_1^d, \mathbf{r}_2^d), \quad (2.53)$$

with the two-body dipole coupling

$$V^{\text{dip}}(\mathbf{r}_1^d, \mathbf{r}_2^d) = \frac{q^2}{R^5} [R^2 (\mathbf{d}_1 \cdot \mathbf{d}_2) - 3 (\mathbf{d}_1 \cdot \mathbf{R}) (\mathbf{d}_2 \cdot \mathbf{R})], \quad (2.54)$$

where $\mathbf{d}_i = \mathbf{r}_i^d - \mathbf{R}_i^O$, $\mathbf{R} = \mathbf{R}_1^O - \mathbf{R}_2^O$ and $R = |\mathbf{R}|$. This quadratic Hamiltonian can be diagonalized analytically and the resulting interaction energy $E^{\text{dip}}(R)$ as a function of the distance between the QDOs is

$$E^{\text{dip}}(R) = E^+(R) + E^-(R) - E_0, \quad (2.55)$$

where

$$E^\pm(R) = \frac{\omega}{2} \left(2\sqrt{1 \pm \frac{q^2}{\mu\omega^2 R^3}} + \sqrt{1 \mp \frac{2q^2}{\mu\omega^2 R^3}} \right), \quad (2.56)$$

and $E_0 = 2 \times \frac{3}{2}\omega$ is the energy of two non-interacting QDOs. The dipole Hamiltonian from Eq. 2.53 can be easily generalized for N_d non-homogeneous interacting QDOs, again leading to a quadratic Hamiltonian, which can be diagonalized numerically. The specific case of two homogeneous QDOs in Eq. 2.53 is sufficient for the scope of this work.

The ground state wavefunction of the of Hamiltonian 2.53 has the form

$$\Psi^{\text{dip}}(\mathbf{r}_1^d, \mathbf{r}_2^d) = N \exp[\bar{\mathbf{r}}_{dO}^\top \mathbf{A} \bar{\mathbf{r}}_{dO}], \quad (2.57)$$

where N is a normalization factor, $\bar{\mathbf{r}}_{dO} = \bar{\mathbf{r}}^d - \bar{\mathbf{R}}^O$ is a 6-dimensional vector containing distances between each drudon and its center, and \mathbf{A} is a matrix containing coupling coefficients which depend on the parametrization of the QDOs and which can be obtained via diagonalization.

The dipole approximation of the QDO Hamiltonian has an important role in the so-called many-body dispersion method [179, 9], which is used to reintroduce dispersion in computational methods lacking the correct description of the intermolecular interactions such as Density Functional Theory. This method is discussed in more detail in Sec. 3.4.3.

2.5 El-QDO embedding

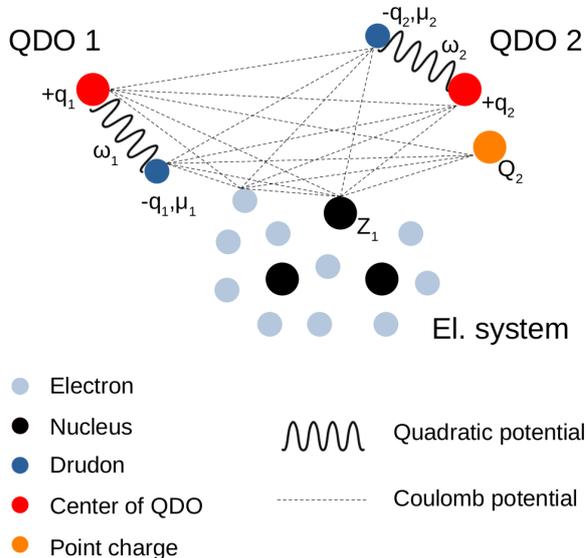


Figure 2.7: A schematic of the El-QDO model described by the Hamiltonian in Eq. 2.58 in the case of an electronic system interacting with two interacting QDOs, where QDO 2 has one additional point charge. Only selected interactions between the QDOs and the electronic system are shown.

This section focuses on the use of the QDO model introduced in Sec. 2.4 as an embedding bath for electronic systems in a novel embedding framework named El-QDO. As mentioned in the introduction and in Sec. 2.3, the goal is to exploit the range separations of molecules in a solvent, where we can clearly distinguish between the interactions within the main fragment, interaction within the solvent and interactions between the solvent and the main fragment (Fig. 2.3). Here for the first time we use the QDO model to represent the molecules in the environment. We construct a single many-body Hamiltonian for the two coupled subsystems, and later in Secs. 3.5 and 3.6.3 we propose a solution of the mathematically well-defined problem using quantum Monte Carlo methods.

The motivation for the El-QDO embedding are the response properties of the QDO model. As shown in Sec. 2.4, the model can be parametrized to correctly describe the leading order responses, namely the polarizabilities and dispersion coefficients of real matter. This in fact means, that when a QDO interacts with an electronic system instead of another QDO, the polarization and dispersion interactions should be captured correctly, exactly as in the case of two or more interacting QDOs. One could in principle show the presence of polarization and dispersion between the two subsystems using the second-order perturbation theory, but this is not within the scope of this work. Instead, we show the presence of these interactions by analyzing the numerical results in Chap. 4.

Up to our knowledge, the QDO model had not been used for an embedding method in the existing literature prior to the publications related to this Thesis. The only example of a single electron coupled with QDOs can be found in Refs. 167, 197, where the authors studied a dipole-bound anion via one-electron model potentials. Yet, the model, levels of approximations, and applications are very different from the scope of this Thesis.

The environment of QDOs and point charges that represent a classical Force-Field reproducing the static dipole moment of polar molecules can be thus, as discussed above, used as a bath for the sub-system of electrons, able to reproduce the quantum effects responsible for the long-range interaction in real matter. An electronic system containing N_n nuclei defined by positions and charges $\{\mathbf{R}_i^n, Z_i\}_{i=1}^{N_n}$, and N_e electrons can be embedded into a bath of N_d QDOs with parameters $\{\mathbf{R}_i^O, q_i, \mu_i, \omega_i\}_{i=1}^{N_d}$ and N_p point charges with parameters $\{\mathbf{R}_i^p, Q_i, p_i\}_{i=1}^{N_p}$, leading to the total El-QDO Hamiltonian of the form

$$\hat{H}^{\text{El-QDO}} = \hat{H}^e + \hat{H}^d + \hat{V}_{\text{int}}^{d-e}, \quad (2.58)$$

where the first term \hat{H}^e represents the standard electronic Hamiltonian within the Born-Oppenheimer approximation from Eq. 2.17, \hat{H}^d is the Hamiltonian of interacting QDOs and point charges from Eq. 2.41 and the last term

$$\begin{aligned} \hat{V}_{\text{int}}^{d-e} = & \sum_{i=1}^{N_e} \sum_{j=1}^{N_d} \left(\frac{q_j}{|\mathbf{r}_i^e - \mathbf{r}_j^d|} - \frac{q_j}{|\mathbf{r}_i^e - \mathbf{R}_j^O|} \right) + \\ & + \sum_{i=1}^{N_n} \sum_{j=1}^{N_d} \left(\frac{Z_i q_j}{|\mathbf{R}_i^n - \mathbf{R}_j^O|} - \frac{Z_i q_j}{|\mathbf{R}_i^n - \mathbf{r}_j^d|} \right) - \\ & - \sum_{i=1}^{N_e} \sum_{j=1}^{N_p} \frac{Q_j}{|\mathbf{r}_i^e - \mathbf{R}_j^p|} + \sum_{i=1}^{N_n} \sum_{j=1}^{N_p} \frac{Z_i Q_j}{|\mathbf{R}_i^n - \mathbf{R}_j^p|} \quad (2.59) \end{aligned}$$

represents the Coulomb interactions between the two subsystems, atoms/molecules and the QDOs with the point charges. A scheme of the interactions in the Hamiltonian in Eq. 2.58 is shown in Fig. 2.7, and an application of the El-QDO to the benzene dimer in an environment composed of 50 water molecules replaced by QDOs is shown in Fig. 2.8.

The Hamiltonian in Eq. 2.58 allows both, the main fragment and the environment, to polarize themselves: *ie* the electrons polarize in the presence of the environment

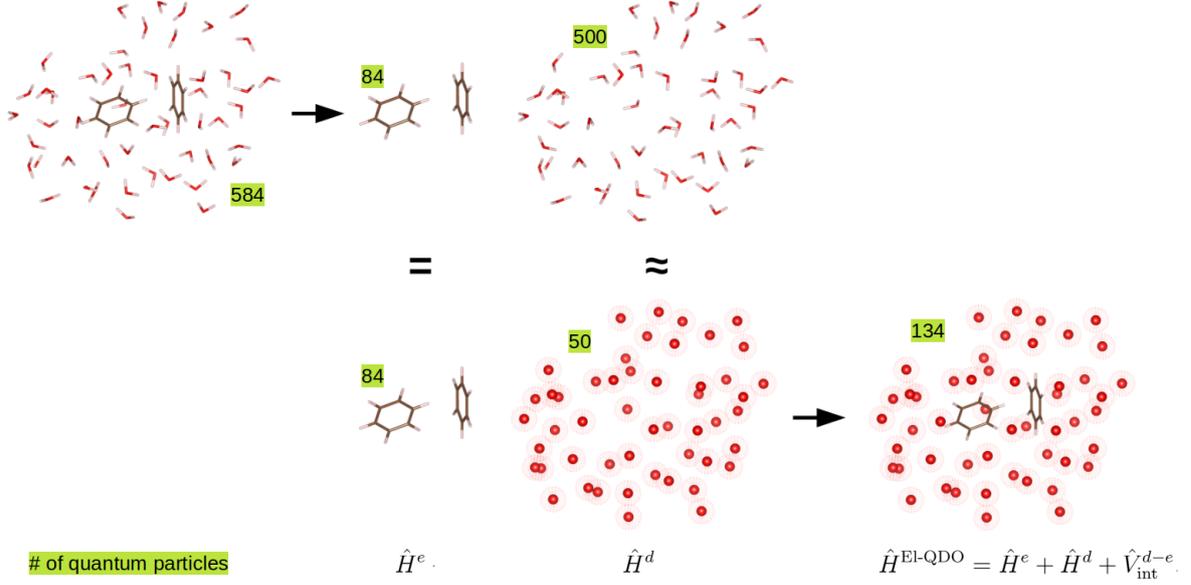


Figure 2.8: An example of the mapping of the T-shaped benzene dimer in an environment composed of 50 water molecules onto the T-shaped benzene dimer embedded in an environment composed of 50 QDOs.

represented by the QDOs and also the QDOs polarize in the presence of the electronic system. The same holds also for the many-body dispersion effects, which are present within both subsystems and also between them. By solving the full Hamiltonian, these many-body correlation effects are included beyond traditional self-consistent approaches, like the ones applied in the existing polarizable FFs.

Since the QDO model is introduced to describe the long-range interactions, it can not efficiently model the short-range repulsive region arising from the exchange between electrons and from the nuclear repulsion. For this reason, the short-range repulsion between the two subsystems and also between the QDOs is missing in the El-QDO embedding. The problem is discussed in more detail in Sec. 2.7.

2.6 Damping of the Coulomb potential

In practice, the use of the bare Coulomb potential for the interactions between the electrons in the QDOs' environment and between interacting QDOs displays an unphysical

overpolarization in the short-range region. This is an effect well known from QM:MM methods [17], often referred to as the polarization catastrophe. The reason is that the missing repulsion between the drudon-drudon, electron-drudon and electron-point charge pairs leads to an unphysical drift of the quantum particles towards the classical particles of the opposite charge, namely the centers of the QDOs or the classical point charges.

Another problem is the necessity of satisfying the cusp conditions [97] coming from the $1/r$ divergences of the Coulomb potential, providing the exact properties of the wavefunction when an electron approaches a drudon, a center of QDO or a point charge or when a drudon approaches other drudons, centers of QDOs and point charges. Missing cusps in the wavefunctions manifest in an error of the corresponding energies in the case of the wavefunction-based numerical methods. For the diffusion Monte Carlo method, used in this Thesis, the missing cusps manifest only in the decrease of the computational efficiency via the increase of the statistical noise.

To minimize these issues, damping functions are introduced to regularize the interactions, changing the shape of the Coulomb potential in the short-range region and removing the $1/r$ divergence. In this work, we test four functional forms of the damping functions from the state-of-the-art literature, previously used for the QM:MM methods. Our goal is to show the robustness of the El-QDO methods with respect to the choice of the damping for distances characteristic of non-covalent interactions.

The four functional forms are namely the error function (erf) [31]

$$V_{\text{erf}}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \operatorname{erf}\left(\frac{r_{ij}}{\sqrt{2}\sigma_{ij}}\right), \quad (2.60)$$

the Gaussian damping (exp2) [61]

$$V_{\text{exp2}}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \left[1 - e^{-\left(\frac{r_{ij}}{\sigma_{ij}}\right)^2} \right], \quad (2.61)$$

the exponential with the 4th power (exp4) [88]

$$V_{\text{exp4}}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \left[1 - e^{-\left(\frac{r_{ij}}{\sigma_{ij}}\right)^4} \right], \quad (2.62)$$

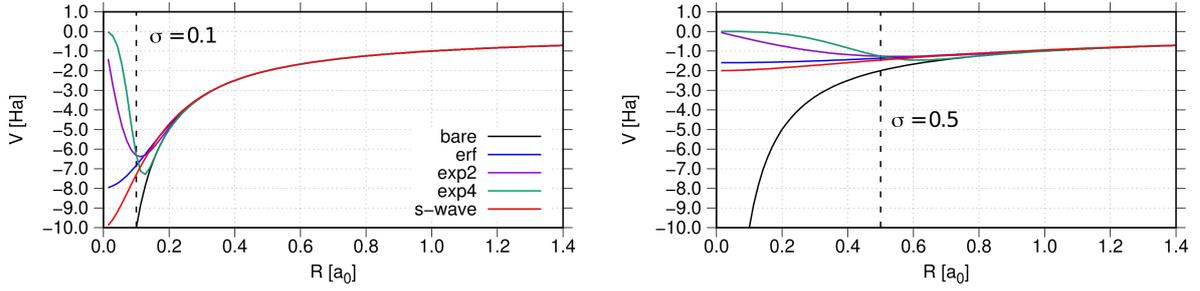


Figure 2.9: Damped Coulomb potentials from Eqs. 2.60,2.61,2.62 and 2.63 compared to the bare Coulomb potential for values of $\sigma = 0.1$ and $\sigma = 0.5$ (indicated via the dashed vertical lines).

and the s-wave expansion (s-wave) [17, 30]

$$V_{\text{s-wave}}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \left[1 - e^{-\frac{2r_{ij}}{\sigma_{ij}}} - \frac{r_{ij}}{\sigma_{ij}} e^{-\frac{2r_{ij}}{\sigma_{ij}}} \right], \quad (2.63)$$

where σ_{ij} is specified for all pairs of particles using the combination rule $\sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2}$. All four damping functions are plotted in Fig. 2.9 for fixed values of $\sigma = 0.1$ and $\sigma = 0.5$.

One problem of these damping functions comes from asymmetric interaction effects, that introduce additional artificial polarization. The standard damping functions in QM:MM methods or for systems of interacting QDOs can be considered as charge regularization as a function of the distance between the particles. We will use V_{erf} as an example, but all the following considerations apply to all the damping functions defined above. Eq. 2.60 can be written as

$$V_{\text{erf}}(r_{ij}) = \frac{q_i q_j}{r_{ij}} \text{erf} \left(\frac{r_{ij}}{\sqrt{2}\sigma_{ij}} \right) = \frac{q_i(r_{ij}) q_j(r_{ij})}{r_{ij}}, \quad (2.64)$$

where $q_i(r_{ij})$ and $q_j(r_{ij})$ are position dependent charges (see Fig. 2.10a). This may lead to an artificial polarization because the particles see other parts of the system as artificially charged.

In order to test the effect of this additional artificial polarization we construct a uniform version of the damping (referred to as uniform damping). Instead of V_{erf} from the Eq. 2.60, we introduce a damping distance r_{damp}

$$V_{\text{erf}}(r_{ij}, r_{\text{damp}}) = \frac{q_i q_j}{r_{ij}} \text{erf} \left(\frac{r_{\text{damp}}}{\sqrt{2}\sigma_{ij}} \right), \quad (2.65)$$

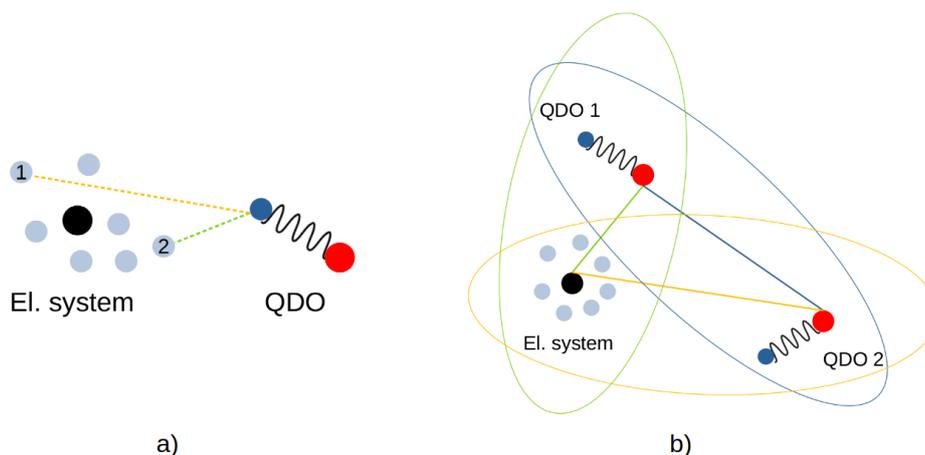


Figure 2.10: a) The artificial polarization problem due to the standard damping functions caused by the position-dependent charges (shown by different colors of the interactions); and b) a schematic of the uniform damping, where the three straight lines represent the damping distances r_{damp} used for the uniform damping functions (e.g. in Eq. 2.65) for all the pairs of particles within the ellipse of the same color.

which is taken as a center-center distance in the case of a pair of QDOs and center of mass of electronic system - QDO center for El-QDO pairs, with σ_i defined for the entire electronic system/QDO. To make it clearer we show a schematic representation of the uniform damping for an electronic system interacting with two QDOs in Fig. 2.10b.

All eight possibilities of the damping functions defined above (4 functional forms, each for normal and uniform version) are tested in Sec 4.4.4 on an example of the argon dimer, in order to show the robustness of the potential energy surface in the region relevant for the non-covalent interactions with respect to the choice of the damping. For all the other results in Chap. 4 we used the normal error function damping from Eq. 2.60 and all the results in Chap. 5 are obtained with the bare Coulomb potential.

2.7 Short-range repulsion

The problem of the inability of the QDO model to describe the short-range repulsion of real systems is usually solved using two different approaches. The first option is to add the Hartree-Fock (HF) interaction energy on top of the QDO one, which partially

works for noble gases [152] because HF misses all the energy contributions apart for repulsion. The second option is to fit the missing part of the potential energy surface using *ab initio* reference data. This approach is common also for the QM:MM methods, where the repulsive part is taken directly from fitted force-fields [30].

In this work, we use a fit of the repulsion using exponential functions of the form

$$V_{\text{rep}}(R_{ij}) = \sum_{k=1}^N a_k e^{-b_k R_{ij}} \quad (2.66)$$

for each center-center and nucleus-center pair. The functional form of Eq. 2.66 is inspired by the dependency of the exchange energy on the overlap of the densities, which decays exponentially as a function of the separation and which was briefly discussed in Sec. 2.2. The parameters in Eq. 2.66 are then balanced with the parameters σ used for the damping functions in order to reproduce a set of reference results. In this Thesis, we used the short-range repulsion only to demonstrate the capability of the QDO model and the El-QDO method to reproduce the full binding curves of the noble gas and water dimers. The distances between the main fragment and the environment in all the other applications in Chap. 4 were chosen so that the fast decaying repulsion can be neglected and the study of the short-range behavior of the QDO model in Chap. 5 was done without the external repulsion.

Computational methods

Parts of Secs. 3.5 and 3.6 in this Chapter have been published in this or a similar form in:

M. Ditte, M. Barborini, L. M. Sandonas, A. Tkatchenko. “Molecules in Environments: Toward Systematic Quantum Embedding of Electrons and Drude Oscillators” *Phys. Rev. Lett.* **131**, 228001 (2023)

and will be published in:

M. Ditte, M. Barborini, A. Tkatchenko. “Electrons embedded in charged oscillators: A quantum embedding approach for molecular systems” In preparation.

and in:

M. Ditte, M. Barborini, A. Tkatchenko. “Quantum Drude Oscillators Coupled with Coulomb Potential as an Efficient Model for Bonded and Non-Covalent Interactions in Atomic Dimers” In preparation.

and have been produced in collaboration with the above authors.

Since, as discussed in the previous Chapter, the Schrödinger's equation is not analytically solvable for many-body systems, many numerical techniques have been developed in order to obtain an approximate description of the ground or excited states.

Within the Quantum Chemistry methods we can fundamentally distinguish between three approaches: 1. wavefunction-based methods that approximate the solution of the eigenvalue problem through a Slater determinant or a combination of them (e.g. Hartree-Fock [44], Configuration Interaction [37], Coupled Cluster [120]); 2. density-based methods that approximate the electronic Hamiltonian or parts of it, leading to a model that can be numerically solved (e.g. Density Functional Theory [91], Density Functional tight-binding [172]); and 3. perturbative approaches that reintroduce the electron correlation through perturbation energy contributions (e.g. Møller-Plesset perturbation theory [36]).

Within the wavefunction methods, we can also find the quantum Monte Carlo (QMC) techniques, which are stochastic methods used to integrate the energy functional and other physical observables over an optimized trial wavefunction. Apart from the computational efficiency of QMC, which comes from the intrinsic parallelization of their algorithms, one of the main advantages resides in the fact that with these methods it is possible to calculate integrals over explicitly correlated wavefunctions, thus guaranteeing more flexibility to tackle also complex Hamiltonians like the one comprised of electrons and drudons defined in Chap. 2.

This Chapter is thus dedicated to the description of the numerical methods, and in particular to the QMC methods, employed in our works. After the description of the variational principle in quantum mechanics (Sec. 3.1) and of the simplest numerical method, *ie* the Hartree-Fock method, used to approximately compute the ground state of the electronic Schrödinger equation (Sec. 3.2), in Sec. 3.3 we briefly discuss the basic concepts of Density Functional Theory (DFT), and the various dispersion methods (Sec. 3.4) that are used in this Thesis to describe dispersion interactions within the DFT framework. Finally, the last sections of this Chapter (Secs. 3.5 and 3.6) discuss the QMC methods, the generalization of the algorithms introduced to integrate the El-QDO Hamiltonian, and the form of the variational *ansatz* developed to approximate the El-QDO ground state.

3.1 Variational Principle

The basic principle behind all the numerical methods used to solve the time-independent Schrödinger equation is the variational principle, which states that given an approximate expression of the ground state eigenfunction $\Psi_T(\bar{\mathbf{r}}; \boldsymbol{\alpha})$ of a Hamiltonian operator \hat{H} , depending on the vector $\boldsymbol{\alpha}$ of variational parameters, we can define the energy functional

$$E[\Psi_T(\bar{\mathbf{r}}; \boldsymbol{\alpha})] = \frac{\langle \Psi_T(\boldsymbol{\alpha}) | \hat{H} | \Psi_T(\boldsymbol{\alpha}) \rangle}{\langle \Psi_T(\boldsymbol{\alpha}) | \Psi_T(\boldsymbol{\alpha}) \rangle} = \frac{\int_V \Psi_T^*(\bar{\mathbf{r}}; \boldsymbol{\alpha}) \hat{H} \Psi_T(\bar{\mathbf{r}}; \boldsymbol{\alpha}) d\bar{\mathbf{r}}}{\int_V |\Psi_T(\bar{\mathbf{r}}; \boldsymbol{\alpha})|^2 d\bar{\mathbf{r}}} \geq E_0, \quad (3.1)$$

which is an upper bound of the exact ground state's energy E_0 . Naturally, the equality holds only when $\Psi_T(\bar{\mathbf{r}}; \boldsymbol{\alpha}) = \Psi_0(\bar{\mathbf{r}})$, being $\Psi_0(\bar{\mathbf{r}})$ the wavefunction corresponding to the exact ground state.

In general however $\Psi_T(\bar{\mathbf{r}}; \boldsymbol{\alpha})$ is only an approximation of the true electronic ground state, and the best estimation of the ground state energy can be achieved through the optimization of the parameters $\boldsymbol{\alpha}$ by minimizing Eq. 3.1.

3.2 Hartree-Fock theory

The first and simplest numerical algorithm derived from the variational principle described above was proposed by Hartree[67], Fock[51] and Slater[165], and is based on the mean-field approximation, for which each particle in the Hamiltonian is considered as an independent particle that exists in the effective potential generated by all the other particles.

In this representation, the fully correlated wavefunction of a system of fermionic particles, such as the electrons, that should depend on all the relative distances between the particles, is approximated by the antisymmetric product of single-particle molecular

orbitals, *ie* the Slater determinant (SD)[51, 165]

$$\Psi_{\text{SD}}(\bar{\mathbf{x}}) = \frac{1}{\sqrt{N_e!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \dots & \psi_1(\mathbf{x}_{N_e}) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \dots & \psi_2(\mathbf{x}_{N_e}) \\ \dots & \dots & \dots & \dots \\ \psi_{N_e}(\mathbf{x}_1) & \psi_{N_e}(\mathbf{x}_2) & \dots & \psi_{N_e}(\mathbf{x}_{N_e}) \end{vmatrix}, \quad (3.2)$$

where the functions $\psi_k(\mathbf{x})$ are build as the product of a spatial function and an eigenvector of the spin operator \hat{S}_z , *ie*. $\psi_k(\mathbf{x}) = \psi_k(\mathbf{r}) \left| \frac{1}{2}, \pm \frac{1}{2} \right\rangle$. The wavefunctions built through the Slater determinant are in fact eigenstates of both the \hat{S}^2 and \hat{S}_z spin operators.

When describing isolated atomic systems, the choice for the molecular orbitals $\psi_k(\mathbf{r}_i)$ is that of a set of ‘atomic orbitals’ $\phi_k(\mathbf{r}_i) = R(r_i)Y_{l,l_z}(\theta_i, \phi_i)$ similar to eigenfunctions of the hydrogen atom. The angular part $Y_{l,l_z}(\theta_i, \phi_i)$ is represented by spherical or cubic Harmonics, while the spatial part $R(r_i)$ is usually written as polynomial functions multiplied by a Gaussian or a Slater-type orbital. The list of the atomic orbitals used to describe a molecular or atomic system is referred to as the atomic basis set of Q functions centered on the various atoms.

To describe the eigenstates of molecular systems, the single electron molecular orbitals are written as a linear combination of Q atomic orbitals that belong to all the atoms:

$$\psi_k(\mathbf{r}_i) = \sum_{a=1}^{N_n} \sum_{\nu=1}^{Q_a} c_{\nu}^a \phi_{\nu}^a(r_{i,a}) \quad (3.3)$$

where N_n is the number of atoms in the system and the atomic wavefunctions $\phi_{\nu}^a(r_{i,a})$ are centered on the a th atom so that $\phi_{\nu}^a(r_{i,a}) = \phi_{\nu}(|\mathbf{r}_i - \mathbf{R}_a|)$.

Thus, within this framework, the Hartree-Fock (HF) theory is a variational method based on the minimization of the energy functional in Eq. 3.1 with respect to the SD in Eq. 3.2

$$E[\Psi_{\text{SD}}] = \langle \Psi_{\text{SD}} | \hat{H}_1 | \Psi_{\text{SD}} \rangle + \langle \Psi_{\text{SD}} | \hat{H}_2 | \Psi_{\text{SD}} \rangle \quad (3.4)$$

where

$$\hat{H}_1 = -\frac{1}{2} \sum_{i=1}^{N_e} \nabla_{\mathbf{r}_i}^2 - \sum_{a=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|} \quad (3.5)$$

is the sum of N_e single-electron Hamiltonians

$$\hat{h}_i = -\frac{1}{2}\nabla_{\mathbf{r}_i}^2 - \sum_{a=1}^{N_n} \frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|}$$

and the two-particle interacting potential

$$\hat{H}_2 = \sum_{i<j=1}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (3.6)$$

Since the first term can be separable into single-particle operators

$$\langle \Psi_{\text{SD}} | \hat{H}_1 | \Psi_{\text{SD}} \rangle = \sum_{i=1}^{N_e} \langle \Psi_{\text{SD}} | \hat{h}_i | \Psi_{\text{SD}} \rangle, \quad (3.7)$$

for the orthogonality of the electronic states $\psi_k(\mathbf{r})$, we can reduce it to

$$\langle \Psi_{\text{SD}} | \hat{H}_1 | \Psi_{\text{SD}} \rangle = \sum_{k=1}^{N_e} \langle \psi_k(\mathbf{x}_i) | \hat{h}_i | \psi_k(\mathbf{x}_i) \rangle = \sum_k I_k, \quad (3.8)$$

where the sum over k indicates the sum over all the eigenstates which are occupied by the N_e electrons.

The average over the second term

$$\langle \Psi_{\text{SD}} | \hat{H}_2 | \Psi_{\text{SD}} \rangle = \sum_{i<j=1}^{N_e} \langle \Psi_{\text{SD}} | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \Psi_{\text{SD}} \rangle, \quad (3.9)$$

in which, for each electronic pair i, j , the only terms that survive in the expansion are those that contain the electronic coordinate \mathbf{r}_i and \mathbf{r}_j , which can be also rewritten as

$$\begin{aligned} \sum_{i<j=1}^{N_e} \langle \Psi_{\text{SD}} | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \Psi_{\text{SD}} \rangle &= \\ &= \frac{1}{2} \sum_{k,l} \left[\langle \psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j) | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j) \rangle - \right. \\ &\quad \left. - \langle \psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j) | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi_l(\mathbf{x}_i) \psi_k(\mathbf{x}_j) \rangle \right], \quad (3.10) \end{aligned}$$

which is the sum of two contributions, that are respectively called the direct term

$$J_{k,l} = \langle \psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j) | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j) \rangle, \quad (3.11)$$

that defines the mean value of the Coulomb interaction between the i th and j th electrons respectively, in the quantum states k and l , and the exchange term

$$K_{k,l} = \langle \psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j) | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi_l(\mathbf{x}_i) \psi_k(\mathbf{x}_j) \rangle, \quad (3.12)$$

that takes into account the energy of the interaction between the two states $\psi_k(\mathbf{x}_i) \psi_l(\mathbf{x}_j)$ and $\psi_l(\mathbf{x}_i) \psi_k(\mathbf{x}_j)$ obtained by exchanging the coordinates of the i th and the j th electrons.

At this point, the final expression of the energy functional can be written as the sum of these three contributions

$$E[\Psi_{\text{SD}}] = \sum_k I_k + \frac{1}{2} \sum_{k,l} [J_{k,l} - K_{k,l}]. \quad (3.13)$$

To optimize the set of molecular orbitals $\psi_k(\mathbf{r}_i)$ that appear in the Slater determinant, we now have to minimize this energy functional, by imposing the condition

$$\delta E - \sum_{k,l} \epsilon_{k,l} \delta \langle \psi_k | \psi_l \rangle = 0 \quad (3.14)$$

that takes into account the orthogonality of the molecular orbitals $|\psi_k\rangle$ through a Lagrangian multiplier $\epsilon_{k,l}$. By expanding the variation δE in this last equation we obtain N_e coupled Hartree-Fock equations

$$\begin{aligned} \epsilon_k \psi_k(\mathbf{r}) = & \left[-\frac{1}{2} \nabla_{\mathbf{r}}^2 - \sum_{a=1}^M \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|} \right] \psi_k(\mathbf{r}) + \\ & + \sum_l \left[\int d\mathbf{r}' \frac{|\psi_l(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \psi_k(\mathbf{r}) - \delta_{\sigma_k \sigma_l} \int d\mathbf{r}' \frac{\psi_l^*(\mathbf{r}') \psi_k(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_l(\mathbf{r}) \right] \quad \forall k, \end{aligned} \quad (3.15)$$

each associated to a single-particle state $|\psi_k\rangle$, that must be solved iteratively. For each of these coupled equations, the first term represents the electronic kinetic contribution; the second term, or Hartree term, describes the mean field contribution which depends on the interaction of the particular electron with all the other electrons; and finally

the last term, *ie* the Fock term or exchange term, introduces the non-local interaction between two electrons having identical spins, *i.e.* $\delta_{\sigma_k\sigma_l}$.

As anticipated at the beginning of this section, the HF equations have the limitation of approximating the full correlated wavefunction through the antisymmetric product of single-particle functions, treating them as independent non-interacting particles ‘dressed’ by the mean-field interaction with the other particles. For this reason the energy $E[\Psi_{\text{SD}}]$ of the HF ground state will always be higher than that of the true eigenvalue

$$E_0 < E[\Psi_{\text{SD}}]. \quad (3.16)$$

The difference in energy between the exact ground state energy and the HF estimation

$$E_{\text{corr}} = E[\Psi_{\text{SD}}] - E_0 \quad (3.17)$$

is thus equal to the missing ‘correlation energy’ E_{corr} of the electrons. The only correlation between the particles, that is included in the HF method, is the one that originates in imposing the antisymmetrization of the wavefunction, for which two fermions with the same spin can not occupy the same spatial state, which is also at the origin of the exchange contribution. For this reason, the HF method is usually used as the ‘uncorrelated’ reference through which it is possible to estimate the correlation energy recovered by the other more accurate numerical approaches such as CI [37] or CC [120].

3.3 Density Functional Theory

As discussed in the previous section, the HF method does not describe the explicit correlation between the electrons, and for this reason, many post-HF numerical methods have been developed with the aim of capturing the missing correlation effects, such as perturbation theory or multi-determinant expansions.

One alternative and computationally less demanding approach is that of Density Functional Theory (DFT), based on the idea of reducing the degrees of freedom of the Schrödinger equation, defining the variational principle on the electronic density and not on the electronic wavefunction.

In fact, being $\Psi_0(\bar{\mathbf{r}})$ the ground state eigenfunction of the electronic Schrödinger Eq. 2.8, we define the electronic density as the integral

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_e} \int_V \delta(\mathbf{r} - \mathbf{r}_i) |\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_{N_e})|^2 d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \dots d\mathbf{r}_{N_e}, \quad (3.18)$$

that, while the wavefunction depends on $3N_e$ degrees of freedom, is a local quantity that only depends on the 3 coordinates of the position in space.

The two main theorems on which DFT is based were proven in 1964 by P. Hohenberg and W. Kohn[78] and state that:

Theo. 1 The external potential, that defines the motion of a single electron in the field of nuclei $v_{ne} = \int \Psi_0(\bar{\mathbf{r}}) \hat{V}_{ne}(\bar{\mathbf{r}}) \Psi_0(\bar{\mathbf{r}}) d\bar{\mathbf{r}}$ is a unique functional of the electron density $\rho(\mathbf{r})$, ie $V_{ne}[\rho] = \int \rho(\mathbf{r}) v_{ne}(\mathbf{r}) d\mathbf{r}$.

Theo. 2 Given the external potential $V_{ne}[\rho]$, there exists a functional $F[\rho]$ so that the ground state energy E_0 and the electronic density $\rho_0(\mathbf{r})$ can be variationally obtained by minimizing the equation

$$E_0[\rho] = F[\rho] + V_{ne}[\rho] \quad (3.19)$$

with the condition that $\int \rho(\mathbf{r}) d\mathbf{r} = N_e$.

Naturally at the minimum value of $E_0[\rho = \rho_0]$ the functional $F[\rho_0]$ is equal to the expectation value of $T[\rho] + V_{ee}[\rho]$ over the ground state, but since the form of $F[\rho_0]$ is unknown the theorems are not applicable for direct calculations.

The spreading of the DFT method was made possible by the formulation introduced by W. Kohn and L. J. Sham (KS) [106], which enabled direct calculations of the energies and other observables of the electronic systems. In the Kohn-Sham (KS) approach, as in the HF theory, we first define a set of independent single-electron orbitals (Eq. 3.3) that represent the motion of non-interacting electrons in the external potential $v_{ne}(\mathbf{r})$ given by the atomic nuclei. The electronic density is then represented as the sum of the square modulus of these single-particle molecular orbitals, called KS orbitals,

$$\rho(\mathbf{r}) = \sum_{k=1}^{N_e} |\psi_k(\mathbf{r})|^2, \quad (3.20)$$

and can in principle give an exact representation of the electronic density of the system. As for the HF method, the KS orbitals must not be considered as the orbitals of the exact wavefunction because they represent non-interacting electrons. Their meaning is only related to Eq. 3.20 and only the density can be considered as a physical observable of the interacting system. Having defined the density through the non-interacting single-particle orbitals, we can define the kinetic energy as

$$T_s[\rho] = -\frac{1}{2} \sum_{k=1}^{N_e} \int \psi_k^*(\mathbf{r}) \nabla_{\mathbf{r}}^2 \psi_k(\mathbf{r}) d\mathbf{r}, \quad (3.21)$$

which describes the motion of free electrons. In the KS formulation, the total energy defined in Eq. 3.19 is now rewritten as

$$E[\rho] = T_s[\rho] + V_{ne}[\rho] + J[\rho] + E_{xc}[\rho], \quad (3.22)$$

where $J[\rho] = \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'$ is the expectation value of the \hat{V}_{ee} potential between the electrons occupying the single-particle states, similar to the HF contribution in Eq. 3.11. Moreover, the exchange-correlation functional $E_{xc}[\rho]$ that appears in Eq. 3.22 is defined as the sum of different contributions,

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] - K[\rho], \quad (3.23)$$

the first part of which is the difference between the unknown kinetic energy of the correlated electrons and that of the ones occupying the single-particle orbitals $T[\rho] - T_s[\rho]$, while the $K[\rho]$ functional, defined as $J[\rho] - V_{ee}[\rho]$, includes all the exchange-correlation contributions to the electronic Coulomb interaction.

The minimization of the energy with respect to Eq. 3.20 leads to the equation

$$\left[-\frac{1}{2} \nabla_{\mathbf{r}}^2 - \sum_{a=1}^{N_n} \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[\rho]}{\delta \rho} \right] \psi_k^{KS}(\mathbf{r}) = \epsilon_k^{KS} \psi_k^{KS}(\mathbf{r}), \quad (3.24)$$

similar to the HF ones, that optimize the KS orbitals in a self-consistent way. The question that remains unsolved in these KS equations, is that of the definition of the exchange-correlation functional $E_{xc}[\rho]$ for which only approximated forms have been proposed[49]. In this Thesis, in particular, we use two of the most common and well-established exchange-correlation functionals, namely PBE [143] and PBE0 [4], the first of which is based on the Generalized Gradient Approximation (GGA), while the second belongs to the class of the Hybrid functionals that include also a percentage of the ‘exact exchange’ energy coming from HF.

3.4 Dispersion methods

Despite the wide applicability and versatility of DFT, as discussed in the previous section only approximate forms of the exchange-correlation E_{xc} functionals are present. Usually, one limitation of these functionals is their inability to correctly reproduce electronic correlation arising from dispersion interactions, leading to an over- or under-binding problem.

For this reason, various methods have been introduced to improve the interaction energies or the E_{xc} functionals, using knowledge coming from perturbation theories (Sec. 2.2) or *ab initio* calculations. In the next sections, we will briefly recall three of these approaches, that have been applied in this Thesis, namely the D4 London-dispersion model [25], the Tkatchenko-Scheffler (TS) method [180] and the many-body dispersion (MBD) method [179, 9, 75]. While the first approach is parametrized independently for different exchange-correlation functionals and its dispersion energy depends on the atom-in-molecule polarizabilities only via the number of surrounding atoms, in the TS and MBD methods, the polarizabilities of the atom-in-molecule depend on the electronic density obtained from the DFT calculations.

3.4.1 D4 London-dispersion model

The dispersion energy in the D4 London-dispersion model[25] is written as

$$E_{\text{disp}}^{D4} = E_{\text{disp}}^{(6,8)} + E_{\text{disp}}^{(9),\text{ATM}}, \quad (3.25)$$

where the first term contains pairwise contributions and the second term is the Axilrod-Teller-Muto (ATM) three-body term.

The pairwise contribution is written as the sum of the $\propto -\frac{1}{R^6}$ and $\propto -\frac{1}{R^8}$ terms

$$E_{\text{disp}}^{(6,8)} = - \sum_{AB} \sum_{n=6,8} s_n \frac{C_{AB}^{(n)}}{R_{AB}^{(n)}} f_{\text{damp}}^{(n)}(R_{AB}), \quad (3.26)$$

where R_{AB} is the distance between atoms A and B, $C_{AB}^{(n)}$ is the corresponding dispersion coefficient, f_{damp} is the Becke-Johnson damping function and s_n is a scaling factor.

Within this approach, the dispersion coefficients are obtained by integrating over the frequency-dependent polarizabilities, so that for C_6 we have

$$C_6^{AB} = \frac{3}{\pi} \int_0^\infty d\omega \alpha^A(i\omega) \alpha^B(i\omega), \quad (3.27)$$

where α^i are atom-in-molecule dipole polarizabilities obtained by reweighing of the reference polarizabilities of free atoms by weights that depend on the coordination number, and take into account the local environment of the atom.

The three-body terms are written as

$$E_{\text{disp}}^{(9),\text{ATM}} = \sum_{ABC} E^{ABC} f_{\text{damp}}^{(9)}(\bar{R}_{ABC}), \quad (3.28)$$

where f_{damp} is a damping function depending on the relative positions of the three atoms \bar{R}_{ABC} and E^{ABC} is the ATM term

$$E^{ABC} = \frac{C_9^{ABC} (3 \cos \Theta_a \cos \Theta_b \cos \Theta_c + 1)}{(R_{AB} R_{BC} R_{CA})^3}, \quad (3.29)$$

with R_{ij} being the distance between atoms i and j , Θ_i are the angles of the triangle formed by the atoms A, B and C, and C_9^{ABC} is the corresponding dispersion coefficient, which is approximated as $C_9^{ABC} \approx \sqrt{C_6^{AB} C_6^{BC} C_6^{CA}}$.

3.4.2 Tkatchenko-Scheffler method

The dispersion energy in the Tkatchenko-Scheffler (TS) [180] method contains only the pairwise $\propto -\frac{1}{R^6}$ term which is written as

$$E_{\text{disp}}^{D4} = -\frac{1}{2} \sum_{A,B} f_{\text{damp}}(R_{AB}, R_A^0, R_B^0) C_{6AB} R_{AB}^{-6}, \quad (3.30)$$

where again R_{AB} is the distance between atoms A and B, f_{damp} is a damping function avoiding divergences in the short-range region and C_{6AB} is the corresponding dispersion coefficient, which is calculated as

$$C_{6AB} = \frac{2C_{6AA}C_{6BB}}{\left[\frac{\alpha_B^0}{\alpha_A^0} C_{6AA} + \frac{\alpha_A^0}{\alpha_B^0} C_{6BB} \right]}, \quad (3.31)$$

where α_i^0 and C_{6ii} are free-atom reference values of the static polarizability and of the homo-nuclear dispersion coefficients. The values of the dispersion coefficient of the atom-in-molecule are obtained via the rescaling of the free values by an effective volume of the atoms, which is obtained via the Hirshfeld partitioning of the electron density obtained from the DFT calculations:

$$C_{6AA}^{\text{eff}} = \left(\frac{V_A^{\text{eff}}}{V_A^{\text{free}}} \right)^2 C_{6AA}^{\text{free}}. \quad (3.32)$$

The only free parameter in the TS method is in the Fermi-type damping function f_{damp} , and it has the role of determining the onset of the dispersion interactions. Also for TS, as for the D4 method, the choice of the parameters is tuned separately for each class of the exchange-correlation functionals.

3.4.3 Many-Body Dispersion method

In the basic many-body dispersion (MBD) method introduced in Ref. 179, the dispersion contribution is calculated as an interaction energy of a set of dipole-coupled quantum oscillators defined through the Hamiltonian

$$\hat{H}^{\text{MBD}} = -\frac{1}{2} \sum_{p=1}^N \nabla_{\mathbf{x}_p}^2 + \frac{1}{2} \sum_{p=1}^N \omega_p^2 \chi_p^2 + \sum_{p>q}^N \omega_p \omega_q \sqrt{\alpha_p \alpha_q} \mathbf{x}_p T_{pq} \mathbf{x}_q, \quad (3.33)$$

where \mathbf{x}_q are proportional to the displacement of the oscillator p from its equilibrium, T_{pq} is the dipole-dipole tensor (including also a damping function), ω_p is a characteristic excitation frequency and the polarizabilities α_p are obtained via the self-consistent screening (SCS) equation

$$\alpha_p^{\text{SCS}}(i\omega) = \alpha_p^{\text{TS}}(i\omega) + \alpha_p^{\text{TS}}(i\omega) \sum_{q \neq p}^N T_{pq} \alpha_q^{\text{SCS}}(i\omega), \quad (3.34)$$

with α_p^{TS} being the polarizabilities coming from the TS method.

The dispersion energy is then obtained via the equation

$$E_{\text{disp}}^{\text{MBD}} = \frac{1}{2} \sum_{p=1}^{3N} \sqrt{\lambda_p} - \frac{3}{2} \sum_{p=1}^N \omega_p^{\text{SCS}}, \quad (3.35)$$

where λ_p are the eigenvalues of the Hamiltonian \hat{H}^{MBD} . The polarizabilities obtained through the SCS equation include the long-range electrostatic screening, coming from the dynamic electric field created by the other atoms. By mapping the problem onto a set of dipole-coupled oscillators, the $E_{\text{disp}}^{\text{MBD}}$ can be shown to describe the explicit many-body dispersion of all orders, beyond the two or two/three-body terms present in TS or D4 methods.

In this Thesis we use two implementations of MBD, which improve on the original method, which are the MBD@rsSCS [9] and the MBD-NL [75] versions. With respect to the original method, MBD@rsSCS includes an effective range-separation of the coupling between the atomic response functions in order to increase its accuracy in the description of non-metallic materials with highly anisotropic electronic responses. The MBD-NL, on the other hand, extends non-local vdW functionals for polarization and interatomic methods for many-body interactions, extending the applicability of the MBD method to ionic and metallic compounds, and hybrid metal-organic interfaces.

3.5 Quantum Monte Carlo methods

Quantum Monte Carlo (QMC) methods are a family of stochastic techniques for integrating the many-body time-independent Schrödinger equation over a chosen trial wavefunction [52, 95, 14].

In this Thesis, we use two of the most common methods, namely variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) [52, 95, 14], implemented in QMeCha $\alpha.0.3.0$ [12], a QMC package published privately on Github and generalized to integrate the Hamiltonian of the El-QDO embedding approach which includes electrons and drudons.

In the following subsections, we first introduce the concept of stochastic integration, followed by the basic concepts regarding the generalized VMC and DMC algorithms. Moreover, we discuss the wavefunction optimization algorithms employed in the VMC framework to optimize the variational parameters of our electron-drudon *ansatz*.

3.5.1 Stochastic integration and the Metropolis-Hastings algorithm

Let us suppose we need to evaluate the integral of a function g , which depends on a vector of N independent variables $\bar{\mathbf{r}} = (r_1, \dots, r_N)$

$$I = \int_{\Omega} g(\bar{\mathbf{r}}) d\bar{\mathbf{r}}, \quad (3.36)$$

where Ω is the integration domain. The simplest approach to perform this integration stochastically is to generate \mathcal{N} uniformly random vectors $\{\bar{\mathbf{r}}_i\}_{i=1}^{\mathcal{N}}$ within the integration domain Ω , reducing the integral

$$I \approx \langle g \rangle = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} g(\bar{\mathbf{r}}_i) \quad (3.37)$$

to the expectation value of the function g over the finite sample \mathcal{N} of the continuous domain Ω . From the central limit theorem, we can associate with this expectation value an error in the estimation of the integral, *ie* $\sigma_I = \sigma_g / \sqrt{\mathcal{N}}$, that depends on the variance of the function g

$$\sigma_g^2 = \int_{\Omega} (g(\bar{\mathbf{r}}) - \langle g \rangle)^2 d\bar{\mathbf{r}}, \quad (3.38)$$

and on the size of the sampling \mathcal{N} , not explicitly on the dimensionality N of the function g .

The efficiency of this basic approach can be improved through the so-called importance sampling, which instead of uniformly sampling the space, introduces an auxiliary positive function that is used to force the sample the domain regions that contribute the most to the total value of the integral. Let us write the function $g(\bar{\mathbf{r}})$ as a product $g(\bar{\mathbf{r}}) = f(\bar{\mathbf{r}})\pi(\bar{\mathbf{r}})$, with the condition that $\pi(\bar{\mathbf{r}}) > 0, \forall \bar{\mathbf{r}} \in \Omega$ and that $\int_{\Omega} \pi(\bar{\mathbf{r}}) d\bar{\mathbf{r}} = 1$, so that π can be interpreted as a probability density. The integral of the function g can thus be rewritten as

$$I = \int_{\Omega} f(\bar{\mathbf{r}})\pi(\bar{\mathbf{r}}) d\bar{\mathbf{r}} \quad (3.39)$$

and evaluated as the sum over the values of the function f

$$I \approx \langle f \rangle_{\pi} = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} f(\bar{\mathbf{r}}_i), \quad (3.40)$$

where the samples $\{\bar{\mathbf{r}}_i\}_{i=1}^{\mathcal{N}}$ are generated according to the probability density π . Again, the associated statistical error will now be defined as $\sigma_f/\sqrt{\mathcal{N}}$, where

$$\sigma_f^2 = \int_{\Omega} (f(\bar{\mathbf{r}}) - \langle f \rangle_{\pi})^2 \pi(\bar{\mathbf{r}}) d\bar{\mathbf{r}} \quad (3.41)$$

will be the variance of the function f over the samplings according to π , which for the finite sample can be written as the finite sum

$$\frac{\sigma_f}{\sqrt{\mathcal{N}}} \approx \sqrt{\frac{1}{\mathcal{N}(\mathcal{N}-1)} [\langle f(\bar{\mathbf{r}})^2 \rangle_{\pi} - \langle f(\bar{\mathbf{r}}) \rangle_{\pi}^2]}. \quad (3.42)$$

The optimal choice of the factorization $g(\bar{\mathbf{r}}) = f(\bar{\mathbf{r}})\pi(\bar{\mathbf{r}})$ is the one minimizing the statistical error in Eq. 3.42, and thus minimizing the variance of the function f , and ideally $f(\bar{\mathbf{r}}) = \text{const.}$

The generation of random variables from an arbitrary high dimensional probability density π is a highly nontrivial task, which usually corresponds to the knowledge of the distribution's integral, *ie* its normalization. One way of overcoming this difficulty comes from the Metropolis-Hastings algorithm [126, 68], which allows us to generate variables from the probability distribution $\pi(\bar{\mathbf{r}})/\int \pi(\bar{\mathbf{r}})d\bar{\mathbf{r}}$ through a random walk. Starting from a random position $\bar{\mathbf{r}}$, a trial move is proposed $\bar{\mathbf{r}}' \leftarrow \bar{\mathbf{r}}$ with a transition probability $T(\bar{\mathbf{r}}', \bar{\mathbf{r}})$. This transition probability is usually assumed to be a multi-dimensional Gaussian centered at $\bar{\mathbf{r}}$ with a variance given by the constant Δt

$$\bar{\mathbf{r}}' = \bar{\mathbf{r}} + \boldsymbol{\eta}\sqrt{\Delta t}, \quad (3.43)$$

where $\boldsymbol{\eta}$ is a N -dimensional vector extracted from a Gaussian distribution with unitary variance and centered at the origin. This random walk corresponds to the Markov process since the new step depends only on the current position and is independent of the history of the walk. The acceptance probability associated with the proposed trial move is given by the ratio

$$A(\bar{\mathbf{r}}', \bar{\mathbf{r}}) = \min \left(1, \frac{T(\bar{\mathbf{r}}, \bar{\mathbf{r}}')\pi(\bar{\mathbf{r}}')}{T(\bar{\mathbf{r}}', \bar{\mathbf{r}})\pi(\bar{\mathbf{r}})} \right), \quad (3.44)$$

for which if $A(\bar{\mathbf{r}}', \bar{\mathbf{r}}) = 1$ the move is always accepted, while for $A(\bar{\mathbf{r}}', \bar{\mathbf{r}}) < 1$, the move is accepted with the probability $A(\bar{\mathbf{r}}', \bar{\mathbf{r}})$. In practice, this is achieved through the comparison to a randomly generated number μ from the uniform distribution $[0, 1]$, so that

the move is accepted when $A(\bar{\mathbf{r}}', \bar{\mathbf{r}}) > \mu$, and rejected otherwise. If the move is accepted the new position $\bar{\mathbf{r}}'$ becomes the new starting point for the next proposed trial move, while if the move is rejected, the old position $\bar{\mathbf{r}}$ remains the starting point for a new trial move. It can be proven that after an initial thermalization, the generated configurations $\{\bar{\mathbf{r}}_i\}_{i=1}^{\mathcal{N}}$ are distributed according to the probability density $\pi(\bar{\mathbf{r}})/\int \pi(\bar{\mathbf{r}}')d\bar{\mathbf{r}}'$.

In order to make the sampling more efficient via better exploration of the configuration space, it is possible in these integration algorithms to take advantage of the ergodicity property, generating \mathcal{M} independent initial configurations, *ie* ‘walkers’, each consisting of its position vector $\bar{\mathbf{r}}$. By evolving the set of walkers independently, accumulating for each a number \mathcal{N} of consecutive steps after thermalization, the value of the integral can be calculated as the average over ‘time’ (\mathcal{N}) and ‘space’ (\mathcal{M}) simultaneously

$$I \approx \langle f \rangle_{\pi(\bar{\mathbf{r}})/\int \pi(\bar{\mathbf{r}}')d\bar{\mathbf{r}}'} = \frac{1}{\mathcal{M}\mathcal{N}} \sum_{i=1}^{\mathcal{M}} \sum_{j=1}^{\mathcal{N}} f(\bar{\mathbf{r}}_{ij}), \quad (3.45)$$

exploiting the perfect parallelization that makes the integration particularly efficient on modern high-performance computing (HPC) facilities.

Here we must add that, by definition, the \mathcal{N} consecutive steps of the random walk are not fully independent since they depend on the previous position and are thus affected by the autocorrelation error. This error is inversely proportional to the increase of the step amplitude Δt . Yet, the increase of Δt is also inversely proportional to the increase of the rejected moves during the dynamics. On the other hand, small Δt decreases the number of rejected moves but increases the autocorrelation error. In order to minimize the autocorrelation of the variables, with an acceptable acceptance ratio, the rule-of-the-thumb consists of tuning Δt so that the acceptance probability of the moves remains approximately around 50%. The remaining autocorrelation error can then be removed by reweighing the variance σ_f^2 by the estimation of autocorrelation time, or through reblocking methods. In all our calculations, we apply the second more straightforward approach that is more efficient in the estimations of the generalized forces necessary for the optimization protocols described below.

3.5.2 Variational Monte Carlo

The variational Monte Carlo (VMC) method is based on the stochastic integration of the energy functional of a given Hamiltonian \hat{H} over the chosen many-body variational *ansatz* $\Psi_T(\bar{\mathbf{r}})$

$$E[\Psi_T] = \frac{\int \Psi_T^*(\bar{\mathbf{r}})\hat{H}\Psi_T(\bar{\mathbf{r}})d\bar{\mathbf{r}}}{\int |\Psi_T(\bar{\mathbf{r}})|^2 d\bar{\mathbf{r}}}, \quad (3.46)$$

which is achieved by separating the integrand into the product of two functions

$$E[\Psi_T] = \int E_{\text{loc}}(\bar{\mathbf{r}})\Pi(\bar{\mathbf{r}})d\bar{\mathbf{r}}, \quad (3.47)$$

that are respectively the probability density $\Pi(\bar{\mathbf{r}}) = \frac{|\Psi_T(\bar{\mathbf{r}})|^2}{\int |\Psi_T(\bar{\mathbf{r}})|^2 d\bar{\mathbf{r}}}$ of finding the system in the configuration $\bar{\mathbf{r}}$ and the local energy $E_{\text{loc}}(\bar{\mathbf{r}}) = \frac{\hat{H}\Psi_T(\bar{\mathbf{r}})}{\Psi_T(\bar{\mathbf{r}})}$ that is the energy of the system in that particular configuration. The integration in Eq. 3.47 is obtained through a random walk (or an arbitrary number of parallel random walks) evolved according to the Metropolis-Hastings algorithm [126, 68] described in Sec. 3.5.1. By computing the local quantities, such as the energy $E_{\text{loc}}(\bar{\mathbf{r}})$, for each configuration, one can get a statistical estimation of the total energy through the mean value of the corresponding local quantities accumulated over the entire evolution

$$E[\Psi_T] \approx \langle E_{\text{loc}} \rangle_{\mathcal{N}} \pm \sqrt{\frac{\langle E_{\text{loc}}^2 \rangle_{\mathcal{N}} - \langle E_{\text{loc}} \rangle_{\mathcal{N}}^2}{\mathcal{N}}}, \quad (3.48)$$

with an error that decreases as $1/\sqrt{\mathcal{N}}$, and is proportional to the square root of the variance $\langle E_{\text{loc}}^2 \rangle_{\mathcal{N}} - \langle E_{\text{loc}} \rangle_{\mathcal{N}}^2$.

The extension of the VMC algorithm to integrate a mixed system of drudons and electrons is rather straightforward. In our approach, the two sets of particles, the positions of which are stored in each walker, are diffused particle-by-particle in a random order starting from the electrons, according to the Metropolis-Hastings algorithm [126, 68]. Each particle's trial move is proposed according to

$$\mathbf{r}'_i = \mathbf{r}_i + \sqrt{\Delta_i}\boldsymbol{\eta}, \quad (3.49)$$

where $\boldsymbol{\eta}$ is a 3-dimensional vector of Gaussian distributed random numbers with zero mean and unitary variance, and Δ_i is an amplitude that depends on the type of particle

and is defined as:

$$\Delta_i = \begin{cases} \delta_e/m_e & \text{for } i \in [1, N_e] \\ \delta_d/\mu_i & \text{for } i \in [N_e + 1, N_e + N_d] \end{cases} . \quad (3.50)$$

The parameters δ_e and δ_d are two amplitudes used respectively for the electrons and for the drudons, and m_e and μ_i are respectively the mass of the electrons (which is equal to 1) and the mass of the i th drudon. δ_e and δ_d are optimized during the thermalization of the VMC run by converging the acceptance probability of the moves to the value of 50%, which is the rule of thumb discussed in the previous section. This procedure is repeated until \mathcal{N} consecutive configurations are sampled.

3.5.3 Wavefunction optimization methods

Within the VMC framework, it is possible to optimize the trial wavefunction through energy or variance minimization [187, 99, 66, 184, 42, 170, 186, 182, 168], exploiting the variational principle discussed at the beginning of this Chapter, or the zero-variance principle that holds for VMC, in which the variance of the local energy is null when the trial wavefunction corresponds to an eigenstate of the Hamiltonian, *ie.* $E_{\text{loc}}(\bar{\mathbf{r}}) = \frac{\hat{H}\Psi_i(\bar{\mathbf{r}})}{\Psi_i(\bar{\mathbf{r}})} = \frac{E_i\Psi_i(\bar{\mathbf{r}})}{\Psi_i(\bar{\mathbf{r}})} = E_i$.

In this work, the set of parameters is optimized through the Stochastic Reconfiguration procedure described in Refs. 169, 170 with the use of the correlated sampling (CS) technique[50] in order to better estimate the energy variation in between parameter updates and to automatically tune the prefactor Δ of the parameters' variation amplitude. The wavefunction's sampling is automatically recomputed after a fixed number of CS steps or if the overlap between two consecutive wavefunctions becomes lower than a chosen threshold.

In the Stochastic Reconfiguration algorithm, the vector of variational parameters $\boldsymbol{\alpha}$ is updated after each step of the optimization as

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha} + \Delta\mathbf{S}^{-1}\mathbf{f}_{\boldsymbol{\alpha}}, \quad (3.51)$$

where Δ is an empirical parameter that scales the change of the parameters, $\mathbf{f}_{\boldsymbol{\alpha}}$ is a

vector of generalized forces

$$f_{\alpha_k} = -2 \{ \langle E_{\text{loc}} \mathcal{O}_{\alpha_k} \rangle_{\Pi} - \langle E_{\text{loc}} \rangle_{\Pi} \langle \mathcal{O}_{\alpha_k} \rangle_{\Pi} \}, \quad (3.52)$$

being $\mathcal{O}_{\alpha_k} = \frac{\partial}{\partial \alpha_k} \ln[\Psi_T(\boldsymbol{\alpha})]$ the derivative of the logarithm of the wavefunction with respect to the parameter α_k , and \mathbf{S} is the covariance matrix between the \mathcal{O}_{α_k} components

$$\mathbf{S}_{kl} = \langle \mathcal{O}_{\alpha_k} \mathcal{O}_{\alpha_l} \rangle_{\Pi} - \langle \mathcal{O}_{\alpha_k} \rangle_{\Pi} \langle \mathcal{O}_{\alpha_l} \rangle_{\Pi}. \quad (3.53)$$

In practice, the variation of the parameters $\delta\boldsymbol{\alpha}$ is calculated using the set of linear equations

$$\mathbf{S}\delta\boldsymbol{\alpha} = \mathbf{f}_{\boldsymbol{\alpha}} \quad (3.54)$$

via a parallel version of the conjugate gradient method, for which \mathbf{S} is never explicitly computed and is preconditioned.

3.5.4 Diffusion Monte Carlo

As previously discussed in the VMC method, the quality of the description of the ground state and its energy strictly depends on the parametrization of the trial wavefunction $\Psi_T(\bar{\mathbf{r}}; \bar{\alpha})$ [52, 95, 14]. In order to obtain a more accurate estimation of the physical observable, overcoming the limitations of the trial wavefunction, also improving the description of the dynamical correlation between the particles in the system, it is possible to use projection schemes such as the diffusion Monte Carlo (DMC) method.

The DMC method is based on the transformation of the time-dependent Schrödinger equation [154]

$$i \frac{\partial}{\partial t} \varphi(\bar{\mathbf{r}}, t) = \hat{H} \varphi(\bar{\mathbf{r}}, t) \quad (3.55)$$

into a diffusion equation in the imaginary time ($t = -i\tau$)

$$\frac{\partial \varphi(\bar{\mathbf{r}}, \tau)}{\partial \tau} = \sum_{i=1}^N \frac{1}{2\mu_i} \nabla_i^2 \varphi(\bar{\mathbf{r}}, \tau) - (V(\bar{\mathbf{r}}) - E_R) \varphi(\bar{\mathbf{r}}, \tau), \quad (3.56)$$

where N is the total number of particles and E_R is a constant energy shift whose role will be made clear below. The solutions of this imaginary time equation can be written

as a linear combination

$$\varphi(\bar{\mathbf{r}}, \tau) = \sum_{k=0}^{\infty} a_k e^{-(E_k - E_R)\tau} \Psi_k(\bar{\mathbf{r}}) \quad (3.57)$$

of the stationary states $\Psi_k(\bar{\mathbf{r}})$ of the time-independent Schrödinger equation, each multiplied by an exponential function that decays as $\tau \rightarrow \infty$ according to the energy differences $E_k - E_R$. Here, the role of the constant energy shift E_R becomes evident [107, 147]. If $E_R \approx E_0$, where E_0 is the ground state's energy, by letting the solution evolve for a reasonable amount of time, the only component that will survive in the expansion 3.57 will be the one corresponding to $\Psi_0(\bar{\mathbf{r}})$. As a matter of fact, all the excited states with $E_k > E_R$ will decay faster in time, while all the states with $E_k \leq E_R$ will gain more and more weight. In general, we must add that whatever the choice of E_R , the weight of the ground state with respect to the others will always grow exponentially in the imaginary time τ , so that for $\tau \rightarrow \infty$ the algorithm will extract the ground state component $\Psi_0(\bar{\mathbf{r}})$ from a general initial wavefunction $\varphi(\bar{\mathbf{r}}, \tau_0)$, given that the latter one is not orthogonal to the former.

For fermionic systems, or a mixed system of fermions and distinguishable particles such as that in the El-QDO method, the general wavefunction $\varphi(\bar{\mathbf{r}}, \tau_0)$ changes the sign and can not be sample directly as a distribution function. In order to solve this sign problem, it is thus necessary to introduce the fixed-node (FN) approximation [147, 107, 132], for which the propagated state $\varphi(\bar{\mathbf{r}}, \tau)$ is replaced by the mixed distribution

$$f(\bar{\mathbf{r}}, \tau) = \varphi(\bar{\mathbf{r}}, \tau) \Psi_T(\bar{\mathbf{r}}), \quad (3.58)$$

and the nodes (the region in which the function changes the sign) are fixed by Ψ_T during the entire propagation. The time evolution of $f(\bar{\mathbf{r}}, \tau)$ is obtained from Eq. 3.56 by multiplying both sides by $\Psi_T(\bar{\mathbf{r}})$ from the right, leading to the equation

$$\frac{\partial}{\partial \tau} f(\bar{\mathbf{r}}, \tau) = \sum_{i=1}^N \frac{1}{m_i} \left[\frac{1}{2} \nabla_i^2 f(\bar{\mathbf{r}}, \tau) - \nabla_i (f(\bar{\mathbf{r}}, \tau) \bar{\mathbf{v}}(\bar{\mathbf{r}})) \right] - (E_{\text{loc}}(\bar{\mathbf{r}}) - E_R) f(\bar{\mathbf{r}}, \tau), \quad (3.59)$$

that now includes two quantities that only depend on the trial wavefunction, that are respectively the local energy $E_{\text{loc}}(\bar{\mathbf{r}})$, previously defined for the VMC algorithm, and the drift velocity

$$\bar{\mathbf{v}}(\bar{\mathbf{r}}) = \frac{\nabla \Psi_T(\bar{\mathbf{r}})}{\Psi_T(\bar{\mathbf{r}})}. \quad (3.60)$$

In order to define the stochastic integration of the diffusion process described above, we consider the time propagation in the integral form, which is written through the Green's function $G(\bar{\mathbf{r}}', \tau; \bar{\mathbf{r}}, \tau_0)$ associated to the given Hamiltonian

$$f(\bar{\mathbf{r}}', \tau) = \int_V G(\bar{\mathbf{r}}', \tau; \bar{\mathbf{r}}, \tau_0) f(\bar{\mathbf{r}}, \tau_0) d\bar{\mathbf{r}}. \quad (3.61)$$

Since the exact Green's function $G(\bar{\mathbf{r}}', \tau; \bar{\mathbf{r}}, \tau_0)$ of the time propagation of the drift-diffusion process defined in Eq. 3.59 is unknown, it is necessary to approximate the problem by separating the action of the kinetic and potential energy operators in Eq. 3.59 which do not commute. A way of dealing with this problem is to discretize the time evolution between τ_0 and τ into \mathcal{M} steps of length $\delta\tau$ applying for each infinitesimal evolution the Trotter-Suzuki decomposition [54] for which the Green's function is approximated as

$$e^{-\delta\tau(\hat{T}+\hat{V}-E_R)} = e^{-(\hat{V}-E_R)\frac{\delta\tau}{2}} e^{-\hat{T}\delta\tau} e^{-(\hat{V}-E_R)\frac{\delta\tau}{2}} + \mathcal{O}(\delta\tau^3). \quad (3.62)$$

Thus, the approximated Green's function in the infinitesimal time interval will now be written as the product

$$G(\bar{\mathbf{r}}', \bar{\mathbf{r}}; \delta\tau) \approx P(\bar{\mathbf{r}}', \bar{\mathbf{r}}; \delta\tau) W(\bar{\mathbf{r}}', \bar{\mathbf{r}}; \delta\tau)$$

of a separated diffusion function process¹

$$P(\bar{\mathbf{r}}', \bar{\mathbf{r}}; \delta\tau) = \prod_{i=1}^N \left(\frac{\mu_i}{2\pi\delta\tau} \right)^{3/2} e^{-\frac{\mu_i}{2\delta\tau} \left[\mathbf{r}'_i - \mathbf{r}_i - \frac{\delta\tau}{m_i} \mathbf{v}_i(\bar{\mathbf{r}}) \right]^2} \quad (3.63)$$

times a reweighting or growth/decay function

$$W(\bar{\mathbf{r}}', \bar{\mathbf{r}}; \delta\tau) = e^{-\frac{\delta\tau}{2}(S(\bar{\mathbf{r}}') + S(\bar{\mathbf{r}}))}, \quad (3.64)$$

where $S(\bar{\mathbf{r}}) = E_{\text{loc}}(\bar{\mathbf{r}}) - E_R$, so that the full propagation in time will correspond to \mathcal{M} consecutive applications of the two processes

$$f(\bar{\mathbf{r}}', \tau) = \int_V \left[\prod_{m=1}^{\mathcal{M}} d\bar{\mathbf{r}}^{(m-1)} \delta(\bar{\mathbf{r}}' - \bar{\mathbf{r}}^{(\mathcal{M})}) \times \right. \\ \left. \times P(\bar{\mathbf{r}}^{(m)}, \bar{\mathbf{r}}^{(m-1)}; \delta\tau) W(\bar{\mathbf{r}}^{(m)}, \bar{\mathbf{r}}^{(m-1)}; \delta\tau) \right] f(\bar{\mathbf{r}}^{(0)}, \tau_0). \quad (3.65)$$

¹The diffusion function in Eq. 3.63 is defined as the product of single-particle consecutive diffusion processes. We must point out that the drift velocity of the i th particle depends on all the particles' positions, thus every time one particle is moved, all the components of the drift velocity have to be updated.

Through the Trotter-Suzuki decomposition, we introduce a time-step error that vanishes as the discretization step is extrapolated to the limit of continuum $\delta\tau \rightarrow 0$.

In practice, the FN-DMC algorithm implemented to integrate both electronic and drudonic degrees of freedom is based on the established method by Umrigar *et al.* described in Ref. 185. In the following paragraphs, we will concentrate on describing the most important details and generalizations of our approach. Let us assume we have \mathcal{N} walkers each representing an electronic and drudonic configuration taken from a VMC random walk (thus distributed according to $|\Psi_T(\bar{\mathbf{r}})|^2 = f(\bar{\mathbf{r}}, \tau_0)$ with initial weights $W_0 = 1$). During the drift/diffusion process, we update the particles' positions with a particle-by-particle scheme such that from time step m to $m+1$ we will have

$$\mathbf{r}_i^{(m+1)} = \mathbf{r}_i^{(m)} + \frac{\delta\tau}{\mu_i} \tilde{\mathbf{v}}_i(\bar{\mathbf{r}}^{(m)}) + \sqrt{\frac{\delta\tau}{\mu_i}} \boldsymbol{\eta}, \quad (3.66)$$

where $\boldsymbol{\eta}$ is a 3-dimensional vector of random variables extracted from a Gaussian distribution with zero mean value and unitary variance, and $\tilde{\mathbf{v}}_i(\bar{\mathbf{r}}^{(m)})$ is the drift velocity rescaled according to the procedure introduced by Umrigar *et al.* [185] to avoid divergences near the nodal surface. The single-particle move is accepted with probability

$$\mathcal{A}(\bar{\mathbf{r}}^{(m+1)}, \bar{\mathbf{r}}^{(m)}; \delta\tau) = \min \left[1, \frac{|\Psi_T(\bar{\mathbf{r}}^{(m+1)})|^2 P(\bar{\mathbf{r}}^{(m)}, \bar{\mathbf{r}}^{(m+1)}; \delta\tau)}{|\Psi_T(\bar{\mathbf{r}}^{(m)})|^2 P(\bar{\mathbf{r}}^{(m+1)}, \bar{\mathbf{r}}^{(m)}; \delta\tau)} \right], \quad (3.67)$$

and always rejected when a particle crosses a node ($\frac{\Psi_T(\bar{\mathbf{r}}^{(m+1)})}{\Psi_T(\bar{\mathbf{r}}^{(m)})} < 0$). Once the trial move is proposed for all the particles the weight of each walker is updated through the term [185]

$$W(\bar{\mathbf{r}}^{(m+1)}, \bar{\mathbf{r}}^{(m)}; \delta\tau) = \begin{cases} e^{-\frac{\delta\tau_{\text{eff}}}{2} (S(\bar{\mathbf{r}}^{(m)}) + S(\bar{\mathbf{r}}^{(m+1)}))}, & \text{accepted move} \\ e^{-\delta\tau_{\text{eff}} S(\bar{\mathbf{r}}^{(m)})}, & \text{rejected move} \end{cases}, \quad (3.68)$$

where we introduce an effective time step [185] defined as the average over the single particle moves in a single time step

$$\delta\tau_{\text{eff}} = \delta\tau \frac{\sum_{i=1}^N \mathcal{A}_i(\bar{\mathbf{r}}^{(m+1)}, \bar{\mathbf{r}}^{(m)}; \delta\tau) \left(\bar{\mathbf{r}}_i^{(m+1)} - \bar{\mathbf{r}}_i^{(m)} \right)^2}{\sum_{i=1}^N \left(\bar{\mathbf{r}}_i^{(m+1)} - \bar{\mathbf{r}}_i^{(m)} \right)^2}. \quad (3.69)$$

To avoid numerical instabilities near the nodal surface of the trial wavefunction for finite time steps $\delta\tau$, we use the cutoff $E_{\text{cut}} = \alpha\sqrt{N/\delta\tau}$ introduced by Zen *et al.* in

Ref. 209, where α is a tunable parameter, here set to 0.2. The local energy that appears in the branching factors $\bar{S}(\bar{\mathbf{r}}) = \bar{E}_{\text{loc}}(\bar{\mathbf{r}}) - E_R$ is thus redefined as

$$\bar{E}_{\text{loc}}(\bar{\mathbf{r}}) = E_{\text{MA}} + \text{sign} [E_{\text{loc}}(\bar{\mathbf{r}}) - E_{\text{MA}}] \times \min\{E_{\text{cut}}, |E_{\text{loc}}(\bar{\mathbf{r}}) - E_{\text{MA}}|\}, \quad (3.70)$$

where E_{MA} is the mixed average estimation of the total energy during the DMC run.

After the weights have been updated we execute the branching procedure through the stochastic reconfiguration [24] only if the total weight fluctuation is larger than 5% of the target weight, or if there is one walker with a weight lower than 0.50 or larger than 1.50. If neither of the conditions are verified the code postpones the branching. Clearly, to control the total weights the reference energy is updated after each DMC step through the formula[185]

$$E_R^{(m+1)} = E_{\text{MA}}^{(m)} - C \frac{\delta\tau}{\delta\tau_{\text{MA}}} \log \frac{W_m}{W_0}, \quad (3.71)$$

where C is a constant (usually set to 1) and $\delta\tau_{\text{MA}}$ is the mixed average estimation of the effective time step defined in Eq. 3.69.

Finally, in this work, to treat the atoms of the electronic systems we use pseudopotentials in order to reduce the computational cost of the calculations. Within the FN-DMC algorithm, the integration of the non-local part is usually treated through the Locality approximation [133] or T-move [27]. Here, to reduce the possible dependency of the results on the Jastrow factor, we apply the Determinant Locality approximation (DLA) [208] for which the non-local operator is evaluated using only the Slater determinant part of the many-body wavefunction. The interaction between the drudons and the pseudopotentials is reduced to the Coulomb potential, using only the effective charge of the nuclei.

3.6 Wavefunctions for many-electron and many-drudon systems

The El-QDO Hamiltonian described in Eq. 2.58 depends explicitly on the $3N_e$ electronic $\bar{\mathbf{r}}^e$ and $3N_d$ drudonic $\bar{\mathbf{r}}^d$ coordinates, while the coordinates of the nuclei $\bar{\mathbf{R}}^n$ and of the

QDOs' centers $\bar{\mathbf{R}}^O$ are assumed to be fixed parameters within the Born-Oppenheimer approximation.

In order to approximately describe the ground state of such a system with the QMC methods, taking the full advantage of their ability to integrate explicitly correlated wavefunctions that include the particles' distances, we construct a variational *ansatz* as a product of three independent terms

$$\Psi_{\text{EL-QDO}}(\bar{\mathbf{r}}^e, \bar{\mathbf{r}}^d; \bar{\mathbf{R}}^n, \bar{\mathbf{R}}^O) = \Psi_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) \Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O) J_{e-d}(\bar{\mathbf{r}}^e, \bar{\mathbf{r}}^d; \bar{\mathbf{R}}^n, \bar{\mathbf{R}}^O), \quad (3.72)$$

which correspond to the electronic wavefunction $\Psi_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)$, the pure QDO wavefunction $\Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O)$ and a positive-defined function $J_{e-d}(\bar{\mathbf{r}}^e, \bar{\mathbf{r}}^d; \bar{\mathbf{R}}^n, \bar{\mathbf{R}}^O)$ which has the purpose of describing the correlation effects between the electrons and drudons. For the calculations of the pure electronic or pure QDO systems only the corresponding $\Psi_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)$ or $\Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O)$ terms are used.

In the next subsections, we will discuss in more detail the various functional forms of these three terms employed in our calculations.

3.6.1 Electronic wavefunction

The electronic part Ψ_e of the total wavefunction, which depends only on the electronic coordinates, and on the nuclear positions that are parameters of the system, is the most commonly used Slater-Jastrow wavefunction, written as the product

$$\Psi_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = \det[\mathbf{S}^\uparrow] \det[\mathbf{S}^\downarrow] J_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) \quad (3.73)$$

of the Slater determinants of the \mathbf{S}^\uparrow and \mathbf{S}^\downarrow matrices of the molecular orbitals occupied respectively by the spin up and spin down electrons, and the electronic Jastrow factor $J_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)$ [41].

The molecular orbitals that define the elements of the Slater matrix $\mathbf{S}(\bar{\mathbf{r}}^e)$ are written as linear combinations

$$\varphi_k(\mathbf{r}^e) = \sum_{q=1}^{Q_S} c_q^k \phi_q(\mathbf{r}^e) \quad (3.74)$$

of Q_S contracted Gaussian-type orbitals $\phi_q(\mathbf{r})$ centered on the various nuclei of the electronic system. The initial values of the linear coefficients of the molecular orbital expansions are in our case obtained from initial DFT calculations.

The positive-definite electronic Jastrow factor[41] used in this work is inspired by the one introduced by Casula *et al.* in Ref. 26, as a sum of two terms

$$J_e(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = e^{\mathcal{J}_2(\bar{\mathbf{r}}^e) + \mathcal{J}_{3/4}(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)}, \quad (3.75)$$

which can be classified as a pure homogeneous two-body term $\mathcal{J}_2(\bar{\mathbf{r}}^e)$ and a three/four-body inhomogeneous term $\mathcal{J}_{3/4}(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n)$ that is used to describe the fermionic pair correlations in the field of the nuclei.

The homogeneous two-body Jastrow factor is written as the sum of functions depending only on the distances between electron pairs (with dropped superscript 'e' for r_{ij})

$$\mathcal{J}_2(\bar{\mathbf{r}}^e) = \sum_{j>i=1}^{N_e} f_{ee}(r_{ij}), \quad (3.76)$$

where the pair correlation functions are written as

$$f_{ee}(r_{ij}) = \begin{cases} -\frac{1}{4b^p(1+b^p r_{ij})} + \sum_{n=1}^N g_n^p e^{-\zeta_n^p r_{ij}^2} & \text{undis.} \\ -\frac{1}{2b^a(1+b^a r_{ij})} + \sum_{n=1}^N g_n^a e^{-\zeta_n^a r_{ij}^2} & \text{dis.} \end{cases} \quad (3.77)$$

respectively for distinguishable (antiparallel spin) electrons and indistinguishable ones (parallel spin). The variational parameters b^p and b^a are related to the cusp functions and are optimized independently[80]. The additional linear combination of Gaussian-type orbitals works as a re-modulating factor which depends on the set of coefficients g_n^p and g_n^a and exponents ζ_n^p and ζ_n^a that are optimized.

The non-homogeneous three/four-body term is written as the linear combination of products of two atomic orbitals

$$\mathcal{J}_{3/4}(\bar{\mathbf{r}}^e; \bar{\mathbf{R}}^n) = \sum_{j>i=1}^{N_e} \sum_{q,p=1}^{Q_{\mathcal{J}}} \gamma_{qp} \chi_q(\mathbf{r}_i^e) \chi_p(\mathbf{r}_j^e), \quad (3.78)$$

where $\chi_q(\mathbf{r}^e)$ are non-normalized atomic orbitals and the parameters γ_{qp} describe the correlation of two electrons in the field of the nuclei. The terms in the sum of Eq. 3.78

with the orbitals χ_q and χ_p both centered at the same nucleus represent the three-body interactions and the terms with the orbitals χ_q and χ_p centered at distinct nuclei represent the four-body interactions. Since the Jastrow factor must be symmetric with respect to the exchange of all the electrons, the γ_{qp} parameters satisfy the condition $\gamma_{qp} = \gamma_{pq}$.

The electronic Jastrow factor introduces dynamical correlations into the wavefunction at the level of VMC. Due to its bosonic nature, it does not affect the nodal surface of the total wavefunction and does not affect the final DMC energy. Yet, its introduction and optimization are useful since they reduce the variance of the observable accumulated during the sampling process, thus increasing the efficiency of the computations.

The free variational parameters of the electronic part of the wavefunction are the expansion coefficients in the molecular orbitals, the contraction coefficients and the exponents in the Gaussian atomic orbitals, the cusp-related parameters, expansion coefficients and exponents in f_{ee} (Eq. 3.77), and the expansion coefficients with the exponents of the orbitals in $\mathcal{J}_{3/4}$ (Eq. 3.78). These can be all optimized at the VMC level, using the optimization algorithms described in Sec. 3.5.3.

We would like to point out that the electronic wavefunction used in this Thesis is not the only possible choice. It is in fact possible to use also other alternative *ansätze* such as multi-determinant expansions, Geminal or Pfaffian wavefunctions among others [195, 102, 28], and the electronic Jastrow can have almost an arbitrary functional form, with many examples available in the literature [195, 102]. In recent years also Neural Network techniques [74] have been developed to represent highly accurate *ansätze* for molecular systems, yet the application of such computationally expensive approaches is beyond the scope of this Thesis.

Another important note is that the electronic Jastrow factor in Eq. 3.75 does not contain the one-body cusp functions due to the divergence of the electron-nuclei Coulomb potential, since all the QMC calculations executed in this Thesis substitute the nuclei and core electrons of the atoms with effective core potentials [198, 16, 10, 15].

3.6.2 Drudonic wavefunction

The QDO part $\Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O)$ of the wavefunction depends only on the drudonic coordinates and parametrically on the QDO centers.

In previous works that approached the solution of the QDO model with the full Coulomb potential, the authors focused essentially on the correct description of the large distance regimes that characterize dispersion interactions [152, 88, 31]. At those distances, the solution of the QDO model can be written as a perturbation of the solutions of non-interacting QDOs, represented by isotropic Gaussian orbitals which are deformed by additional couplings. For this reason, in the CI approach of M. Sadhukhan and F. R. Manby [152] the authors constructed the variational space of the homogeneous QDO dimer as a linear combination of a set of single QDO excitations of the two localized oscillators, highlighting the convergence difficulties in the short-range limit, which will be shown in Chap. 5. On the other hand, in the QMC approach from G. J. Martyna and co-workers [87, 88] the authors used a trial wavefunction built from the product of the solution of non-interacting harmonic oscillators adding a fixed three-body correlation function constructed analytically to reproduce the long-range interactions between the QDOs that are responsible for the reciprocal polarization.

The first *ansatz*, referred to as Dip. in Chap. 5 and used in the calculations of the El-QDO approach in Chap. 4, is inspired by the exact solution of the QDO model within the dipole approximation and is written as the exponential function

$$\Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O) = \exp [\bar{\mathbf{r}}_{dO}^\top \mathbf{A} \bar{\mathbf{r}}_{dO}] \quad (3.79)$$

of the vector-matrix-vector product between the vector $\bar{\mathbf{r}}_{dO} = \bar{\mathbf{r}}^d - \bar{\mathbf{R}}^O$ of the $3N_d$ components of the distances between each drudon and its center, and the square symmetric matrix \mathbf{A} containing $3N_d(3N_d + 1)/2$ independent parameters. While the diagonal elements of the parameters' matrix \mathbf{A} contain the Gaussian solutions of the isolated QDOs, and upon relaxation they describe the on-site polarization of the oscillators, the off-diagonal elements represent the coupling of the various oscillators along the 3-dimensions. This wavefunction converges to the exact solution of the QDO model in the limit of large spatial separations between the QDOs, where the dipole potential becomes a better and better approximation of the Coulomb potential. Interestingly,

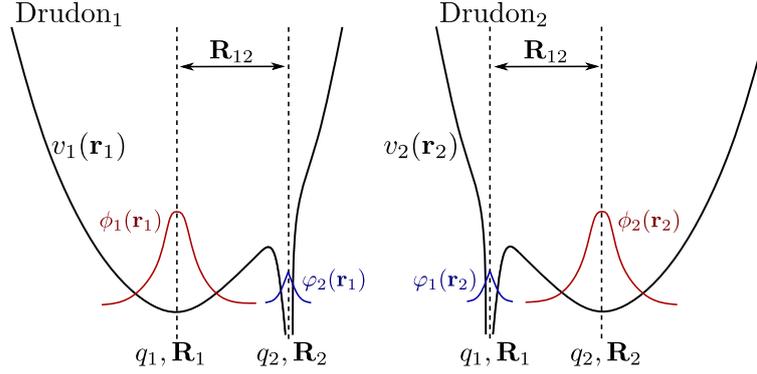


Figure 3.1: Schematic representations of the one-body potentials $v_1(\mathbf{r}_1)$ and $v_2(\mathbf{r}_2)$ of each drudon particle for a system of two interacting QDOs. $\phi_1(\mathbf{r})$ and $\phi_2(\mathbf{r})$ are the Gaussian functions on the center of each QDO, that represent the exact solution in the limit of dissociation. $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ are the 1s exponential Slater orbitals that describe the states of the drudons when drifting towards the attractive Coulomb potential on the opposite center.

when used for the QDOs interacting via the full Coulomb potential, it possesses multipolar contributions, beyond the dipole potential case. As it will be shown later in Chap. 4, it can serve as a very reasonable and simple first approximation of the QDO wavefunction in the region characteristic for dispersion interactions.

Despite the fact that for distances characteristic of the dispersion interactions this *ansatz* is a good approximation, it fails to converge to the proper state for the QDO model for all values of the QDOs' distances since it does not take into account the form of the single-particle potential energies in Fig. 3.1. In order to construct a more general and accurate solution to the Hamiltonian defined in Eq. 2.41, here we want to address all the main aspects and properties that the solution should incorporate, following a similar approach to the one used for many-electron systems in the basis of molecular orbitals [70, 164, 20].

First, let us define the distances between a drudon and the various QDO centers $d_{ij} = |\mathbf{r}_i^d - \mathbf{R}_j^O|$, the distances between the drudonic particles and their center of oscillation $d_i = |\mathbf{r}_i^d - \mathbf{R}_i^O|$ and the distances between two drudons $r_{ij} = |\mathbf{r}_i^d - \mathbf{r}_j^d|$ that will be used in the following sections. If we consider the i th drudon, since the quantum particle interacts with its center through the harmonic potential, it is clear that the solution in this region should be similar to a Gaussian primitive function of the type $\phi_i(\mathbf{r}_i^d) \propto \exp(-\frac{\mu_i \omega_i d_i^2}{2})$. Yet, the same drudon will interact with all the other $N - 1$

QDO centers via the Coulomb potential, meaning that the wavefunction must satisfy the cusps' conditions [97]:

$$\frac{1}{\langle \Psi_d \rangle} \left. \frac{\partial \langle \Psi_d \rangle}{\partial d_{ij}} \right|_{d_{ij}=0} = q_i q_j \mu_j \quad \forall \quad j \neq i, \quad (3.80)$$

where $\langle \Psi_d \rangle$ represents the angular average of the wavefunction. Thus, around the j th centers the single-particle orbital of the i th drudon will have the form of an exponential function proportional to $\varphi_j(\mathbf{r}_i^d) \propto \exp(-q_i q_j \mu_j d_{ij})$.

Considering these characteristics, the approximate single-particle orbitals should be written as a linear combination of Gaussian and Slater-type functions (see Fig. 3.1)

$$\Phi^i(\mathbf{r}_i^d) \approx \alpha_i^i e^{(-\frac{\mu_i \omega_i}{2} d_i^2)} + \sum_{j \neq i}^N \alpha_j^i e^{(-q_i q_j \mu_j d_{ij})}, \quad (3.81)$$

where α_j^i is a matrix of parameters, with indices $i, j \in [1, N]$. These orbitals can be generalized following the same approach used to define electronic molecular orbitals, and thus in this work, they are written as a linear combination

$$\Phi^i(\mathbf{r}_i^d) = \sum_{q=1}^Q \alpha_q^i \phi_q(\mathbf{r}_i^d), \quad (3.82)$$

where $\phi_q(\mathbf{r}_i^d)$ are Q atomic-like orbitals that constitute the basis set of the system of QDOs that are written as a linear combination of Gaussian-type orbitals (GTO), centered at the positions of the various QDOs $\bar{\mathbf{R}}^O$, whose linear and exponential parameters are fully optimized. In particular, in all our calculations each QDO is represented through a (3s1s)/[1s1s] contracted basis set, where the uncontracted 1s orbital is used to describe the Gaussian solution around the quadratic potential and the contracted (3s)/[1s] orbital is used to describe the Slater solution centered on the Coulomb attractive potential.

The use of GTOs in place of exponential functions clearly introduces an error in the description of the one-body cusp in Eq. 3.80 that can be eliminated through a drudonic Jastrow factor, which is normally used in quantum Monte Carlo trial wavefunctions, of the form

$$J_1(\bar{\mathbf{r}}^d) = \exp \left(\sum_{i \neq j}^N q_i q_j \mu_j f_{1b}(d_{ij}; \gamma) \right), \quad (3.83)$$

where $f_{1b}(d_{ij}; \boldsymbol{\gamma})$ is a parametric function that only depends on the distance between two particles (in this case the i th drudon and the j th QDO center) that for $d_{ij} \rightarrow 0$ has the property of going linearly to zero, *ie* $f(d_{ij}) \approx d_{ij}$, and decays to zero as the distance increases $f(d_{ij}) \rightarrow 0$ as $d_{ij} \rightarrow \infty$.

Another property of the wavefunction will be the explicit correlation between the pairs of drudons, which is introduced through the two-body Coulomb potential $\frac{q_i q_j}{r_{ij}}$ in the Hamiltonian from Eq. 2.41. A first correlation function between drudonic pairs can be introduced considering that in the limit of two overlapping drudons the exact wavefunction has a two-body cusp of the form:

$$\frac{1}{\langle \Psi_d \rangle} \left. \frac{\partial \langle \Psi_d \rangle}{\partial r_{ij}} \right|_{r_{ij}=0} = q_i q_j \frac{\mu_i \mu_j}{\mu_i + \mu_j} \quad \forall \quad j \neq i. \quad (3.84)$$

This requisite can be satisfied again through a drudonic Jastrow factor of the type

$$J_2(\bar{\mathbf{r}}^d) = \exp \left(\sum_{i>j}^N q_i q_j \frac{\mu_i \mu_j}{\mu_i + \mu_j} f_{2b}(r_{ij}; \boldsymbol{\eta}) \right), \quad (3.85)$$

where $f_{2b}(r_{ij}; \boldsymbol{\eta})$ is a parametric function with similar properties to the one defined in Eq. 3.83 to reproduce the one-body cusp condition.

The total drudonic Jastrow factor can be thus written as $J_d(\bar{\mathbf{r}}^d) = J_1(\bar{\mathbf{r}}^d) J_2(\bar{\mathbf{r}}^d)$. From these considerations, by combining Eqs. 3.82, 3.83 and 3.85 we can write a first approximation to the explicitly correlated wavefunction as

$$\Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O) = \left[\prod_{i=1}^N \Phi^i(\mathbf{r}_i^d) \right] J_d(\bar{\mathbf{r}}^d), \quad (3.86)$$

that will be referred to as Mol.+J in Chap. 5, where the one-body cusp function is written as

$$f_{1b}(d_{ij}; \boldsymbol{\gamma}) = -\frac{e^{-\gamma_0 d_{ij}}}{\gamma_0} + \sum_{m=1}^M \gamma_m e^{-\gamma_{M+m} d_{ij}^2}, \quad (3.87)$$

and the two-body has a form inspired by the Padé's approximation[141]

$$f_{2b}(r_{ij}; \boldsymbol{\eta}) = -\frac{1}{\eta_0(1 + \eta_0 r_{ij})} + \sum_{m=1}^M \eta_m e^{-\eta_{M+m} r_{ij}^2}. \quad (3.88)$$

In Eqs. 3.87 and 3.88, the two vectors of $2M+1$ parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ are both optimized. In our calculations, we assume $M = 5$. Notice, that by construction, the functions f_{1b} and f_{2b} do not only satisfy the cusp conditions but also re-modulate the total wavefunction. This will be important in Chap. 5, in which we compare the performances of the various *ansätze*.

The third trial wavefunction for the QDOs considered in this Thesis is the dipole function from Eq. 3.79 enhanced with the drudonic Jastrow factors defined above:

$$\Psi_d(\bar{\mathbf{r}}^d; \bar{\mathbf{R}}^O) = \exp[\bar{\mathbf{r}}_{dO}^\top \mathbf{A} \bar{\mathbf{r}}_{dO}] J_d(\bar{\mathbf{r}}^d), \quad (3.89)$$

that will be referred to as Dip.+J.

An important consideration that should be made here is that due to the distinguishable nature of drudons, the DMC results do not depend on the different *ansätze* since the algorithm always converges to the correct ground state. This will be discussed in more detail in Chap. 4, confirming that already the simplest dipole *ansatz* from Eq. 3.79 is a very reasonable first approximation of the QDO wavefunction.

3.6.3 Electron-drudon coupling

The last part of the El-QDO wavefunction is the coupling $J_{e-d}(\bar{\mathbf{r}}^e, \bar{\mathbf{r}}^d; \bar{\mathbf{R}}^n, \bar{\mathbf{R}}^O)$ between two types of quantum particles in the El-QDO system, in the external field of the nuclei and of the QDOs' centers. This part of the wavefunction is inspired by the dipole *ansatz* defined in Eq. 3.79 for QDOs, considering that the Coulomb potential is well approximated by the dipole potential for distances characteristic of the dispersion interactions. Thus the approximated electron-drudon coupling has the form

$$J_{e-d}(\bar{\mathbf{r}}^e, \bar{\mathbf{r}}^d; \bar{\mathbf{R}}^n, \bar{\mathbf{R}}^O) = \exp[\boldsymbol{\mu}^\top \mathbf{B} \bar{\mathbf{r}}_{dO}], \quad (3.90)$$

where $\bar{\mathbf{r}}^{dO}$ is the vector of the distances between each drudon and its corresponding center, defined in the previous section, and $\boldsymbol{\mu}$ is the 3-dimensional vector of the dipole moment of the electronic part of the system (explicitly depending on the positions of the electrons $\bar{\mathbf{r}}^e$)

$$\boldsymbol{\mu} = \sum_{j=1}^{N_n} Z_j \mathbf{R}_j^n - \sum_{i=1}^{N_e} \mathbf{r}_i^e, \quad (3.91)$$

and \mathbf{B} is a matrix containing $3 \times 3N_d$ variational parameters.

The bias introduced in DMC calculations from the use of this approximated *ansatz* is more difficult to estimate. In theory, the drudonic part should explicitly affect the nodal structure of the electronic part, which is not possible through the functional form defined in Eq. 3.90. In practice, it is well known that non-covalent interactions in the case of standard electronic calculations do not greatly depend on the nodal surface in between the bond [43], and from the accuracy of the results presented in the Chap. 4 we can conclude that the *ansatz* in Eq. 3.90 is a reasonable first approximation for the El-QDO coupling in the total wavefunction.

An improvement of the *ansatz* is of course possible, for example by introducing a Jastrow term similar to that proposed for the QDO wavefunctions in Eqs. 3.86 and 3.89, which would satisfy the cusp conditions between the electrons, drudons, and centers of the QDOs. Yet, this would most likely give just a very small improvement on the VMC energy estimations for distances that are characteristic of dispersion interactions, while the contributions to the DMC calculations would be only in the efficiency of the sampling.

El-QDO applications

Parts of this Chapter have been published in this or a similar form in:

M. Ditte, M. Barborini, L. M. Sandonas, A. Tkatchenko. “Molecules in Environments: Toward Systematic Quantum Embedding of Electrons and Drude Oscillators” *Phys. Rev. Lett.* **131**, 228001 (2023)

and will be published in:

M. Ditte, M. Barborini, A. Tkatchenko. “Electrons embedded in charged oscillators: A quantum embedding approach for molecular systems” In preparation.

and have been produced in collaboration with the above authors.

In this Chapter, we present the first applications of the El-QDO method introduced in Sec. 2.5. First, we discuss the case of the noble gas dimers (Sec. 4.1), which are systems dominated by dispersion interactions, and of the water dimers (Sec. 4.2), in order to show the capability of the method to describe hydrogen bonds. The calculations of these dimers, which serve as proof of concept, are followed by the study of the solvation energies of benzene, the benzene dimer and ortho-benzyne in a water environment containing up to 50 explicit water molecules (Sec. 4.3). Moreover, we present

calculations of the singlet-triplet gap of ortho-benzyne in a water environment. All the results are compared to state-of-the-art *ab initio* methods.

4.1 Noble gas dimers

The long-range interaction energies of noble gas dimers are dominated by dispersion. For this reason, the QDO model constructed from the dispersion coefficients and polarizabilities, together with an external short-range repulsion to account for

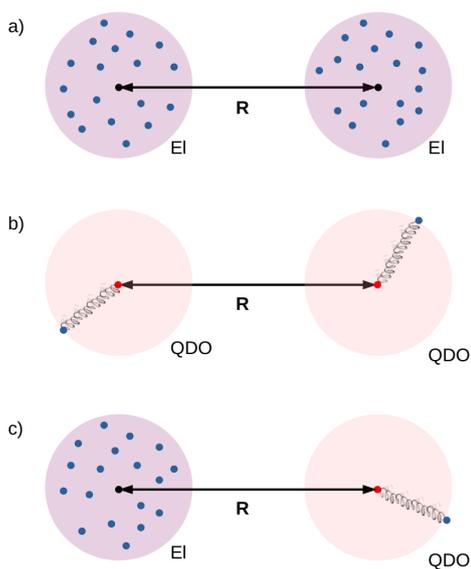


Figure 4.1: The three variants of the noble gas dimers studied in this work.

the exchange effects present by the antisymmetrization conditions of the electrons in the regular atoms has been successfully applied to these systems [88, 87, 204, 152]. Here we study the dissociation curves of Ne_2 , Ar_2 , Kr_2 and Xe_2 , each with three different Hamiltonians, as illustrated in Fig. 4.1: a) EI-EI, where both atoms are described at the fully electronic level using Hamiltonian from the Eq. 2.17 with 16 valence electrons, which is solved using the reference Coupled Cluster method [55, 142, 196, 71]; b) QDO-QDO, where two atoms are replaced by two QDOs, described by the Hamiltonians with the Coulomb and dipole potential in the Eqs. 2.41 and 2.53

with two drudons; and finally c) EI-QDO, where one of the atoms is described at the electronic level and the second one is replaced by a QDO, described by the Hamiltonian in Eq. 2.58 with 8 valence electrons and 1 drudon.

First, we study the dissociation curves without the external short-range repulsion, in order to show the interaction energies without introducing additional fitting parameters. The results are shown in Fig. 4.2 and are compared to results obtained from the symmetry-adapted perturbation theory (SAPT) decomposition [161], namely the

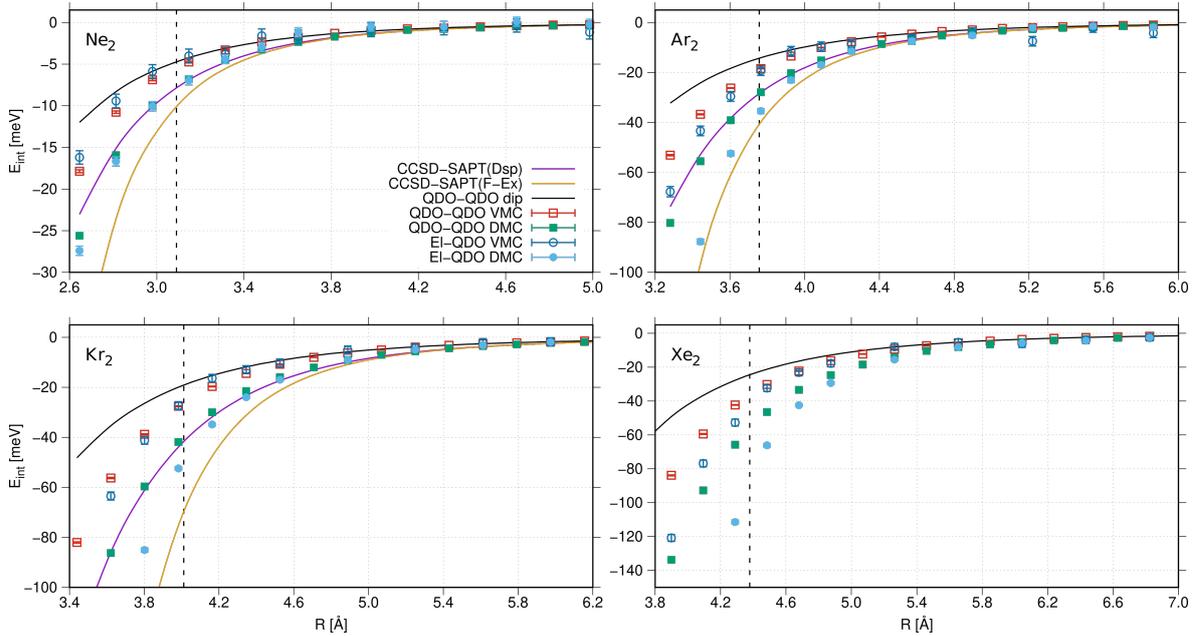


Figure 4.2: The interaction energies of Ne_2 , Ar_2 , Kr_2 and Xe_2 using the QDO-QDO and EI-QDO models from Fig. 4.1 at the VMC and DMC level of theory without the external short-range repulsion. The results are compared to the pure dispersion (Dsp) and the sum of all the terms not including exchange (F-Ex) from the EI-EI SAPT decomposition [161] and to the exact solution of two QDOs interacting via the dipole potential (QDO-QDO dip). The SAPT decomposition for Xe_2 is not available. The dashed vertical lines represent the equilibrium distance of the atoms.

pure dispersion contribution (Dsp) and the total energy minus all the terms including exchange (F-Ex) accounting also for electrostatics and polarization effects, and to the exact solution of the QDO-QDO model interacting via the dipole potential from Eq. 2.55 (QDO-QDO dip). Here we note that the SAPT decomposition is not unique and depends on the choice of the wavefunctions of the non-interacting monomers. At the VMC level, both the QDO-QDO and EI-QDO models give energies that are similar to the QDO model interacting via the dipole potential. This underestimation of the interactions at the VMC level is due to the approximate wavefunction in Eq. 3.79. Yet, at the DMC level, the QDO-QDO model reproduces the pure dispersion curves almost exactly for all the cases. Our EI-QDO model at the DMC level leads to lower interaction energies in all the cases, closer to the F-Ex curve due to the additional interactions present. The accuracy of the The DMC results clearly depend on the fact, that being QDOs distinguishable particles, their trial wavefunction does not have nodes and the

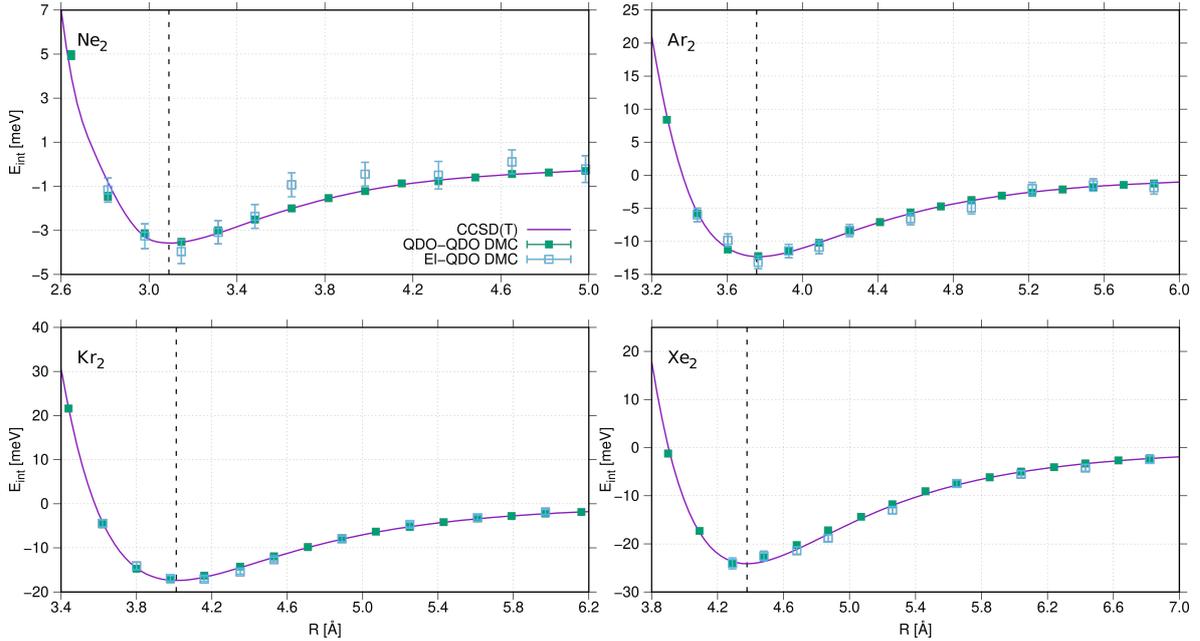


Figure 4.3: Binding curves of Ne_2 , Ar_2 , Kr_2 and Xe_2 obtained from the DMC data of the QDO-QDO and El-QDO models shown in Fig. 4.2 to which we have added the short-range repulsion potentials shown in Fig. 4.4 and whose parameters are reported in Tab. A1.3. The results are compared to the *ab initio* CCSD(T) calculations from Refs. 55, 142, 196 and 71 respectively. The dashed vertical lines represent the equilibrium distance of the atoms.

energy converges more accurately to the ground state. Furthermore, also electrons are distinguishable from QDOs. Both these conditions justify the choice of the *ansätze* in Eqs. 3.79 and 3.90. Interestingly, the small differences between the exact solution of the QDO-QDO model interacting via the dipole potential and the VMC for QDO-QDO using the full Coulomb potential at the equilibrium distances are due to the presence of multipolar contributions, beyond the dipole potential case, which are introduced through the optimized *ansatz* in VMC. The most important result from Fig. 4.2 is the agreement between QDO-QDO DMC, El-QDO DMC and the SAPT decomposition in the long-range region.

On top of these DMC results, as previously done in the literature for QDO dimers [88], we added an external short-range repulsion to the QDO-QDO and El-QDO models. This was done by fitting the difference between the reference CCSD(T) and the DMC results from Fig. 4.2 with an exponential function defined in Eq. 2.66 with $N=2$, with 4 parameters (see Tab. A1.3). Clearly, the binding curves shown in Fig. 4.3 perfectly

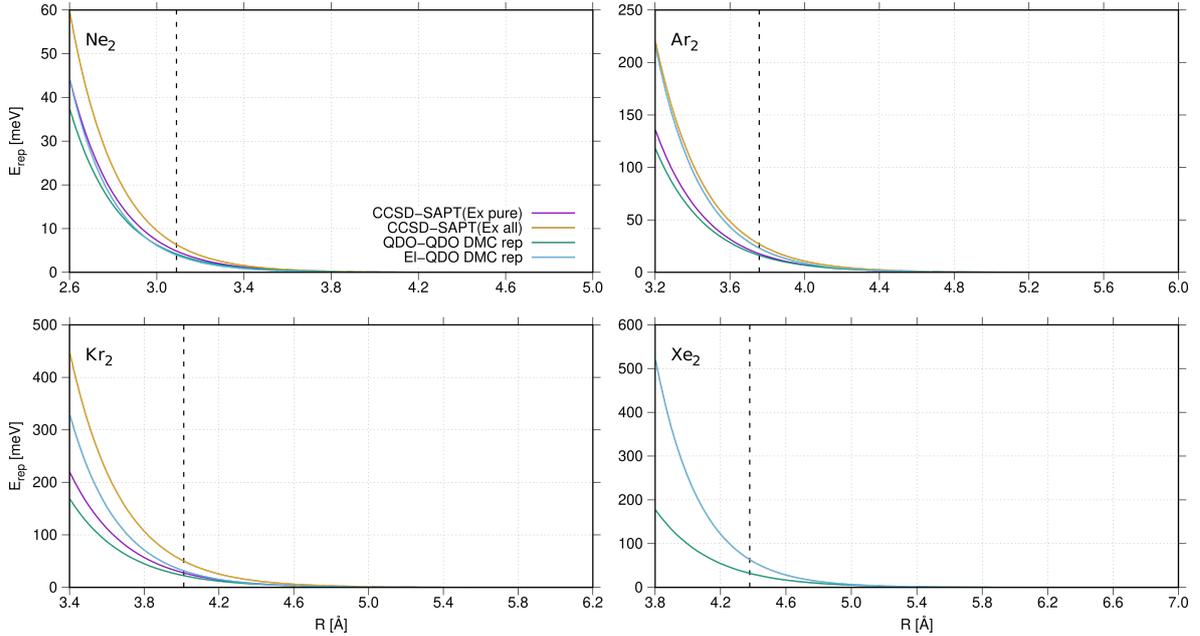


Figure 4.4: Plots of the external short-range repulsion for the noble gas dimers using Eq. 2.66 with $N=2$ and with parameters from Tab. A1.3, obtained by fitting the differences between the *ab initio* CCSD(T) data and the DMC dissociation curves for the QDO-QDO and El-QDO models. The repulsions from the QDO-QDO and El-QDO calculations are compared to the El-El SAPT decomposition [161], namely to the pure exchange contribution (Ex pure) and to all the terms containing exchange (Ex all). The SAPT decomposition for Xe_2 is not available. The dashed vertical lines represent the equilibrium distance of the atoms.

match the reference CCSD(T) results for all four dimers.

The pure short-range repulsive parts are plotted in Fig. 4.4, where we compare our fitted repulsions for the QDO-QDO and the El-QDO models at the DMC level, with the pure exchange curves from SAPT (Ex pure) and with the sum of all the mixed terms, which include exchange (Ex all). The goal here is not to reproduce the SAPT curves, which are not uniquely defined anyway, but to obtain a qualitative agreement with the fast exponential decay, that confirms the fact, that already at distances approximately 0.5 \AA longer than the equilibrium ones, these exchange contributions can be neglected.

4.2 Water dimers

Next, we apply the El-QDO model to the water dimer in its two energetically lowest geometries, labeled as C1 and C2 respectively, in order to confirm its ability to capture the hydrogen bond.

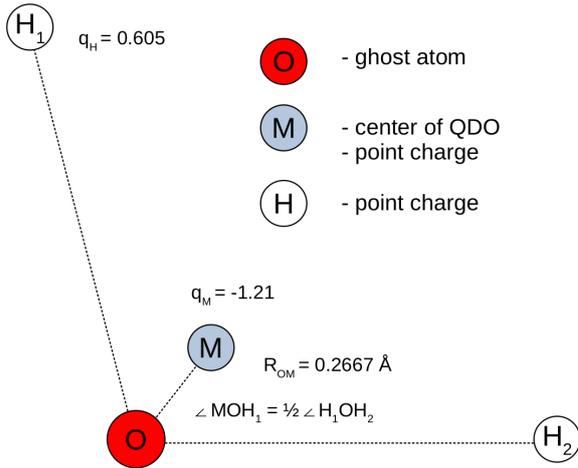


Figure 4.5: A schematic of the QDO-based model of the water molecule from Ref. 31. The $\angle HOH$ angles and R_{OH} distances are taken from real molecular structures.

In order to reproduce the dipole moment of an isolated water molecule, we used the QDO-based model of water [31], built on the classical TIP4P force field model [92], consisting of 2 point charges H ($q_H = 0.605$), 1 point charge M ($q_M = -1.21$) and a QDO placed at point M . A ghost atom O serves to define an angle $\angle HOM = \frac{1}{2} \angle HOH$ and the distance $R_{OM} = 0.2667 \text{ \AA}$. The $\angle HOH$ angles and R_{OH} distances were taken from atomistic geometry relaxation, which is different from the previous applications, where the angle $\angle HOH$ and the distance R_{OH} had been fixed. A schematic of the QDO model of water is shown in Fig. 4.5.

Similarly to the noble gas dimers, we compare the dissociation curves of standard electronic calculations to the ones obtained with the QDO-QDO and El-QDO models, as shown in Fig. 4.6: 1.) El-El in Figs. a) and e), where both atoms are described at the fully electronic level using the Hamiltonian in Eq. 2.17 with 16 valence electrons, which is solved using the reference Coupled Cluster method [127]; 2.) QDO-QDO in Figs. b) and f), where both monomers are replaced by QDOs, described by the Hamiltonian in Eq. 2.41 with two drudons; 3.) El-QDO in Figs. c) and g), where the acceptor is kept at the fully electronic level and the donor is replaced by a QDO (Eq. 2.58) with 8 valence electrons and 1 drudon; and 4.) El-QDO in Figs. d) and h), where the donor is described at the fully electronic level and the acceptor is replaced by a QDO, with 8 valence

electrons and 1 drudon. Here we recall that the acceptor molecule is the monomer donating the electrons of the electronegative oxygen atom toward the donor monomer. The reason for discussing both cases, in which the QDO plays the role of the donor and in which the QDO plays the role of the acceptor, comes from the necessity to verify how the model behaves while describing an anisotropic environment. In QM:MM methods [17], it is usually assumed, that the QM part has to model the acceptor molecule, enabling the charge transfer towards the donor. The opposite case becomes complicated to model within the QM:MM methods, because it does not enable the charge transfer. Here we study the ability of the El-QDO method to partially polarize the QDO monomer playing the role of the acceptor toward the electronic donor.

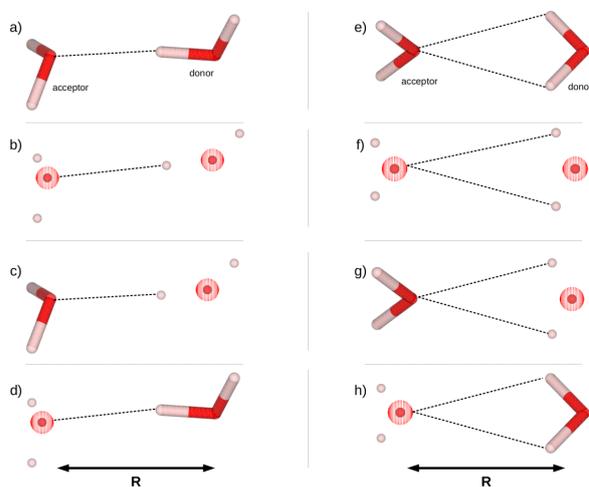


Figure 4.6: The variants of the water dimers studied in this work. The geometries are taken from Ref. 127

First, in Fig. 4.7, we again present the bare dissociation obtained with the QDO-QDO and El-QDO models. Configuration C1 is compared to pure dispersion (Dsp) and to the sum of all terms minus the exchange contributions (F-Ex) from SAPT decomposition [6]. Differently from the case of the noble gas dimers, here the results from the QDO-QDO and El-QDO models differ from the pure dispersion contribution due to the inclusion of the point charges that reproduce the electrostatic interactions between the permanent dipoles of the water monomers. We again observe that the El-QDO model at the DMC level is closer to

the F-Ex curve when compared to the QDO-QDO model. In particular, the El-QDO_D curve matches the F-Ex curve better than the El-QDO_A one due to the ability of the electrons to delocalize toward the QDO region, increasing the polarization and increasing the binding energy. Furthermore, the discrepancy between the El-QDO_A and El-QDO_B results increases in the case of the C2 conformer, in which the charge transfer from the acceptor's oxygen and the donor's hydrogens is enhanced. Both models underestimate the strength of the interactions at the VMC level, exactly as is expected

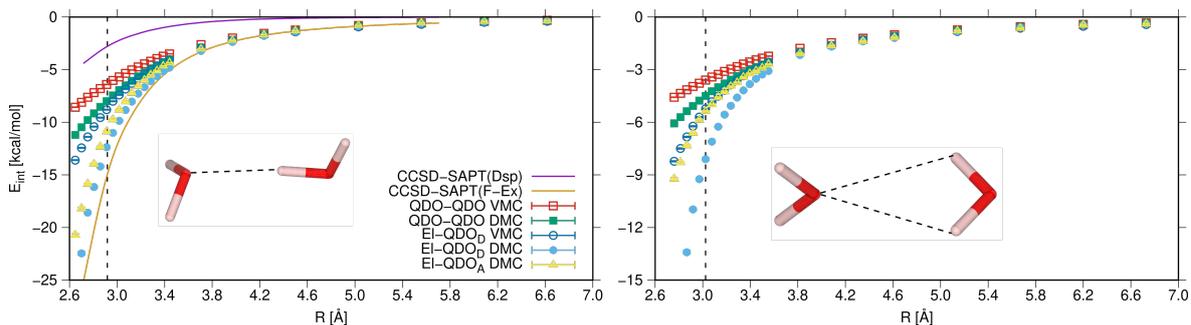


Figure 4.7: The interaction energies of the water dimer in its two geometries (C1 and C2) using the QDO-QDO and El-QDO models from Fig. 4.6 at the VMC and DMC level of theory without the external short-range repulsion. The subscripts in QDO_D and QDO_A denote the QDO playing the role of the donor and acceptor respectively. The results for the energetically lowest geometry C1 are compared to the pure dispersion (Dsp) and the sum of all terms not including exchange (F-Ex) from the El-El SAPT decomposition [6]. The dashed vertical lines represent the equilibrium distance of the monomers.

due to the limited form of the variational *ansatz*.

As for the noble gas dimers, in Fig. 4.8 we construct the repulsion potential (Tab. A1.3) that, when added on top of the bare results of Fig. 4.7, reproduces the PES of *ab initio* electronic calculations. For the QDO-QDO model, here we use the repulsion potential from Ref. 31, which as shown in Fig. 4.8 does not perfectly match the reference CCSD(T) curves due to the fact that it is an average repulsion that does not fully consider the anisotropy of the interactions. For this reason, we have chosen to fit the repulsion independently for each of the two configurations for the El-QDO model. It is important here to recall, that all four potential energy curves coincide for intermolecular distances larger than 3.8 Å that will be exploited in the following sections.

The external short-range repulsion for the two geometries of water dimer, with parameters reported in Tab. A1.3, are plotted in Fig. 4.9.¹ The repulsions for the QDO-QDO and El-QDO models, fitted using Eq. 2.66 with $N=2$, are again compared to two curves from SAPT decomposition [6], namely the pure exchange (Ex pure) and

¹Due to the discrepancies of the QDO-QDO model in Fig. 4.8 we decided to refit the repulsion for the C1 and C2 geometries separately. The new refits are labeled as refit (R) and are compared to the universal one (U) proposed in Ref. 31.

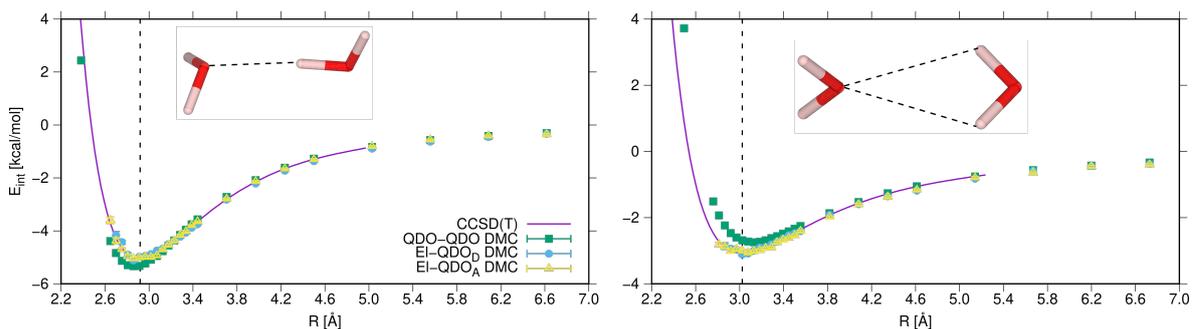


Figure 4.8: The binding curves of the water dimer in its two geometries (C1 and C2) obtained from the DMC data of the QDO-QDO and EI-QDO models shown in Fig. 4.6 to which we have added the short-range repulsion potentials shown in Fig. 4.9 and whose parameters are reported in Tab. A1.3 (the QDO-QDO repulsion is labeled as U). The subscripts in QDO_D and QDO_A denote the QDO playing the role of the donor and acceptor respectively. The results are compared to the *ab initio* CCSD(T) calculations [127]. The dashed vertical lines represent the equilibrium distance of the monomers.

to sum of all terms including exchange (Ex all). Again, the goal here is not to reproduce the SAPT curves, but to show a qualitative agreement. From these results, it is clear, that at the distances of about 4 Å the repulsion contribution can be ignored.

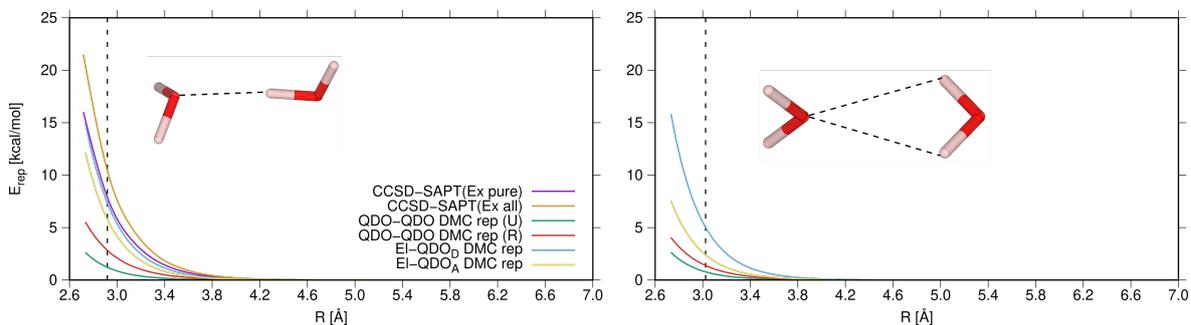


Figure 4.9: Plots of the external short-range repulsion for the water dimers in its two geometries (C1 and C2) using the Eq. 2.66 with $N=2$ and with parameters from Tab. A1.3, obtained by fitting the differences between the reference CCSD(T) and the DMC dissociation curves for the QDO-QDO and EI-QDO models. The subscripts in QDO_D and QDO_A denote the QDO playing the role of the donor and acceptor respectively. Labels U (uniform) and R (refit) denote the universal repulsion from Ref. 31 and our geometry-dependent refit respectively. The repulsions of the C1 geometry from the QDO-QDO and EI-QDO models are compared to the EI-EI SAPT decomposition [6], namely to the pure exchange contribution (Ex pure) and to all terms containing exchange (Ex all). The dashed vertical lines represent the equilibrium distance of the monomers.

4.3 Molecules in a water environment

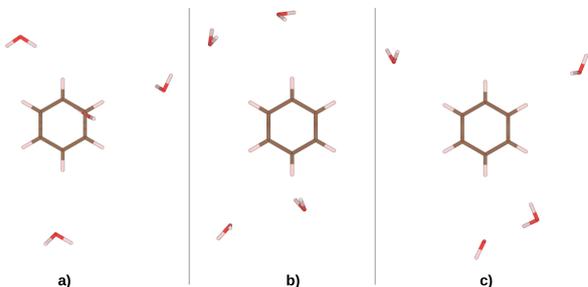


Figure 4.10: Benzene embedded in 3 different water cages: a) cage 1 ($R_{\min} = 3.54 \text{ \AA}$); b) cage 2 ($R_{\min} = 3.54 \text{ \AA}$); and c) cage 3 ($R_{\min} = 3.57 \text{ \AA}$).

After studying the atomic and molecular dimers, here we apply the El-QDO approach to a more complex anisotropic environment studying small molecules in explicit water environments of various sizes. We compare the El-QDO method to the state-of-the-art Density Functional Theory (DFT) calculations using PBE0 [4] functional with additional dispersion interactions, namely with pairwise Tkatchenko-Scheffler (PBE0+TS) [180] approach and

with many-body dispersion (PBE0+MBD) [179, 9] method. Furthermore, we compare our calculations also to the standard QM:MM approach described in Sec. 2.3, that couples DMC with a traditional FF [92, 93], and for the smallest systems, for which the calculations are computationally feasible, also with the pure electronic DMC method.

4.3.1 Benzene in 4 water molecules

The first studied system is benzene embedded in three different small cages (cage 1-3) composed of four water molecules, shown in Fig. 4.10. The cages are artificially generated geometries, with the oxygens of the water molecules randomly placed on a sphere of radius 6 \AA around the center of benzene and with random directions of the HOH

Table 4.1: The solvation energies (in kcal/mol) of benzene in the three different cages composed of 4 water molecules from Fig. 4.10.

	PBE0	PBE0+TS	PBE0+MBD	DMC	DMC El-FF	DMC El-QDO
cage 1	-1.38	-2.03	-1.97	-1.97(6)	-1.83(5)	-1.94(4)
cage 2	-1.37	-1.96	-1.91	-1.91(7)	-1.77(6)	-1.83(5)
cage 3	-1.53	-2.06	-2.01	-2.02(8)	-2.05(5)	-1.86(5)

Table 4.2: The decomposition of the El-FF solvation energies (in kcal/mol) for benzene embedded in the three various cages composed of 4 water molecules from Fig. 4.10.

	El-TIP3P	Dsp.	El-FF
cage1	-1.20(5)	-0.63	-1.83(5)
cage2	-1.19(6)	-0.58	-1.77(6)
cage3	-1.53(5)	-0.52	-2.05(5)

and the water molecules being more than 3.5 Å, so that the additional repulsion potentials can be neglected. The solvation energies of these three systems, defined as the energy difference $E_{\text{mol-in-env}} - E_{\text{mol}} - E_{\text{env}}$, are shown in Tab. 4.1. All the methods, except for the bare PBE0, agree up to the sub-chemical accuracy. The small difference between the pairwise PBE0+TS and many-body PBE0+MBD/DMC indicates a small importance of the many-body effects for this small simple system. The El-FF method for these small systems is comparable with the El-QDO one, but it will not be the case as the size of the system increases. The decomposition of the DMC El-FF into the electrons/nuclei - point charges contribution [92] (El-TIP3P) and the pairwise dispersion contribution [93] (Disp.) in Tab. 4.2 shows on average 31% of the interaction energy coming from the pairwise contribution. Most importantly, DMC El-QDO gives correct solvation energies.

plane with fixed $R_{OH}=0.958$ Å and $\angle\text{HOH}=104.4^\circ$.

The final chosen cages were the ones with the largest solvation energies at PBE0+MBD level of theory. The reason for using these artificial water cages is to completely avoid the short-range repulsion problem discussed in the previous sections, so that the binding energy given by the model will correspond to that of the *ab initio* predictions, with the minimal distance between the benzene molecule

4.3.2 Ortho-benzyne in 4 and 30 water molecules

Table 4.3: The solvation energies (in kcal/mol) of ortho-benzyne in the singlet and adiabatic triplet states in the cages composed of 4 and 30 water molecules from Fig. 4.11.

	PBE0	PBE0+TS	PBE0+MBD	DMC	DMC El-FF	DMC El-QDO
S (4W)	-1.01	-1.50	-1.43	-1.5(1)	-1.30(5)	-1.39(5)
T (4W)	-0.77	-1.24	-1.18	-1.0(1)	-1.11(5)	-1.12(5)
S (30W)	-2.82	-6.78	-6.11	-	-5.44(5)	-6.20(7)
T (30W)	-2.37	-6.41	-5.71	-	-4.99(5)	-5.73(7)

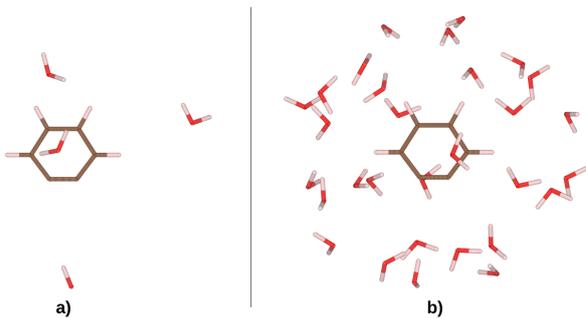


Figure 4.11: Ortho-benzyne in the singlet state geometry embedded in two cages composed of a) 4 water molecules (4W, $R_{\min}=3.61$ Å) and b) 30 water molecules (30W, $R_{\min}=2.99$ Å). The geometries of ortho-benzyne are taken from Ref. 160.

The second studied system is ortho-benzyne in the singlet (S) and the adiabatic triplet (T) states in two cages, one composed of 4 water molecules (4W) and the second composed of 30 water molecules (30W), both shown in Fig. 4.11. The cages were generated in the same way as in the previous example of benzene, except here the criterion for the final cages was the maximal change of the singlet-triplet (ST) gap in the cage compared to the ST gap in the vacuum at the PBE0+MBD level of theory. We report the solvation energies of these four systems

in Tab. 4.3. Similar to benzene in the small water cages, the solvation energies of ortho-benzyne in both states for the 4W cage agree for all the methods, except the bare PBE0, up to the sub-chemical accuracy. This is not the case for the larger cage composed of 30 water molecules, where we see a 0.7 kcal/mol difference between PBE0+TS and PBE0+MBD, indicating the importance of the many-body effects. Unfortunately, fully electronic DMC calculations are not feasible due to the size of the systems, but DMC El-QDO gives solvation energies in the match with the PBE0+MBD results, while DMC El-FF underestimates the interactions between ortho-benzyne and the environment by 0.75 kcal/mol. The decomposition of the DMC El-FF into the electrons/nuclei - point charges contribution [92] (El-TIP3P) and the pairwise dispersion contribution [93] (Disp.) in Tab. 4.5 shows that the discrepancy between the El-QDO and the El-FF results comes from the fact, that the dispersion contributions are greatly

Table 4.4: The singlet-triplet (ST) excitation energies (in kcal/mol) of ortho-benzyne in the vacuum (V) and the two cages composed of 4 and 30 water molecules from Fig. 4.11. All the DMC results are corrected to the ECP error of 2.85 kcal/mol from Tab. A1.4.

	PBE0	PBE0+TS	PBE0+MBD	DMC	DMC El-FF	DMC El-QDO
ST (V)	28.23	28.23	28.26	37.23(5)	-	-
ST (4W)	28.46	28.48	28.51	37.7(1)	37.43(5)	37.50(5)
ST (30W)	28.68	28.61	28.66	-	37.68(5)	37.70(8)

increasing between the 4W (35%) and 30W (66%) cages and probably the pairwise approximation in the El-FF method underestimates the correlations between the molecule and the environment with respect to the many-body interactions in El-QDO.

Table 4.5: The decomposition of the El-FF solvation energies (in kcal/mol) for ortho-benzyne in the singlet and adiabatic triplet states in the cages composed of 4 and 30 water molecules from Fig. 4.11.

	El-TIP3P	Dsp.	El-FF
S (4W)	-0.87(5)	-0.43	-1.30(5)
T (4W)	-0.69(5)	-0.42	-1.11(5)
S (30W)	-2.04(5)	-3.40	-5.44(5)
T (30W)	-1.56(5)	-3.43	-4.99(5)

The main reason for studying ortho-benzyne is its ST gap, which is known to be challenging at the DFT level due to the charge delocalization error [160]. The ST gap, called also the singlet-triplet excitation energy in the text, is calculated as $\Delta E_{ST} = E_T - E_S$, where E_S is the energy of the ground state with total spin equal to 0 and E_T is the energy of the lowest triplet state with the total spin equal to 1 (corresponding to a single flipped electron). Tab. 4.4 contains the adiabatic ST gaps of ortho-benzyne in the vacuum (V) and in the two previously discussed water cages, where all the DMC calculations using effective core potential are corrected to the ECP error of 2.85 kcal/mol coming from all-electrons vs ECP calculations at UPBE0 level of theory reported in Appendix A1 in Tab. A1.4. We see the PBE0 calculations in the vacuum being 9 kcal/mol far from the experimental value of 37.5(3) kcal/mol [202], independently of the used dispersion method. However, DMC, which does not suffer from the delocalization error, is able to recover the experimental value. For the smallest cage composed of 4 water molecules, DMC El-QDO agrees with DMC within the sub-chemical accuracy. Interestingly DMC El-FF agrees with DMC El-QDO for both water cages. The reason is a negligible change in the dispersion between the singlet and triplet states observed for all the methods, lead-

Table 4.6: The solvation energies (in kcal/mol) of the benzene dimer and its two monomers in the cage composed of 50 water molecules from Fig. 4.12.

	PBE0	PBE0+TS	PBE0+MBD	DMC El-FF	DMC El-QDO
Dimer	-4.9	-10.9	-10.0	-8.9(1)	-10.3(2)
Monomer ₁	-1.5	-4.1	-3.8	-3.48(6)	-4.0(1)
Monomer ₂	-3.4	-6.6	-6.1	-5.18(6)	-6.4(1)

ing to the cancellation of the pairwise contribution in the El-FF embedding (Tab. 4.7), which is the source of the error in the previously shown solvation energies.

4.3.3 Benzene dimer in 50 water molecules

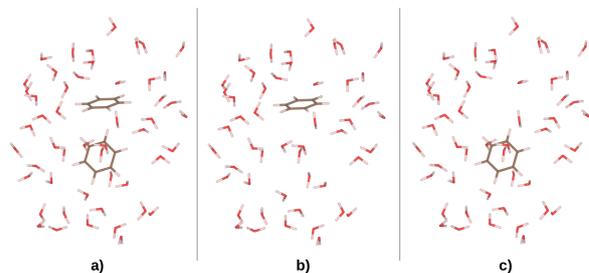


Figure 4.12: a) The T-shaped benzene dimer ($R_{\min}=3.4 \text{ \AA}$); b) its monomer 1; and c) monomer 2 in a cage composed of 50 water molecules [163].

The last studied system is the T-shaped benzene dimer and its monomers in a cage composed of 50 water molecules (Fig. 4.12), which is an example of a computationally unfeasible system for the pure electronic DMC and thus a method like the El-QDO embedding is necessary. The water cage, originally obtained from molecular dynamics, has been expanded from the dimer’s center by 1.5 \AA in order to increase the minimal distance between the benzene dimer and the cage to 3.4 \AA .

The solvation energies, reported in Tab. 4.6, show discrepancies between PBE0+TS and PBE0+MBD, again indicating the importance of the many-body effects. The DMC El-QDO method agrees with the PBE0+MBD one up to the sub-chemical accuracy, while DMC El-FF underestimates the interaction energies by up to 1 kcal/mol . The reason for the failure of the DMC El-FF can be seen from the decomposition of the DMC El-FF energy into the electrons/nuclei - point charges contribution [92] (El-TIP3P) and the pairwise dispersion contribution [93] (Disp.) in Tab. 4.8, where the dispersion contribution corresponds to 50-70% of the total solvation energy, and again can be expected to underestimate the correlation effects, when compared to the many-body treatment in the El-QDO method.

4.4 Computational details

Table 4.7: The decomposition of the El-FF excitation energies (in kcal/mol) of ortho-benzyne in the cages composed of 4 and 30 water molecules from Fig. 4.11. The energies are without the 2.85 kcal/mol ECP correction from Tab. A1.4.

	El-TIP3P	Dsp.	El-FF
ST (4W)	34.57(5)	0.01	34.58(5)
ST (30W)	34.86(5)	-0.03	34.83(5)

All the reference PBE0+TS and PBE0+MBD calculations in this Chapter have been done in FHI-aims package [2], using the all-electron tight basis set, consisting of numerical orbitals [69]. For all the QMC calculations we have used ccECP effective core potentials [198, 16, 10, 15] with the corresponding (aug)-cc-pVDZ Gaussian basis sets for the noble gas and water dimers and cc-pVTZ for all the calculations of benzene/benzene dimer/ortho-benzyne in water cages and in the vacuum. All the molecular orbitals for the QMC wavefunctions were taken from DFT calculations using PBE0 functional [4] in GAMESS (2016 R1) [13] and Orca 5.0 [136] softwares. The dynamical Jastrow factor is constructed from 3s2p1d uncontracted Gaussians (GTOs) for all the heavy atoms and from 2s1p GTOs for the hydrogen atoms. When not mentioned explicitly, $dt = 0.005$ a.u. has been used for the DMC calculations, with standard error function damping from Eq. 2.60, using $\sigma = 0.1$ for all species except for the center of QDO of the water model, where $\sigma = 1.2$ [31] and the Slater determinant of the electronic part was fixed during the optimization at VMC level of theory. For all the QMC calculations of QDO and El-QDO models the dipole *ansatz* of the wavefunction without the Jastrow factor from Eq. 3.79 has been used.

In the rest of this section, we discuss the convergence and computational efficiency of the El-QDO method for some of the systems from Sec. 4.3.

4.4.1 Computational efficiency

In Tab 4.9 we show the relative runtimes of the DMC calculations in the water environment compared to the vacuum. The test calculations have been done on 56 CPUs using 12 walkers per CPU and 20 bins with 100 steps per block. The time per block is an

Table 4.8: The decomposition of the El-FF solvation energies (in kcal/mol) for the benzene dimer and its monomers in the cage composed of 50 water molecules from Fig. 4.12.

	El-TIP3P	Dsp.	El-FF
Dimer	-3.8(1)	-5.11	-8.9(1)
Monomer ₁	-1.04(6)	-2.45	-3.48(6)
Monomer ₂	-2.52(6)	-2.66	-5.18(6)

average over all 20 blocks and the relative runtime is calculated as the ratio of time per block times the ratio of the square of the root mean square deviation σ . In the case of ortho-benzyne (S), it is 11.99 times more expensive to perform fully electronic calculations in the environment of 4 water molecules when compared to the vacuum. In the case of the El-QDO method, this is only 1.02 and 1.25 times more expensive for 4 and 30 QDOs respectively. For the biggest cage composed of 50 QDOs, we see a slowdown by a factor of 1.47 and 1.27 for benzene monomer and benzene dimer respectively when compared to calculations in the vacuum.

Table 4.9: The relative runtime of the El-QDO embedding at the DMC and VMC level of theory with respect to the vacuum and the fully electronic cage.

	DMC			VMC		
	time per block [s]	σ [Ha]	runtime relative to vacuum	time per block [s]	σ [Ha]	runtime relative to vacuum
ortho-benzyne (S) from Fig. 4.11						
El (V)	6.36	0.80	1.00	1.26	0.80	1.00
El (4W)	25.10	1.40	11.99	4.63	1.42	11.47
El-QDO (4W)	6.45	0.80	1.02	1.29	0.82	1.05
El-QDO (30W)	6.78	0.87	1.25	1.45	0.88	1.36
benzene (monomer ₁) from Fig. 4.12						
El (V)	7.15	0.80	1.00	-	-	-
El-QDO (50W)	8.34	0.89	1.47	-	-	-
benzene dimer from Fig. 4.12						
El (V)	30.95	1.31	1.00	-	-	-
El-QDO (50W)	32.90	1.57	1.27	-	-	-

Tab. 4.9 contains also a similar comparison for the VMC calculations for ortho-benzyne (S). The test calculations have been done on 56 CPUs using 1 walker per CPU

Table 4.10: The number of parameters and the total runtime of the El-QDO optimization of ortho-benzyne (S) from Fig. 4.11.

	# of parameters	runtime [s]
El (V)	376	56.00
El (4W)	692	198.21
El-QDO (4W)	490	58.04
El-QDO (30W)	4741	72.65

4 and 30 QDOs compared to the vacuum.

Table 4.11: The mean square deviations σ of the wavefunctions of ortho-benzyne (S) from Fig. 4.11 in the vacuum and after optimizing various parts of the wavefunction embedded in 30 QDOs for the fixed and optimized nodes of the Slater determinant.

Optimization block	σ [Ha] fixed nodes	σ [Ha] optimized nodes
vacuum	0.80	0.74
in unopt. QDOs	0.91	0.91
QDO	0.86	0.80
El	0.86	0.81
El-QDO	0.86	0.81
all	0.85	0.81

ortho-benzyne in the vacuum.

There are three reasons for the high computational efficiency of the El-QDO method: 1) the reduced number of quantum degrees of freedom; 2) the faster evaluation of the drudonic wavefunction compared to the more complicated electronic wavefunction; and

and 20 bins with 100 steps per block. The time per block is an average over all 20 blocks and the relative runtime is calculated as the ratio of time per block times the square of the ratio of the root mean square deviation σ . The fully electronic calculations in the environment of 4 water molecules are 11.47 times more expensive than calculations in the vacuum. On the other hand in the case of the El-QDO method, it is only 1.05 and 1.36 times more expensive to perform the calculations in

Runtimes of the optimization of the wavefunctions with the corresponding number of variational parameters for ortho-benzyne (S) are shown in Tab. 4.10. The test calculations have been done on 56 CPUs using 1 walker per CPU, 200 bins with 2 steps per block and 10 optimization steps without the correlated sampling. Here we report only the absolute values for 10 optimization steps including the sampling. Almost all the differences come from the sampling, indicating the robustness of the optimization algorithm, which is able to optimize 4741 parameters of ortho-benzyne in 30 QDOs almost as fast as 376 parameters of

3) the lower statistical noise in the sampling.

4.4.2 Optimization of the wavefunctions

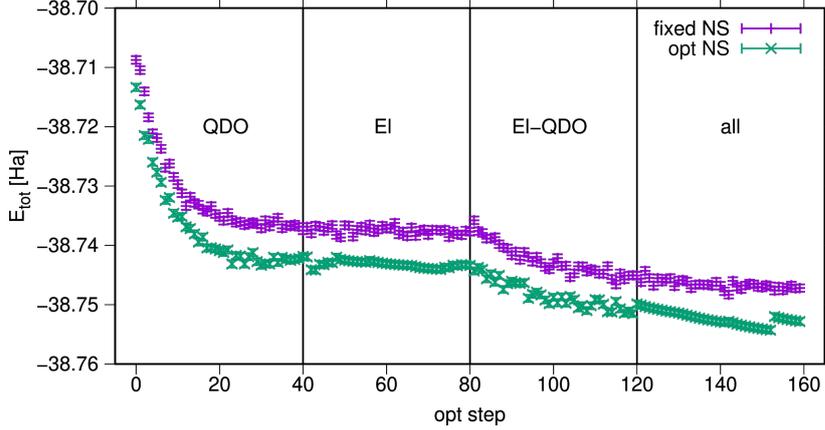


Figure 4.13: Total energy of ortho-benzene (S) in the cage composed of 30 QDOs from Fig. 4.11 as a function of the optimization step for the fixed and optimized nodal surface (NS). The 'QDO', 'EI', 'EI-QDO' and 'all' blocks are explained in the text of Sec. 4.4.

the presence of the QDOs (vacuum). In the next step, we add the QDOs, set the drudonic wavefunction to non-interacting QDOs (a diagonal form of the matrix \mathbf{A} in Eq. 3.79), set the EI-QDO coupling (matrix \mathbf{B} in Eq. 3.79) to zero, and we do not perform an optimization (in unopt. QDOs). This is followed by optimization of the drudonic part of the wavefunction (QDO), optimization of the electronic part (EI), optimization of the EI-QDO coupling part (EI-QDO), and by optimizing all the parameters together (all). All these steps are done with both, fixed and optimized nodal structure of the wavefunction (molecular orbitals in the Slater determinant). The increase of σ from vacuum to unoptimized QDOs is only 12% for the fixed node case, with a further decrease to 6% increase at the end of the optimization. The biggest gain is coming from the optimization of the QDO part, while the rest is contributing at a lower order of magnitude, meaning that the contributions of the correlations recovered by this term are essential for the convergence of the system's energy. The total energies of the same system during the optimization are shown in Fig. 4.13. We see the

In Tab. 4.11 we compare the mean square deviations σ of the wavefunctions optimized at various levels of the ortho-benzene molecule in its singlet state shown in Fig. 4.11. The value of σ is a measure of the quality of the *ansatz*, with the exact solution having $\sigma=0$. First, we optimize the electronic part of wavefunction in the vacuum, without

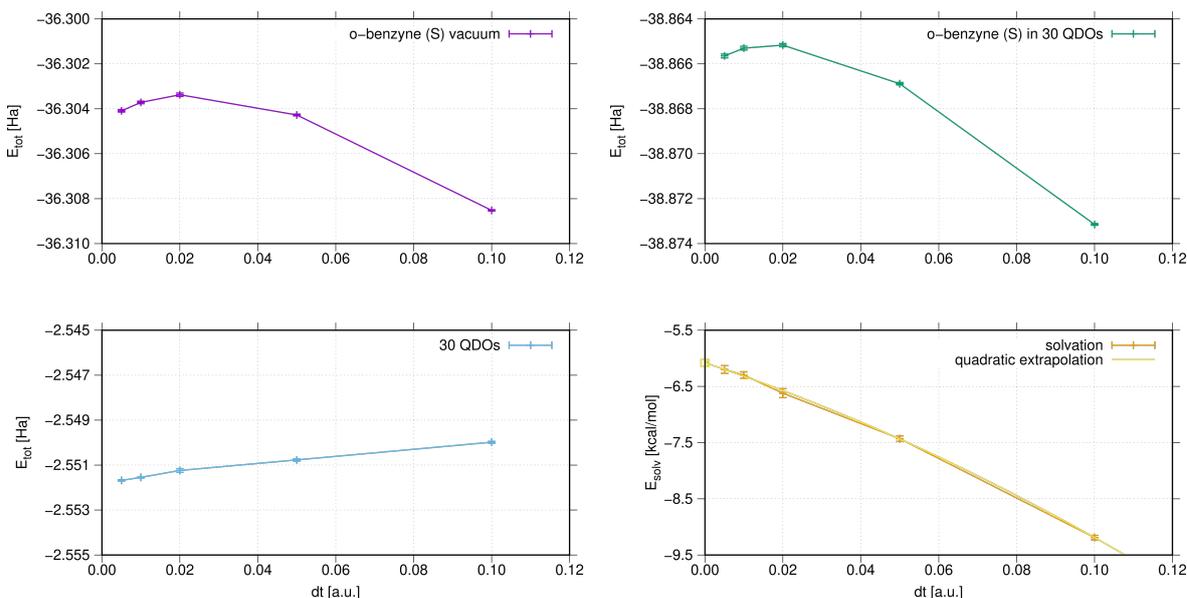


Figure 4.14: The first three bars: the total DMC energies of ortho-benzyne (S) in the vacuum; ortho-benzyne (S) in 30 QDOs; and the empty cage composed of 30 QDOs as a function of the time step $\delta\tau$. The bottom right bar: the DMC solvation energy of ortho-benzyne (S) embedded in 30 QDOs as a function of the time step $\delta\tau$, with a quadratic extrapolation.

same trend as reported in Tab. 4.11, for which the biggest energy improvement comes from the QDO part, followed by the El-QDO coupling. The optimization of the nodal surface leads to a reduction of σ by 5%. The optimization curves with and without the optimization of the nodal structure in Fig. 4.13 are identical up to a constant shift, implying that the presence of the environment has only a small effect on the nodal surface of the molecule in the solvent. This may not be the case with a wavefunction with more explicit interactions between the nodes and the QDOs, but in general, we expect this effect to be negligible at these distances. For completeness, we calculated the DMC energy of ortho-benzyne (S) in 30 QDOs with the optimized nodal surface. The solvation energy with the optimized nodes is equal to -6.04(6) kcal/mol, which is very close to the value obtained with the fixed nodes (-6.20(7) kcal/mol) and it is within the error bars from the PBE0+MBD reference value (-6.11 kcal/mol).

4.4.3 Time step $\delta\tau$ convergence

In order to verify the convergence of the results with respect to the time step $\delta\tau$ for the DMC calculation we report a convergence analysis for the ortho-benzyne molecule in its singlet state embedded in 30 QDOs. The total energies and the solvation energy as a function of $\delta\tau$ are shown in Fig. 4.14. Here we fixed the scale for the y-axis for the total energies in order to not overamplify the convergence of the QDO cage. We performed a quadratic extrapolation for the solvation energy, which for the limit $\delta\tau \rightarrow 0$ leads to the value -6.08(3) kcal/mol. This value is very close to the solvation energy for $\delta\tau=0.05$ (-6.20(7) kcal/mol) and it is within the error bars from the PBE0+MBD value (-6.11 kcal/mol).

4.4.4 Effects of the damping functions

Finally, in this section, we study the effects of the damping functions mentioned in Sec. 2.6 on the example of Ar_2 . The normal damping for the QDO-QDO model in Fig. 4.15 shows the independence of the results on the choice of the functional form of the damping for parameter $\sigma \leq 0.5$. The uniform damping for QDO-QDO is even more stable, with all the damping functions matching with σ as large as 2.5 (and thus the plot is not shown). For the normal damping of the El-QDO model, shown in Fig. 4.16, the behavior is very similar, with the curves being independent of the choice of the damping function up to $\sigma = 1.0$, except for the shortest distance between the atoms considered in this work, where $\sigma > 0.5$ gives results closer to the F-Ex curve. The uniform version of the damping for El-QDO is plotted in Fig. 4.17. The dissociation curves are again stable with respect to the choice of the damping function for σ up to 3.5, except for the shortest separation of the atoms. For the larger values of σ the dissociation curves go closer to the pure dispersion from SAPT, which is an opposite trend when compared to the normal damping for El-QDO from Fig. 4.16. Here we conclude that the results are robust with respect to the choice of the damping used to avoid the divergences of the Coulomb potential and the unphysical overpolarization.

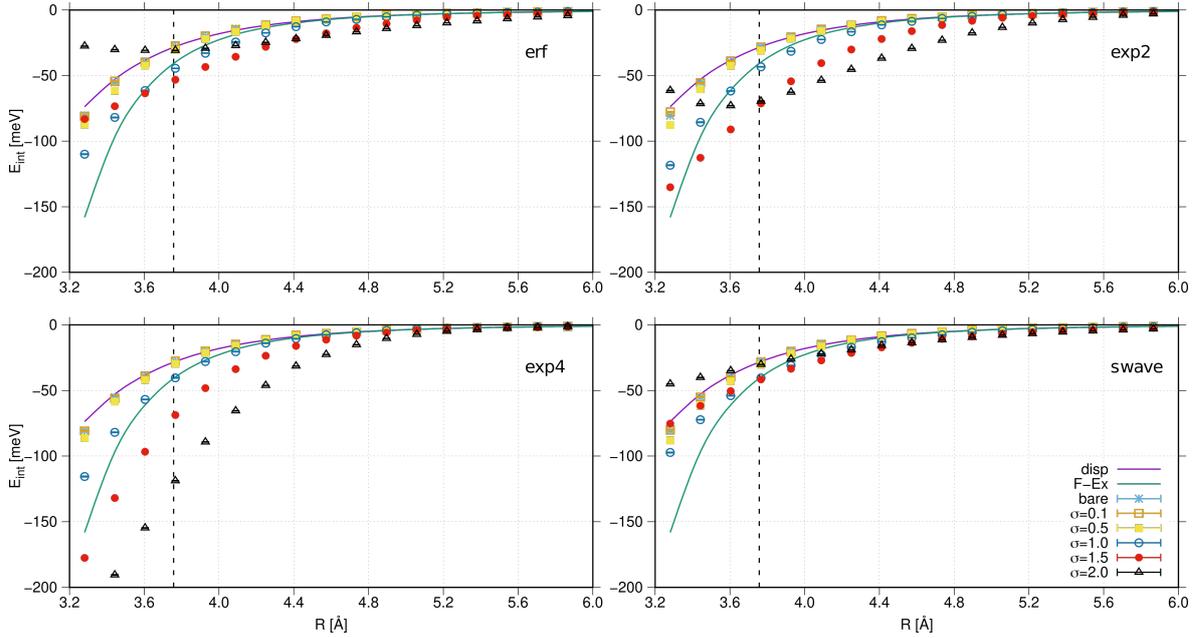


Figure 4.15: Interaction energies of Ar_2 using the QDO-QDO model at the DMC level of theory using the normal damping functions from Eqs. 2.60, 2.61, 2.62 and 2.63 for various values of the parameter σ . The results are compared to the pure dispersion (Dsp) and the sum of all terms not including exchange (F-Ex) from the El-El SAPT decomposition [161] and to the QDO-QDO model interacting via the bare Coulomb potential without the damping (bare). The dashed vertical lines represent the equilibrium distance of the atoms.

4.5 Conclusions

In this Chapter, we have presented applications of the novel El-QDO embedding method, where QDOs represent the environment and its dynamic response. We have used the full Hamiltonian of electrons and QDOs coupled via the Coulomb potential whose ground state is approximated by means of the product between the electronic Slater-Jastrow wavefunction, the dipole wavefunction of QDOs and the electron-QDO Jastrow factor. This variational *ansatz* has been optimized by means of the variational Monte Carlo methods and an accurate estimation of the energy has been obtained through diffusion Monte Carlo. The implemented methods were used to compute the binding energies of four noble gas dimers, Ne_2 , Ar_2 , Kr_2 and Xe_2 , prototypes for pure dispersion interactions and two configurations of the water dimer as prototypes for the hydrogen bonds, finding excellent agreement with *ab initio* reference data. After these

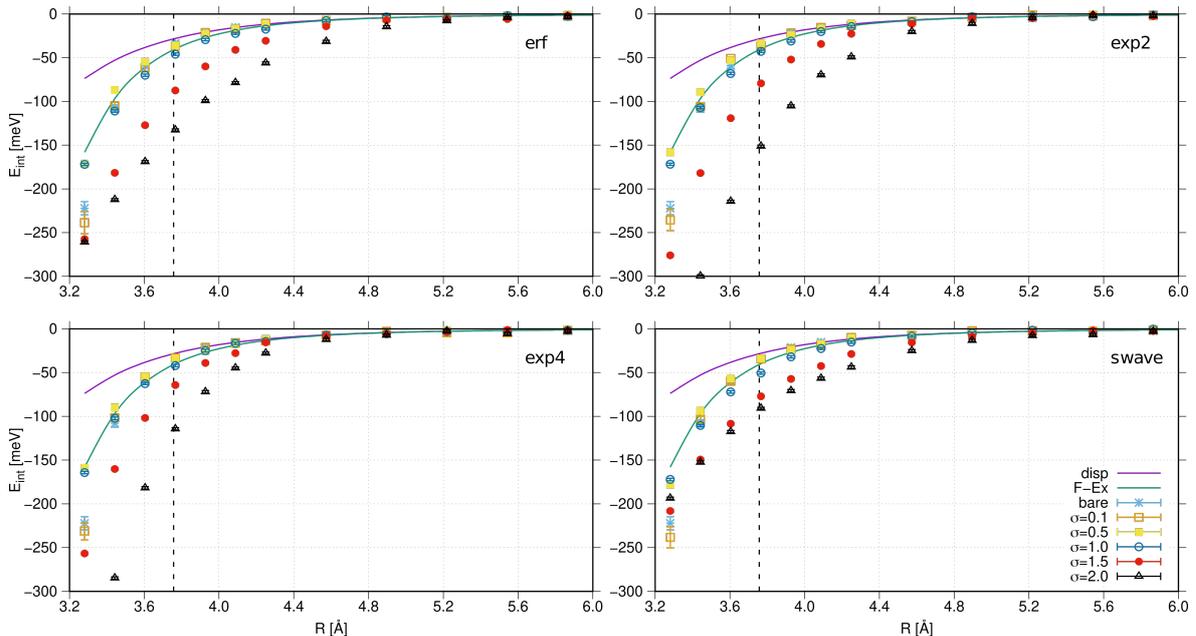


Figure 4.16: Interaction energies of Ar_2 using the El-QDO model at the DMC level of theory using the normal damping functions from Eqs. 2.60, 2.61, 2.62 and 2.63 for various values of the parameter σ . The results are compared to the pure dispersion (Dsp) and the sum of all terms not including exchange (F-Ex) from the El-El SAPT decomposition [161] and to the El-QDO model interacting via the bare Coulomb potential without the damping (bare). The dashed vertical lines represent the equilibrium distance of the atoms.

proofs of concept, we presented applications to molecules embedded in a water environment, namely solvation energies of benzene, the benzene dimer, ortho-benzyne and the singlet-triplet gaps of ortho-benzyne in explicit water environments of various sizes. For all these systems we obtained accurate results in the description of the interactions between the electrons and QDOs, demonstrating how the latter can be used to create a polarizable environment dynamically correlated to the electronic subsystem.

Moreover, this approach has the advantage of introducing an embedding technique for QMC methods, and especially DMC, in which both the environment (QDOs) and the subsystem of interest (the electrons) are treated at the same level of theory able to capture the relevant physical properties of the total system. This overcomes the intrinsic difficulties that appear for example when one tries to embed other Quantum Chemistry methods or Density Functional Theory with a stochastic approach such as QMC [193].

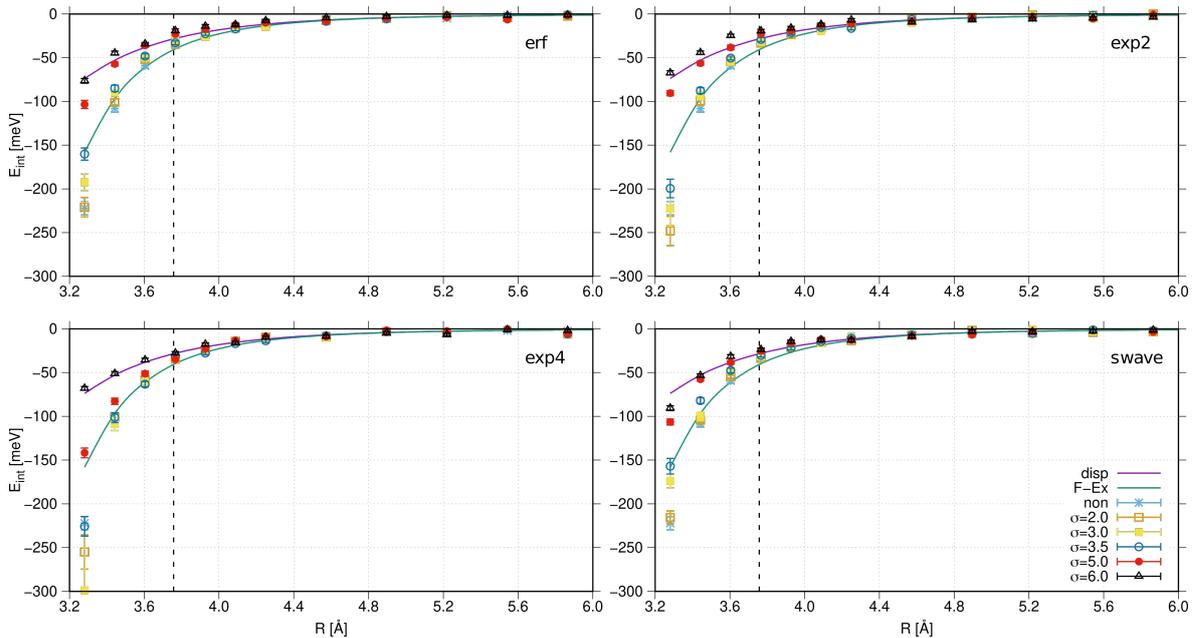


Figure 4.17: Interaction energies of Ar_2 using the El-QDO model at the DMC level of theory using the uniform damping functions from Eqs. 2.60, 2.61, 2.62 and 2.63 for various values of the parameter σ . The results are compared to the pure dispersion (Dsp) and the sum of all terms not including exchange (F-Ex) from the El-El SAPT decomposition [161] and to the El-QDO model interacting via the bare Coulomb potential without the damping (bare). The dashed vertical lines represent the equilibrium distance of the atoms.

We presented the results of the El-QDO method calculated using the QMC algorithms, but these can be implemented in other QCH methods, such as Coupled Cluster or Configuration Interaction frameworks, which are also able to capture the non-covalent interactions.

Clearly, some challenges, such as the description of the short-range repulsive region of the interactions, still remain and will be the inspiration for future investigations. After overcoming these challenges, the El-QDO method might be an alternative to the standard QM:MM methods, with lower levels of empiricism and explicit polarization and dispersion effects for post-Hartree-Fock methods.

Thus, to conclude, the results presented in this Chapter open the way to a general embedding approach of electrons in a bath of QDOs that can be parametrized to reproduce, together with auxiliary point charges, the behavior of molecular sys-

tems in a realistic environment. This will allow us to study for example the collective effects of long-range non-covalent interactions of large environments on the polarization, structural properties and electronic excitations of subsystems in an accurate and computationally feasible way.

Binding mechanism of the QDO model

Parts of this Chapter will be published in this or a similar form in:

M. Ditte, M. Barborini, A. Tkatchenko. “Quantum Drude Oscillators Coupled with Coulomb Potential as an Efficient Model for Bonded and Non-Covalent Interactions in Atomic Dimers” In preparation.

and have been produced in collaboration with the above authors.

In this Chapter we investigate the behavior of the QDO model described in Sec. 2.4 in the short-range region. The understanding of the model, when the distances between QDOs are shorter than the typical distances of dispersion interactions, is crucial for the further generalization of the model, with the goal of describing a larger set of chemical bonds.

Without loss of generality, we study the general homogeneous QDO dimer (both QDOs have the same parametrization). First, we compare the performance of all the *ansätze* defined in Sec. 3.6.2. Next, we study the effect of the three parameters of the model (ω, μ, q) on the potential energy surface at all distances. In the last, part we

define observables, which help us to underline the differences between the QDO dimer and the H_2 molecule.

5.1 The homogeneous QDO dimer

Before analyzing the variation of the QDO model's properties in terms of their frequency ω , the mass μ , and the charge q varied independently with respect to the reference case $\mu = q = \omega = 1$, we must study the convergence of the variational wavefunctions described in Sec. 3.6.2, namely Dip., Dip.+J. and Mol.+J, constructing the PESs at VMC and DMC levels for the homogeneous QDO dimer ($\mu = q = \omega = 1$).

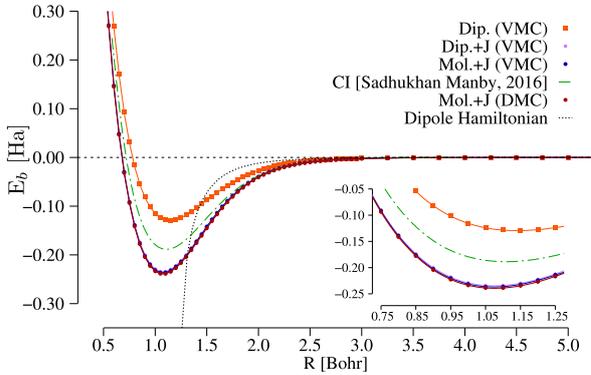


Figure 5.1: The binding energy E_b of the QDO homogeneous dimer with $\omega = \mu = q = 1$. The black dotted line represents the solution of the Hamiltonian in the limit of dipole approximation (Eq. 2.53). The CI results are taken from Ref. 152.

From the results displayed in Fig. 5.1 we can see how the Dip. wavefunction is unable to properly reproduce the binding energy profile in the short range since it can not describe the tunneling effects (see Fig. 3.1) of the drudon from its center towards the attractive Coulomb potential on the opposite QDO. The introduction of the Jastrow factor in the Dip.+J wavefunction greatly improves the variational results since it is able to take into account both the correlation of the drudon around the attractive Coulomb potential and the two-body correlation of the two drudons that become important as the two QDOs

come together. In fact, at the VMC level, the latter wavefunction recovers more than 0.10 Ha in the binding energy with respect to the former. Finally, a small improvement, of about 2 mHa at the equilibrium distance, with respect to the Dip.+J wavefunction can be obtained with Mol.+J, due to the fact that the single-particle orbitals also account for the tunneling effects of the drudons in the short range. Moreover, the VMC PES obtained with the Mol.+J trial wavefunction is nearly identical to the converged

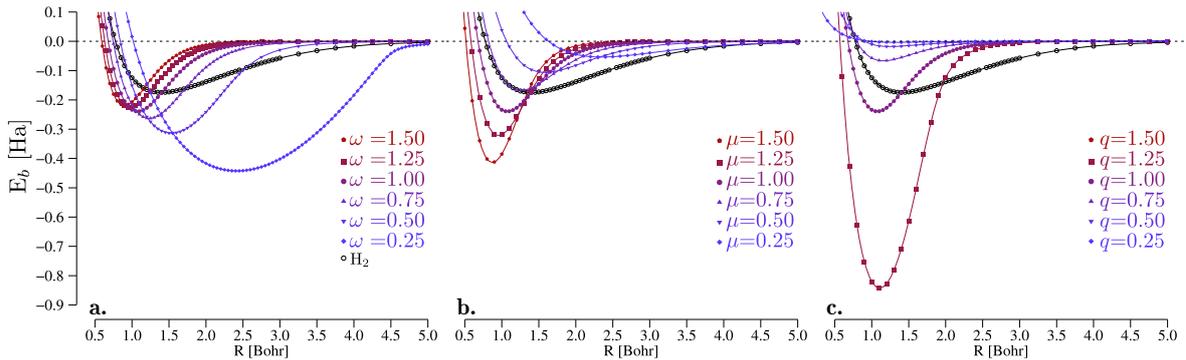


Figure 5.2: The DMC binding energies E_b of the homogeneous QDO dimer obtained with the Mol.+J wavefunction (Eq. 3.86): **a.** $\mu = q = 1$ and variable $\omega \in [0.25, 1.50]$; **b.** $\omega = q = 1$ and variable $\mu \in [0.25, 1.50]$; **c.** $\omega = \mu = 1$ and variable $q \in [0.25, 1.50]$.

DMC energy obtained with the same wavefunction, with a difference of about 2.3 mHa at the equilibrium distance. As discussed in Sec. 3.5.4, since the system is built from distinguishable particles DMC will always converge to the exact solution no matter the trial wavefunction employed. In Fig. 5.1 we also compare our results with the full CI PES, previously reported in Ref. 152 for the same system. As previously anticipated, it is clear that due to how the configurational space was constructed, the authors were unable to properly converge their calculations towards the exact solution obtained with DMC, reporting a maximum discrepancy with our results of more than 0.05 Ha.

Considering these initial results, from now on, we will always employ the Mol.+J trial wavefunction as the guiding function for the DMC calculations used to study the behavior of the QDO model in the limit of full convergence.

5.2 Comparison between the homogeneous QDO dimer and the H_2 molecule

In order to better understand the behavior of the QDO model in describing atomic and molecular interactions, let us compare the homogeneous QDO dimer with $\omega = \mu = q = 1$ to the analogous Coulomb coupled two-electron system, *ie.* the H_2 molecule, considering also how the independent variation of each parameter affects the binding

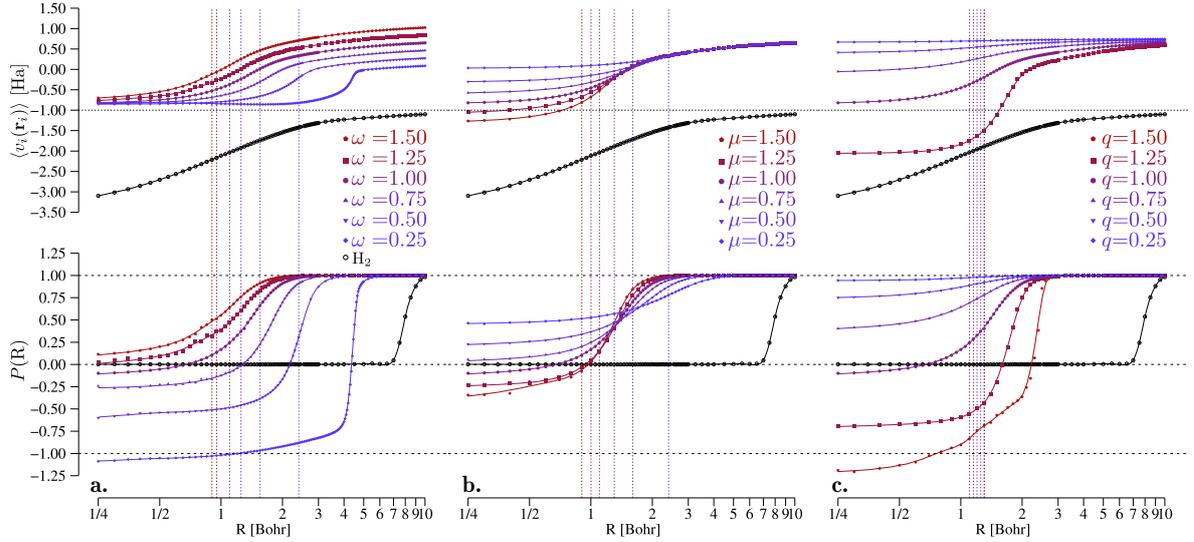


Figure 5.3: One-body DMC observables obtained with the Mol.+J wavefunction, as a function of the relative distance R between the QDO centers (or the nuclei) for the QDO dimer (compared to the H_2 molecule). The upper panels correspond to the single-particle potential $\langle v_i(\mathbf{r}_i) \rangle$ (Eq. 2.42), while the lower panels correspond to the single-particle distances $P(R) = \frac{-\langle z_1 \rangle + \langle z_2 \rangle}{R}$ (the dimers are positioned along the z axis and the origin of the coordinate system is in the center of the bond). As for the previous figures, **a.** $\mu = q = 1$ and variable $\omega \in [0.25, 1.50]$; **b.** $\omega = q = 1$ and variable $\mu \in [0.25, 1.50]$; **c.** $\omega = \mu = 1$ and variable $q \in [0.25, 1.50]$.

energies and other properties.

In both systems, the drudons (for $\mu = q = 1$) and the electrons are distinguishable particles of unitary charge and mass, and the QDO centers act as fixed nuclei of unitary charge. The only difference between the QDO dimer and the H_2 molecule is in the single-particle potential defined for the QDOs in Eq. 2.42 (Fig. 3.1). While the single-particle potential of the electrons is always identical and defined as $v(\mathbf{r}_i) = \sum_{n=1}^{N_n} \frac{q_n Z_n}{|\mathbf{r}_i - \mathbf{R}_n|}$ (where N_n is the number of atomic nuclei in the system, Z_n is the nuclear charge and \mathbf{R}_n is the nuclear position), for the QDOs this potential is always different and differs only in the quadratic interaction between the drudonic particle and its QDO center. Here the characteristic frequency of the QDOs ω that is directly connected to the localization of the drudons around their oscillating centers is used as a parameter to study the behavior of the model in describing the interactions.

First, in Fig. 5.2a we report the PES of the QDO dimer with different values of

$\omega \in [0.25, 1.50]$, compared to the PES of the H_2 molecule. In the QDO model, as the frequency diminishes, the drudons become progressively delocalized with a consequent lowering of the binding energy and elongation of the equilibrium ‘bond’ length. For H_2 at distances of around 8~9 Bohr the binding energy in the PES is already negative, with respect to that of the QDOs which is still basically null, due to the fact that the two electrons of the molecule are already delocalized (we will see this later), while the drudons remain localized due to the height of the quadratic potential’s well that is governed by the frequency (Fig. 5.2a) and the mass (Fig. 5.2b). Yet, as ω diminishes, the two drudons become monotonically more delocalized, while for μ , since it also affects the kinetics of the system, the change in the form of the PES is not consistent for all values of R (Fig. 5.2b). At short distances, the decrease of the mass increases the kinetic energy and the overall energy, while at large distances the decrease of the mass increases the binding. Finally, since the charge affects in particular the Coulomb attraction, between the drudons and their opposite QDO centers, in the short range it increases the binding energy, without affecting the long-range interaction region (Fig. 5.2c) for which the QDOs remain in average neutrally charged objects and only polarization affects the interaction. Although the charge doesn’t seem to greatly affect the binding energy in the long range, this is not the case for the response properties on which the model is usually parameterized[88, 62, 38].

In order to better understand the drudons’ delocalization as a function of the parameters it is best to examine some of the single-particle and two-particle observables.

Regarding the single-particle observables, it is fundamental to study the external one-body potential as a function of the distance between the two QDO centers (or atomic nuclei, for the H_2 molecule) (upper panels in Fig. 5.3). Moreover, we can define a parameter that is connected to this quantity and is related to the probability of the two drudonic (or fermionic) particles to form a bond. By assuming that the molecular axis is along the z coordinate and that the center of mass of the homogeneous dimers is at the origin of the reference system, we can define the function

$$P(R) = \frac{-\langle z_1 \rangle + \langle z_2 \rangle}{R}, \quad (5.1)$$

where $\langle z_1 \rangle$ and $\langle z_2 \rangle$ are the average positions of the two particles along the z axis computed through the DMC sampling and $R = |\mathbf{R}_1 - \mathbf{R}_2|$ is the distance between the two QDOs or atomic centers (lower panels in Fig. 5.3).

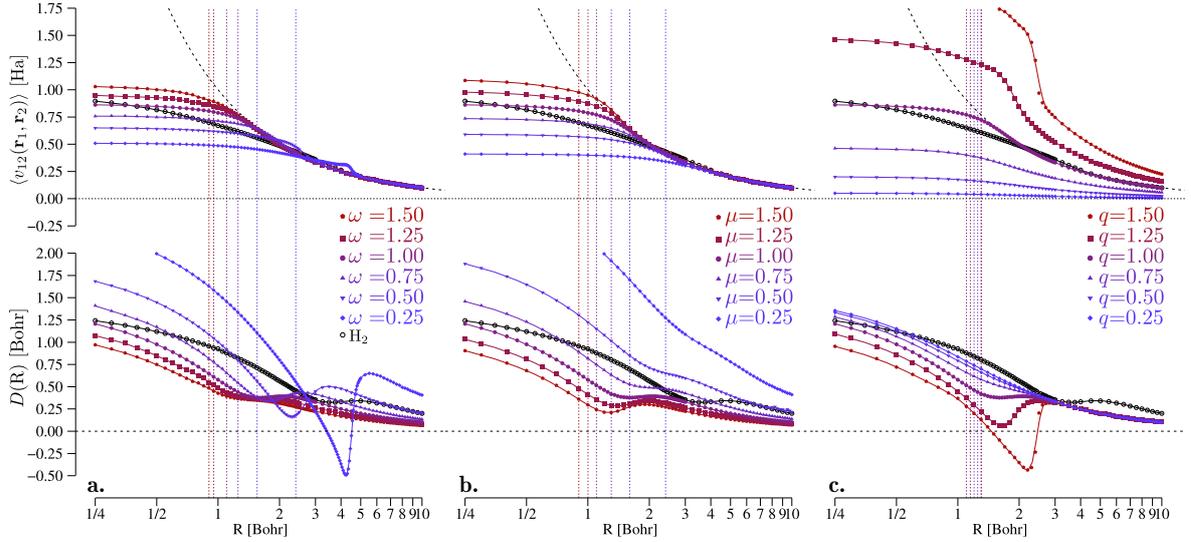


Figure 5.4: Two-body DMC observables obtained with the Mol.+J wavefunction, as a function of the relative distance R between the QDO centers (or the nuclei) for the QDO dimer (compared to the H_2 molecule). The upper panels correspond to the average two-body Coulomb potential energy $\langle v_{12}(r_1, r_2) \rangle$ (Eq. 2.41), while the lower panels correspond to the two-particle distances minus the two-center distance $D(R) = \langle r_{12} \rangle - R$ (the dimers are positioned along the z axis and the origin of the coordinate system is in the center of the bond). As for the previous figures, **a.** $\mu = q = 1$ and variable $\omega \in [0.25, 1.50]$; **b.** $\omega = q = 1$ and variable $\mu \in [0.25, 1.50]$; **c.** $\omega = \mu = 1$ and variable $q \in [0.25, 1.50]$.

In the limit of non-interacting particles, that are on average positioned around their oscillating centers, the values of the average positions $\langle z_1 \rangle$ and $\langle z_2 \rangle$ with respect to the origin of the reference will be equal respectively to $\langle z_1 \rangle \approx -R/2$ and $\langle z_2 \rangle \approx R/2$ (depending on the choice of the position of the first and second atom or QDO), *ie.* $P(R \rightarrow \infty) = 1$.

In the case of the H_2 molecule, the average value of the symmetric single-particle potential $\langle v_i(\mathbf{r}_i) \rangle$ which is identical for both electrons, in the dissociation limit converges to the value of -1.0 Ha (average potential energy of the isolated hydrogen atom) and progressively decreases due to the interaction of the electrons with the opposite attractive center (upper panes Fig. 5.3). This same behavior is mimicked by the function $P(R)$ that converges from the value of 1, in the non-interacting limit, towards 0 signaling the formation of a chemical bond, for which the two electrons are ‘shared’ between the two atomic centers and are thus in average distributed around the center of the

bond.

For the homogeneous QDO dimer, on the other hand, in the dissociation limit the single-particle energies converge towards the average potential energies of isolated harmonic oscillators, *ie.* $\frac{3}{4}\omega$ (upper panels Fig. 5.3), and as the QDOs come close to each other, the single-particle energy tends towards that of a charged particle in the attractive Coulomb potential (with a shift due to the quadratic potential) signaling again that the drudons are shifting towards the opposite center, as suggested also by the behavior of the function $P(R)$, in the lower panels of Fig. 5.3. As the two QDOs start interacting they come closer to the center of mass of the molecule, *ie.* $P(R) < 1$, until the value of the function becomes negative, which corresponds to the case in which the two drudons are now closer to the opposite center. The limit of $P(R) = -1$ corresponds to the exact inversion of the drudonic particles that now on average are localized around the opposite QDO center. This happens because, for small values of the frequency ω or large values of the charge q , the drudonic particles tend to tunnel through the potential barrier, drifting away from their oscillating center and getting caught in the attractive Coulomb potential of the opposite QDO. Mass (Fig. 5.3b) and charge (Fig. 5.3c) have a different effect on the single-particle energy in the short-range limit since the former affects the kinetic energy and the Coulomb attraction, while increasing the latter increases proportionally the attraction lowering the single-particle energy. This polarization effect is responsible for the absolute minimum of the PESs of the QDO model (Fig. 5.2). The tunneling effect is also induced by the increase of μ that lowers the kinetic energy contribution in comparison with the attractive Coulomb. It is thus clear that in the QDO model, we can not really identify a chemical bond (for which $P(R) = 0$) in the proper sense, even in the short-range limit.

A second set of interesting observables is that of the two-body Coulomb potential $v_{12}(\mathbf{r}_1, \mathbf{r}_2)$ (upper panels of Fig. 5.4) and of the average two-particle distance minus the distance between the QDO centers (or atoms)

$$D(R) = \langle r_{12} \rangle - R, \quad (5.2)$$

where $\langle r_{12} \rangle$ is the Monte Carlo average of the module distance $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ between the two quantum particles computed on the entire DMC sampling (lower panels in Fig. 5.4). In the non-interacting limit, both the two-body repulsion and the function $D(R)$ converge to zero, *ie.* $D(R \rightarrow \infty) \rightarrow 0$, since the average position of the two

quantum particles will be around their oscillation centers (or nuclear centers), and the interaction will be zero.

On average for both the QDO model and the H₂ molecule, $D(R)$ is positive, signaling that the average distance of the two quantum particles is larger than that between the QDO or nuclear centers. Despite this, it is interesting to see that for small values of the frequency (for example $\omega = 0.25$) or high values of the charge (for example $q = 1.50$) it happens that $D(R) < 0$, meaning that the two drudons are closer to each other compared to their centers (lower panels of Fig. 5.4a and Fig. 5.4c). Thus, this drifting phenomenon happens when the frequency of the quadratic potential becomes too low compared to the Coulomb attraction between the drudon and the opposite center, leading to the tunneling phenomena previously discussed. Consequently, due to this drift the average repulsive energy between the two particle potential $v_{12}(\mathbf{r}_1, \mathbf{r}_2)$ shown in the upper panels of Fig. 5.4 has a corresponding local maximum due to the increase of the repulsion energy.

In conclusion, the comparison between the QDO dimer with the H₂ molecule has highlighted the overall behavior of the model, which is ascribed to the asymmetry of the single-particle potential energies and the form of the quadratic potentials of the oscillating centers. Due to the overall single-particle potential energies, the model is unable to define a covalent chemical bond via accumulation of the charge density between the QDOs, and the absolute minimum in the PES is due to the Coulomb potentials that attract the drudonic particles towards the opposite QDO center counterbalancing the localization induced by the quadratic potential energy. If the harmonic potential is not sufficiently high (low values of ω), the induced polarization effect becomes too strong and the drudons spontaneously drift towards the opposite centers, undergoing a process that is similar to a phase transition in which the distance between the QDO centers is the order parameter.

Despite this limitation, it is still possible to find a set of parameters ω , μ , and q , so that the QDO model can also reproduce the PES of covalent bonds. This can be achieved for instance by defining a set of parameters that are dependent on the relative distances between the QDO centers. For example, considering the simplest case of the H₂ molecule, it is possible to reproduce its PES by varying the frequency ω of the QDO model as a function of the distance (Fig. 5.5).

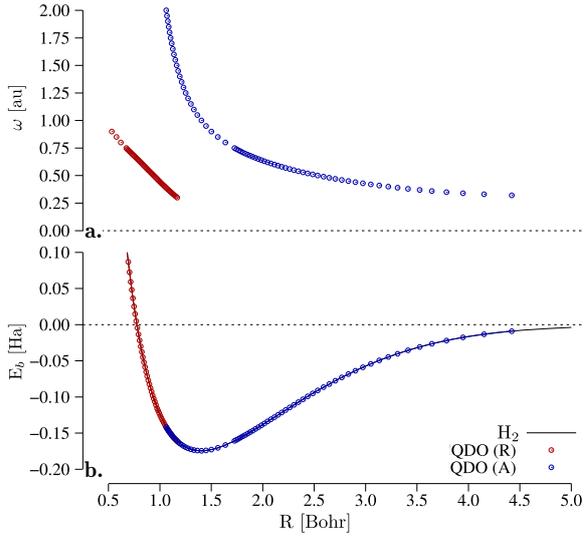


Figure 5.5: **a.** ω as a function of the position R for which the corresponding QDO(ω) curves intersect the PES of the H_2 molecule. For a small distance of ≈ 1 bohr, we observe that the QDO(ω) curves intersect the PES of H_2 in two points. **b.** QDO(ω : R) PES (red and blue circles) obtained by varying the frequency according to the values in panel **a**, compared with the PES of H_2 (black line) obtained with DMC method.

By varying the frequency ($\mu = q = 1$) we obtain a set of QDO PESs each intersecting the H_2 PES at various values of the distance R , *ie.* $E[H_2(R)] + 2E[H] = E[QDO_2(\omega; R)] - 2E[QDO(\omega)]$, that corresponds to setting $E[H_2(R)] - E[QDO_2(\omega; R)] + 1 + 3\omega = 0 \quad \forall \quad R$. By plotting the values of the frequencies ω for which the corresponding QDO curves intersect the H_2 PES, as a function of the interatomic distance R , we obtain the data displayed in Fig. 5.5a. By selecting for each R the values of ω to match the H_2 PES we obtain the binding profile shown in Fig. 5.5b.

Here, we must remark that these are only a set of preliminary tests where only the frequency of the QDO model was changed to reproduce a given binding energy curve for a chemically bonded system. Yet, we hypothesize, that matching

the response properties (polarizability and dispersion coefficients) for QDO and electronic systems for different values of R might lead to a more promising QDO model for chemical bonding. In addition, a QDO model for all relevant interatomic distances might need to be based on a more general effective single-particle potential. We defer the assessment of these promising directions to future work.

5.3 Conclusions

In the last decades, the QDO model has been successfully applied to describe dispersion interactions in Quantum Chemistry and solid state physics, being at the basis of modern dispersion theories[115, 72] such as MBD[179]. In order to describe these long-

range interactions the model is usually exploited in the limit of dipole interactions since the Hamiltonian can be easily diagonalized without introducing a significant computational overhead to the methods in which it is introduced, such as DFT[9, 146]. Thus far, the focus has primarily been on the long-range distances of those interactions that the method is able to predict based on a parametrization obtained from the knowledge of the physical response properties of the systems[8, 31, 47, 62]. However, in the short-range limit, the dipole-Coupled QDO model has imaginary solutions, while the QDO model with the Coulomb potential, although numerically stable, lacks the general parametrization necessary to reproduce the correct interaction energies, requiring additional corrections, such as dumping functions[204, 38, 9], and empirical potentials[87, 86, 88, 31, 38] or energy contributions coming from other Quantum Chemistry methods, such as the HF exchange[152]. One approach for reintroducing the short-range repulsion was suggested in Ref. 47, assuming that in a homogeneous QDO dimer the two Drude particles representing the entirety of the valence electrons of a closed shell atom in a spinless state, could be considered as indistinguishable bosons. Starting from this assumption the authors symmetrized the wavefunction following the same approach used by W. Heitler and F. London[70] to symmetrize the spatial part of the electronic wavefunction of the H_2 molecule. Although this procedure has been shown to give consistent results for a perturbative approach[47, 190, 62], and could be justifiable within the limit of the dipole interaction, for which the homogeneous dipoles can be considered as identical quasi-particles, it becomes inapplicable in the limit of full Coulomb interactions in which the drudonic degrees of freedom become decoupled with respect to the positions of their oscillating centers[152, 31, 38], also making the extension of this approach to higher asymmetric multipole moments problematic[190, 62]. As discussed in this work, in the limit of full Coulomb interaction each drudon exists in a distinct single-particle potential, making them distinguishable particles, no matter their parametrization, and consequently, the application of a particular spin statistics becomes non-trivial.

To define a more general parametrization of the short-range potential, the ‘nearly exact’ solution of the QDO model presented in this work stands as an important milestone since it presents a variational framework to pursue the construction of a univocal and general potential to better mimic the short-range interaction of a wider set of chemical bonds, starting in particular with vdW and hydrogen bonds. This

can be achieved, for example through the construction of univocal anisotropic effective potentials, parameterized through modern machine-learning methods or Neural-Networks [189, 175, 188, 73], that apart from avoiding the critical behavior of the model in the short range, would also correctly reproduce the electronic impermeability of the chemical complexes, preserving its intrinsic efficiency[31, 38].

An alternative approach, shown in this Chapter for the case of the H_2 molecule, could be that of introducing an explicit dependence of the QDO parametrization on the relative distances between the oscillation centers.

Among the many possible directions, our future works will explore a broader set of chemically bonded systems, different ways to parameterize one-particle and two-particle potentials in the QDO model, the transition regime between bonded and non-bonded interactions, going beyond atomic dimers, and developing an efficient parameterization of the variational wavefunctions for Coulomb-coupled QDOs.

Computational study of TeMePh_2^+ interacting with OPPh_3 in the DCE solvent

Parts of this Chapter have been published in this or a similar form in:

L. Gros Lambert, Y. Cornaton, M. Ditte, E. Aubert, P. Pale, A. Tkatchenko, J-P. Djukic, V. Mamane. “Affinity of Telluronium Chalcogen Bond Donors for Lewis Bases in Solution: A Critical Experimental-Theoretical Joint study” *Chem. Eur. J.* e202302933 (2023)

and have been produced in collaboration with the above authors. The LED/DLPNO-CCSD(T) calculations and the calculations including the COSMO solvation model have been done by Dr. Y. Cornaton.

This Chapter is dedicated to the computational study of the affinity of a Lewis acid, namely the diphenylmethyltelluronium cation TeMePh_2^+ with triphenylphosphine oxide OPPh_3 Lewis base, which form a Lewis donor-acceptor complex through a localized chalcogen bond in the solution of ethylene dichloride (DCE).

This work was motivated by the necessity to interpret the experimental measurement of the enthalpies of telluronium salts $[\text{Ar}_2\text{MeTe}]\text{X}$, where X are various counterions, using Isotherm Titration Calorimetry (ITC). The enthalpies for the first

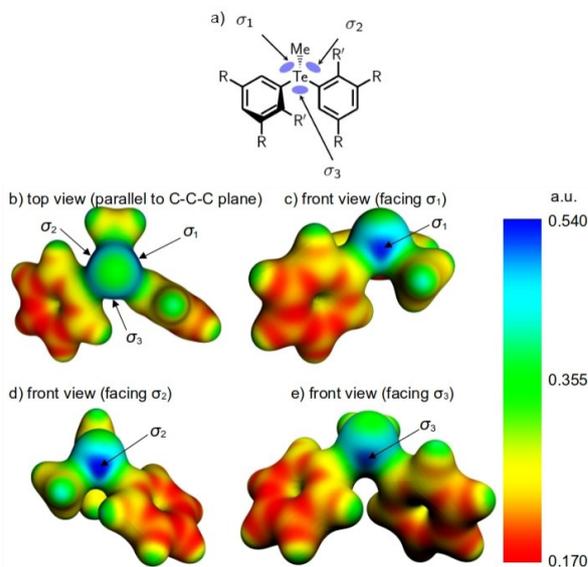


Figure 6.1: a) Numbering and position of the σ -holes σ_{1-3} in Ar_2MeTe^+ ; b)-e) different views of the DFT-determined electrostatic potential map for TeMePh_2^+ drawn on an isosurface of the density (0.03 \AA^{-3}) showing the localization of σ_{1-3} .

association of OPPh_3 with a series of substituted telluronium salts in the DCE solvent were measured to be in the range from -0.5 to -5 kcal/mol . The comparison of the measured values with the theoretical predictions using state-of-the-art methods was necessary to understand the association mechanism. We are especially interested in the importance of the van der Waals contributions to the binding mechanism and to the effects of the solvent. Due to the size of the studied system, we have chosen to apply state-of-the-art DFT calculations using the PBE functional with different dispersion schemes described in Secs. 3.3 and 3.4. We compare the performance of the various dispersion methods and compare various solvation approaches in order to understand the effects of the solvent, namely: the reactants in the vacuum; the molecules in a conductor-like screening model of implicit solvation, *ie* COSMO [103], where the environment is included as a continuum with a fixed value of the permittivity; the molecules in a partial explicit solvent, using one explicit DCE molecule per free σ -hole (Fig. 6.1); and a hybrid approach, where one explicit DCE molecule per free σ -hole is complemented by the implicit COSMO model.

The molecules involved in the reactions are: the diphenylmethyltelluronium cation TeMePh_2^+ (Te - tellurium; Me - methyl CH_3 ; Ph - phenyl C_6H_5), which is one of the Lewis acids from the diarylmethyltelluronium cations Ar_2MeTe^+ group, where Ar is a general aryl group (see Fig. 6.2a); triphenylphosphine oxide OPPh_3 (O - oxygen; P - phosphorus; Ph - phenyl C_6H_5), which is on the other hand a organophosphorus

The molecules involved in the reactions are: the diphenylmethyltelluronium cation TeMePh_2^+ (Te - tellurium; Me - methyl CH_3 ; Ph - phenyl C_6H_5), which is one of the Lewis acids from the diarylmethyltelluronium cations Ar_2MeTe^+ group, where Ar is a general aryl group (see Fig. 6.2a); triphenylphosphine oxide OPPh_3 (O - oxygen; P - phosphorus; Ph - phenyl C_6H_5), which is on the other hand a organophosphorus

Lewis base, often abbreviated also as TPPO (see Fig. 6.2b); and ethylene dichloride $C_2H_4Cl_2$ often written as 1,2-dichloroethane, DCE or EDC, which is a solvent in which the association of $TeMePh_2^+$ and $OPPh_3$ was studied (see Fig. 6.2c).

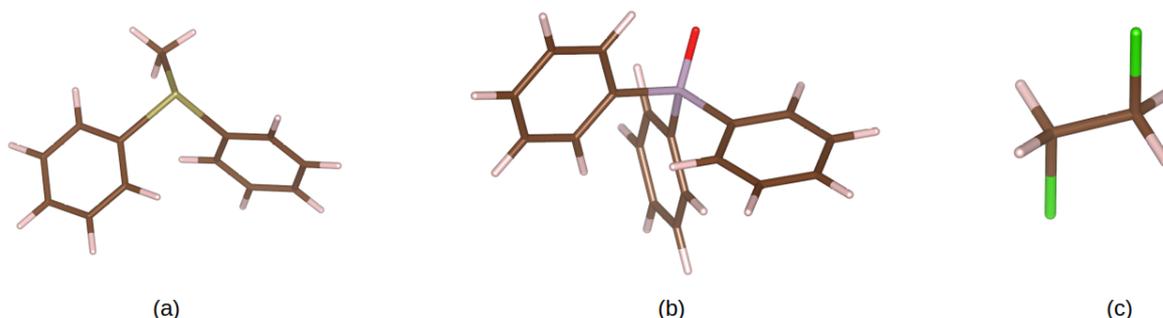


Figure 6.2: (a) $TeMePh_2^+$, (b) $OPPh_3$, and (c) DCE. The color coding: pink - H; brown - C; red - O; purple - P; green - Cl; and yellow - Te.

Through this Chapter, we study 1:1, 1:2 and 1:3 adducts of $TeMePh_2^+$ with $OPPh_3$, defined as:

- 1:1 adducts (3 options for the σ -hole occupancy: $\sigma_1, \sigma_2, \sigma_3$):

$$TeMePh_2^+ + OPPh_3 \rightarrow [(Ph_3PO) \cdot TeMePh_2]^+$$

$$(C_2H_4Cl_2)_3 \cdot TeMePh_2^+ + OPPh_3 \rightarrow [(C_2H_4Cl_2)_2(Ph_3PO) \cdot TeMePh_2]^+ + C_2H_4Cl_2$$
- 1:2 adducts (3 options for the σ -hole occupancy: $\sigma_{1,2}, \sigma_{1,3}, \sigma_{2,3}$):

$$TeMePh_2^+ + 2 OPPh_3 \rightarrow [(Ph_3PO)_2 \cdot TeMePh_2]^+$$

$$(C_2H_4Cl_2)_3 \cdot TeMePh_2^+ + 2 OPPh_3 \rightarrow [(C_2H_4Cl_2)(Ph_3PO) \cdot TeMePh_2]^+ + 2 C_2H_4Cl_2$$
- 1:3 adduct (1 option for the σ -hole occupancy: $\sigma_{1,2,3}$):

$$TeMePh_2^+ + 3 OPPh_3 \rightarrow [(Ph_3PO)_3 \cdot TeMePh_2]^+$$

$$(C_2H_4Cl_2)_3 \cdot TeMePh_2^+ + 3 OPPh_3 \rightarrow [(Ph_3PO) \cdot TeMePh_2]^+ + 3 C_2H_4Cl_2,$$

where for each adduct the first chemical equation corresponds to the association in the vacuum or in the COSMO implicit solvent model, and the second chemical equation corresponds to the association in the partial explicit or hybrid solvation with one explicit DCE molecule per free σ -hole. An example of the 1:1 adduct in the partial explicit DCE solvent is shown in Fig. 6.3. All the other possibilities can be found in Appendix A2.

The Chapter is structured as follows: in Sec. 6.1 we specify the computational

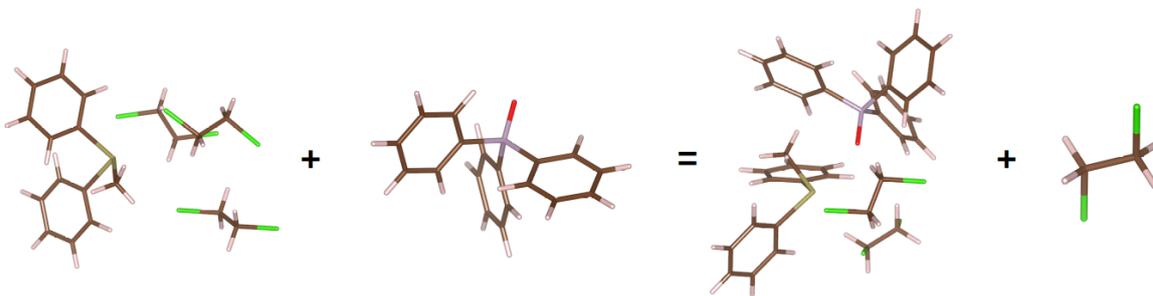


Figure 6.3: 1:1 adduct for σ_1 with 1 explicit DCE molecule per free σ -hole: $(\text{C}_2\text{H}_4\text{Cl}_2)_3 \cdot \text{TeMePh}_2^+ + \text{OPPh}_3 \rightarrow [(\text{C}_2\text{H}_4\text{Cl}_2)_2(\text{Ph}_3\text{PO}) \cdot \text{TeMePh}_2]^+ + \text{C}_2\text{H}_4\text{Cl}_2$

details; in Sec. 6.2 we describe the calculations of the enthalpies and Gibbs free energies in the ideal gas approximation; in Sec. 6.3 we apply the local energy decomposition of domain-based local pair natural orbital CCSD(T) to the 1:1 adduct in the vacuum, in order to analyze the energy components; in Sec. 6.4 we calculate the thermochemistry quantities of all the adducts in the implicit solvent model and hybrid solvation approach using pairwise and many-body dispersion methods, and in the vacuum and partial explicit solvent, analyzing in more detail the effects of a broader palette of the dispersion methods; and finally in Sec. 6.5 we show supporting analysis of the convergence of some of the previously discussed results.

6.1 Computational details

TightSCF and TightPNO thresholds were used for the domain-based local pair natural orbital CCSD(T) [150] calculations in the Orca program package, version 4.2.0. [136]. The Karlsruhe 2nd generation default triple- ζ valence with polarization (def2-TZVP) basis set [210] was used in these calculations, as well as RIJK auxiliary basis set [201].

Fixed-node diffusion Monte Carlo [52, 95, 14] calculations were carried out using the QMeCha quantum Monte Carlo package [12]. For the trial wavefunction we have used a Slater determinant, whose molecular orbitals are obtained from DFT calculations executed with the ORCA package [136], using the PBE functional, BFD effective core potentials (ECPs) and the corresponding cc-pVTZ basis set [22] with the addition of augmented functions from the all-electron aug-cc-pVTZ basis set [98, 205]. To reduce

the wavefunction’s variance we added a Jastrow factor similar to that introduced in Marchi et al. [123, 26] and optimized with the stochastic reconfiguration approach [26]. To guarantee the absence of the size-consistency error in the DMC calculations, we compare the convergence of DMC energies as a function of the time step $d\tau$ obtained with two different energy cut-off schemes, a novel size-consistent one from Barborini et al. (labeled as 'MB') [1] and that proposed by Zen et al. (labeled as 'Zen') [207].

All the calculations including COSMO implicit solvent were done in the SCM ADF2022.01 package [178], using scalar relativistic corrections with the Zeroth Order Relativistic Approximation [192], all-electron triple- ζ Slater-type with single polarization function basis set (TZP) [191] and PBE functional [143]. The parameters of the COSMO implicit solvent were set to $\epsilon = 10.66$ and $r = 3.15 \text{ \AA}$, corresponding to the DCE environment.

All the calculations in the vacuum and in the partial explicit solvent without the COSMO implicit solvent were done in FHI aims (version 221103) [2], ASE [112], and the dftd4 package [25] softwares, using scalar relativistic corrections with the Zeroth Order Relativistic Approximation [192], PBE [143] and PBE0 [4] functionals, with all-electron light, intermediate and tight basis sets composed of numerical atomic orbitals [69].

6.2 Thermochemistry quantities

In this section, we discuss the calculations of two thermochemistry quantities, namely the enthalpy H and Gibbs free energy G . The equations in this section are written as implemented in the Atomic Simulation Environment (ASE) [112], in the SI units. More detailed derivation can be found for example in Ref. 35.

Both, the enthalpy and Gibbs free energy, are calculated in the limit of an ideal gas, assuming that all the spatial degrees of freedom are independent and thus can be separated into the translational, rotational, and vibrational degrees of freedom.

The enthalpy of an ideal gas, which is independent of pressure, at temperature T

can be calculated as

$$H(T) = E_{\text{elec}} + E_{\text{ZPE}} + \int_0^T C_p dT, \quad (6.1)$$

where the first term is the electronic energy, the second term is the zero-point energy and the third term is an integral over the heat capacity at constant pressure. The heat capacity can be separated into the translational, rotational, vibrational and electronic heat capacities

$$C_p = k_B + C_{V,\text{trans}} + C_{V,\text{rot}} + C_{V,\text{vib}} + C_{V,\text{elec}}, \quad (6.2)$$

where the first term is for switching from constant volume to constant pressure. The translational heat capacity of a 3-dimensional gas is $3/2k_B$, and the rotational heat capacity is

$$C_{V,\text{rot}} = \begin{cases} 0, & \text{if monatomic} \\ k_B, & \text{if linear} \\ \frac{3}{2}k_B, & \text{if nonlinear.} \end{cases} \quad (6.3)$$

The electronic heat capacity is assumed to be 0 and the integrated form of the vibrational part is

$$\int_0^T C_{V,\text{vib}} dT = \sum_i^{\text{vib DOF}} \frac{\epsilon_i}{e^{\epsilon_i/k_B T} - 1}, \quad (6.4)$$

with $3N-6$ and $3N-5$ degrees of freedom for nonlinear and linear molecules respectively (N is the number of atoms), where $\epsilon_i = h\omega_i$ are the energies corresponding to the individual vibrational frequencies ω_i .

The Gibbs free energy of an ideal gas is given as

$$G(T, P) = H(T) - TS(T, P), \quad (6.5)$$

where $S(T, P)$ is the entropy as a function of the temperature and pressure, which can be separated into the translational, rotational, electronic and vibrational contributions

$$S(T, P) = S(T, P^0) - k_B \ln \frac{P}{P^0} = S_{\text{trans}} + S_{\text{rot}} + S_{\text{elec}} + S_{\text{vib}} - k_B \ln \frac{P}{P^0}. \quad (6.6)$$

The translational term is given by

$$S_{\text{trans}} = k_B \left\{ \ln \left[\left(\frac{2\pi M k_B T}{h^2} \right)^{3/2} \frac{k_B T}{P^0} \right] + \frac{5}{2} \right\}, \quad (6.7)$$

the rotational term is given by

$$S_{\text{rot}} = \begin{cases} 0, & \text{if monatomic} \\ k_B \left[\ln \left(\frac{8\pi^2 I k_B T}{\sigma h^2} \right) + 1 \right], & \text{if linear} \\ k_B \left\{ \ln \left[\frac{\sqrt{\pi I_A I_B I_C}}{\sigma} \left(\frac{8\pi^2 k_B T}{h^2} \right)^{3/2} \right] + \frac{3}{2} \right\}, & \text{if nonlinear} \end{cases}, \quad (6.8)$$

where I_A, I_B and I_C are the three principle moments of inertia, I is the degenerate moment of inertia for linear molecules and σ is the symmetry number of the molecule. The vibrational contribution is given by

$$S_{\text{vib}} = k_B \sum_i^{\text{vib DOF}} \left[\frac{\epsilon_i}{k_B T (e^{\epsilon_i/k_B T} - 1)} - \ln (1 - e^{-\epsilon_i/k_B T}) \right], \quad (6.9)$$

again with $3N-6$ and $3N-5$ degrees of freedom for nonlinear and linear molecules respectively (N is the number of atoms). And finally, the electronic contribution to the entropy is

$$S_{\text{elec}} = k_B \ln [2 \times (\text{total spin}) + 1], \quad (6.10)$$

which is 0 for singlet electronic states (total spin = 0).

The inputs needed from the quantum chemical calculations are thus the electronic energy E_{elec} and the frequencies of the phonons (obtained in the harmonic approximation at the equilibrium geometries) in order to find the energies ϵ_i and the zero-point energy correction E_{ZPE} .

6.3 LED/DLPNO-CCSD(T) analysis

The analysis of non-local interactions supporting the formation of a molecular complex from two or more fragments requires the physically sound decomposition of an inter-fragment interaction energy term into meaningful canonical interactions. Compared to DFT methods with additional dispersion that consist of native DFT functionals patched with a term accounting for dispersion interactions, wavefunction-based methods give access to physically consistent energy decompositions and to the quantification of non-covalent contributions [79, 156, 116]. By the use of adequate energy decomposition schemes, the mapping of the interactions acting in the periphery of the main

fragment-anchoring interaction is possible and allows us to evaluate the importance of non-covalent interactions in the process of molecular aggregation [59, 58]. Neese and Bistoni outlined the analytical power of the local energy decomposition (LED) [153] and domain-based local pair natural orbital Coupled Cluster with singles-doubles and perturbative triples (DLPNO-CCSD(T)) [150, 149, 148] method which was used to trace down the importance of dispersion in the structuration of agostic C-H metal interactions.

Table 6.1: The binding energies ΔE_0 of $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ with OPPh_3 at position σ_2 optimized at the PBE-D4-COSMO level of theory obtained using the DMC method as a function of the time step $\delta\tau$, for two different energy cut-off schemes, one from Barborini et al. (MB) [1] and that proposed by Zen et al. (Zen) [207].

$\delta\tau$ [a.u.]	ΔE_0 (MB) [kcal/mol]	ΔE_0 (Zen) [kcal/mol]
0.10	-29.7(3)	-31.8(2)
0.05	-32.6(4)	-31.2(4)
0.01	-30.7(6)	-32.0(9)

The DLPNO-CCSD(T) method combines the high accuracy of the coupled-cluster approach with the reduced computational cost [117] stemming from the localization of the orbitals constructed at the Hartree-Fock level. The LED [153] analysis allows us to extract the various contributions of user-defined inter-fragment interactions in a way reminiscent of the energy decomposition analysis (EDA) [211] as applied to DFT calculations. The LED/DLPNO-CCSD(T) analysis was therefore applied to $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ in its PBE-D4-COSMO optimized geometry (see Sec. 6.4 for details) wherein OPPh_3 faces the σ -hole σ_2 (see Fig. A2.2): TeMePh_2^+ being the first fragment and OPPh_3 the second one. The analysis, shown in Fig. 6.4, was carried out in the gas phase by vary-

ing the O-Te distance keeping the fragments' geometries rigid. This approach, although rather simplified, provides a good picture of the interactions that play a significant role in the formation of the molecular complex.

It is also worth to mention here, that due to the known discrepancies between the CCSD(T) and diffusion Monte Carlo methods [52, 95, 14] for large molecular complexes [5], we calculated the gas phase binding energies of $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ at the equilibrium distance using both, the DLPNO-CCSD(T) and DMC methods. The DMC binding energies ΔE_0 of $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ with OPPh_3 at position σ_2 as a function of the time step $d\tau$ can be found in Tab. 6.1. Based on the monotonic behavior of ΔE_0 (Zen) we conclude, that the converged value of the formation energy

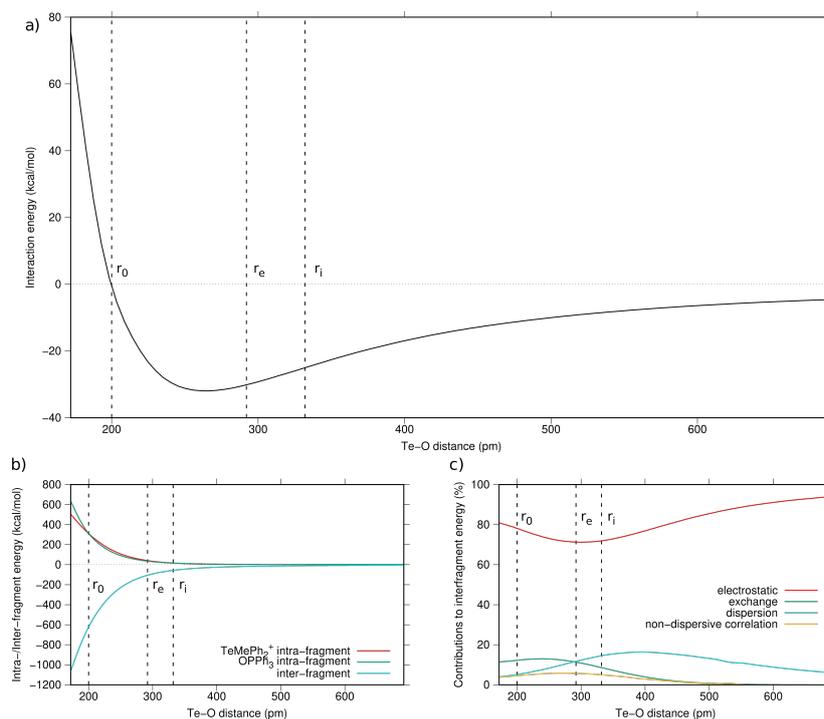


Figure 6.4: a) Interaction energy curve for the $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ adduct along the O-Te distance at the DLPNO-CCSD(T)/def2-TZVP level of theory; b) plot of the intra- and inter-fragment energies; c) the contributions to the inter-fragment energy. The vertical dashed lines show the r_i , r_e and r_0 distances, described in Sec. 6.3.

for given molecular orbitals is $-32.0(9)$ kcal/mol. This value is statistically indistinguishable from -32.5 kcal/mol obtained by DLPNO-CCSD(T) (see Fig. A2.2), which increases the reliability of the LED/DLPNO-CCSD(T) analysis.

The validity of the chosen fragmentation used in the LED [153] analysis was verified by the Mülliken population analysis showing that none of the localized molecular orbitals are delocalized on both fragments for O-Te distances above 152 pm. This also confirmed that the fragments are not strongly covalently bonded even at such O-Te distances shorter than the equilibrium distance r_e (Fig. A2.2). Within the chosen fragmentation scheme, ongoing from a long distance to $r_0 = d(\text{O-Te}) = 200$ pm, the stabilizing inter-fragment contribution surpasses the sum of the intra-fragment ones. At distances longer than ca. 200 pm, even if both fragments are partly individually destabilized in the presence of the other, their mutual interaction nonetheless favors the formation of the adduct through an overly stabilizing inter-fragment energy term. Con-

sidering both intra-fragment contributions, the OPPh₃ fragment is more destabilized than TeMePh₂⁺, except around the equilibrium O-Te distance r_e ($r_e = d(\text{O-Te}) = 292$ pm) (Fig. A2.2b). The stabilizing electrostatic contribution, dominated by electron-nucleus attraction energy terms, represents more than 72% of the inter-fragment energy at any O-Te distance (Fig. A2.2c) with a minimum close to r_e . In turn, the exchange contribution accounts for less than 13% of the inter-fragment energy, with a maximum around r_e . The dispersion force contribution surpasses the exchange one at distances longer than r_e , reaching apex with 17% of the inter-fragment energy around the inflection point distance r_i ($r_i = d(\text{O-Te}) = 332$ pm) of the interaction energy curve. These results demonstrate that the chalcogen bond between TeMePh₂⁺ and OPPh₃ is overwhelmingly non-locally non-covalent in nature and only weakly covalent at a local level as the system reaches r_e . From r_e to shorter distances, the exchange becomes effective and acts as an evanescent intermolecular anchor that opens a narrow channel for charge density transfer. In conclusion, the study carried out at the non-local level confirms the drive of electrostatics supported by dispersion in the self-aggregation of the telluronium cation and OPPh₃ in the gas phase.

6.4 Thermochemistry results

In solution, solvent screening, ion pairing and possible explicit interactions of solvent molecules may temper the attractive electrostatic interactions existing between the telluronium cation and a neutral Lewis base. This complex problem is approached from the viewpoint of thermodynamics by using DFT with additional dispersion methods (see Sec. 3.4) and comparing implicit and explicit solvation schemes using DFT with various dispersion methods.

As shown in the previous section with the DLPNO-CCSD(T) investigation on a rigid model, dispersion force plays a fundamental role in the cohesion of $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ complex; the scrutiny of the quality of the scaling of the dispersion interactions provided for large molecular complexes at the DFT level is therefore justified.

The first calculations of the enthalpy for the 1:1 adduct of the first association of OPPh₃ with TeMePh₂⁺ at position σ_2 using PBE functional with D4 pairwise method [32]

and with the COSMO solvent model (PBE-D4-COSMO), led to a value of $\Delta H = -11$ kcal/mol, which diverges significantly from the best experimental values provided by ITC experiments and the associated thermogram deconvolutions ($\Delta H = -2.9(1)$ kcal/mol or $-1.8(1)$ kcal/mol for two different approaches). This is not surprising, though.

Table 6.2: The enthalpies and Gibbs free energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ at 298.15 K computed at the PBE-D4 and PBE-MBD@rsSCS level of theory in the implicit and hybrid solvation.

method	occ. σ -hole	D4		MBD @rsSCS	
		ΔH	ΔG	ΔH	ΔG
implicit solvent (COSMO DCE)					
1:1	σ_1	-10	+5	-9	-7
	σ_2	-11	+3	-9	+7
	σ_3	-10	+2	-9	+5
1:2	$\sigma_{1,2}$	-23	+3	-25	+2
	$\sigma_{1,3}$	-24	+3	-26	+2
	$\sigma_{2,3}$	-24	+4	-25	0
1:3	$\sigma_{1,2,3}$	-46	-10	-41	0
hybrid solvation (DCE)					
1:1	σ_1	-14	-8	-9	-2
	σ_2	-14	-9	-13	-6
	σ_3	-14	-7	-13	-8
1:2	$\sigma_{1,2}$	-24	-16	-	-
	$\sigma_{1,3}$	-24	-14	-	-
	$\sigma_{2,3}$	-22	-17	-22	-15
1:3	$\sigma_{1,2,3}$	-39	-33	-34	-22

the COSMO model applied to isolated ions may reputedly produce inaccurate solvation

While there exist occurrences of a good match between DFT with additional dispersion and with continuous screening solvation model-computed thermochemistry parameters with ITC data for systems implying neutral molecules or bearing highly delocalized charges [65, 64, 129, 130], significant discrepancies have been often reported for chemical reactions involving salts in which the counterion is systematically omitted in the DFT calculations and in which the solvent is a discrete actor [145, 128, 84].

Due to the discrepancy between the experiment and the first calculations of the thermochemistry quantities, the thermodynamics of the association of one, two and three molecules of OPPh_3 with TeMePh_2^+ were studied with various solvation formulations and with various versions of the dispersion methods.

The question of solvation was addressed using COSMO either fully implicitly or with hybrid implicit COSMO with an explicit amount of solvent in the vicinity of the Te center (1 explicit DCE molecule per free σ -hole) knowing, that

energy contributions [104]. For the DFT methods, the PBE functional was used with two dispersion approaches, namely Grimme’s empirical pairwise dispersion method PBE-D4 [32] and the Tkatchenko’s many-body dispersion method employing the range-separation of the self-consistent screening of polarizabilities PBE-MBD@rsSCS [9] that differ by the construction of their dispersion treatment, which may induce slight differences in the thermodynamics of the association of extended molecular systems. The implicit and hybrid solvation approaches are followed by a more detailed analysis of a broader set of dispersion methods in the vacuum and in the partial explicit DCE solvent without the COSMO model.

For all the results in this section, we have to mind another inaccuracy arising from the rigid rotor harmonic oscillator approximation that questions the reliability of the computed enthalpies and Gibbs free energies for the weakly bonded molecular complexes, for which critical intermolecular vibrational modes in van der Waals complexes are present at low frequencies in the range of $\sim 0\text{-}200\text{ cm}^{-1}$ [94, 171, 57].

So three situations were hence considered:

- a) the implicit solvation by ways of the standard COSMO model;
- b) the hybrid implicit plus partial explicit solvation considering the competition of the Lewis base binding to the Te center with explicitly weakly Te-interacting DCE molecules, i.e. one DCE molecule per available σ -hole; and
- c) in the vacuum and in the partial explicit DCE solvent without the COSMO model.

6.4.1 Implicit solvation

At this level of solvation, the association of the OPPh_3 molecules with TeMePh_2^+ is slightly endergonic, in contradiction with experimental observations (Tab. 6.2). The difference of enthalpy of the first association between the three 1:1 $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ adducts falls below the DFT accuracy limit, seemingly suggesting that the three σ -holes have the same probability of hosting a Lewis base. The same observation still holds when considering the 1:2 adducts, suggesting that the position of the first associated Lewis base has no influence over the position of the association of a second base. The experimental data suggest that real solvation and counterion binding do indeed cancel roughly 80-90% of the stabilizing interaction energy that can be calculated by DFT for

static models taken in the COSMO implicit solvent (Tab. 6.2). To verify the influence of the type of the dispersion method on the divergence between the computed thermodynamics and the experimental ones, the results obtained with the D4 [25] method of the dispersion were compared with calculations with the same native DFT functional, i.e. PBE, but using instead the MBD@rsSCS [9, 179] method for dispersion. From the optimized structures obtained from D4 and MBD@rsSCS with PBE, no major differences were observed in the geometries of the adducts. Again, no major differences between the enthalpies of the association for the two approaches were observed for the 1:1 and 1:2 adducts $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ and $[(\text{Ph}_3\text{PO})_2\cdot\text{TeMePh}_2]^+$ respectively; most values remained within 1-5 kcal/mol of difference. The Gibbs energies of the association diverged significantly though for the 1:3 adduct $[(\text{Ph}_3\text{PO})_3\cdot\text{TeMePh}_2]^+$, with a higher overestimation of exergonicity with the D4 model. If many-body correlation effects can be ruled out as the main source of discrepancy with experimental data for the 1:1 and 1:2 adducts, the MBD@rsSCS method tends to overestimate the Gibbs energies of association less, particularly for the 1:3 adduct.

6.4.2 Hybrid solvation

This hybrid solvation model introduces competition between OPPh_3 and DCE for interacting with the Te center in the thermochemical balance: in the chosen model this entails that each DCE solvent molecule arbitrarily occupies one “free” σ -hole of the Te in TeMePh_2^+ , while the COSMO implicit solvation is kept active. Tab. 6.2 shows that accounting for explicit molecules of solvent does not influence greatly the enthalpies of association of OPPh_3 with TeMePh_2^+ , whereas it significantly influences the Gibbs energies, evidencing the importance of the entropic factor. The hybrid implicit/explicit solvation approach predicts the associations of OPPh_3 with TeMePh_2^+ to be exergonic; the second association being less favored than the first, which somewhat matches the experimental trend. Using the MBD@rsSCS method for the dispersion, instead of the D4 method, induces no significant variations in the geometries and only slight differences in the enthalpies and Gibbs free energies of association when applying the hybrid solvation approach. It is however noticed that for the formation of the 1:3 adduct $[(\text{Ph}_3\text{PO})_3\cdot\text{TeMePh}_2]^+$ the Gibbs free energy of association displays a difference of ca. -9 kcal/mol between the two dispersion methods, where the D4 method again

significantly overestimates the exergonicity of the 1:3 association.

6.4.3 Vacuum and explicit solvation

Due to the large discrepancy between the D4 and MBD@rsSCS for the 1:3 adduct in the implicit and hybrid solution, a further investigation of the impact of the dispersion method was necessary. We thus compared the pure PBE, D4 pairwise method [25], Tkatchenko-Scheffler (TS) pairwise approach [180], the range-separation self-consistent screening version of the many-body dispersion method (MBD@rsSCS) [179, 9] and the non-local many-body dispersion method (MBD-NL) [75]. Due to the missing implementation of the COSMO implicit solvent model in all the software used, the comparison of the various dispersion schemes has been performed in the vacuum and in the partial explicit DCE solvent. Despite the error in total values of ΔH and ΔG , which is due to the over-stabilization in the vacuum compared to the solvent, we argue that there are consistent differences between the pairwise methods and the many-body ones, which are relevant for this work: as a matter of fact MBD@rsSCS and MBD-NL both increase the values of ΔH and ΔG (i.e. make them more positive).

The thermochemistry results obtained using PBE functional and the intermediate basis set are summarized in Tab. 6.3, and show a 1 to 2 kcal/mol agreement between TS and D4 methods (both pairwise) for all the systems except for the 1:3 adduct, where TS gives lower values of about 5 kcal/mol for the ΔG . Both flavors of MBD are within 1 to 2 kcal/mol for the 1:1 and 1:2 adducts and within 3 kcal/mol for the 1:3 adduct for both enthalpies and the Gibbs free energies, with MBD-NL being always higher. In fact, MBD-NL substantially improves the treatment of charge transfer effect and ionicity, when compared to MBD@rsSCS [75]. The many-body treatment of the dispersion interactions persistently leads to higher enthalpies and Gibbs free energies, when compared to the pairwise methods, with differences of 2 kcal/mol, 5 kcal/mol and 9 kcal/mol for the 1:1, 1:2 and 1:3 enthalpies respectively and of 3 kcal/mol, 4-5 kcal/mol and 8 kcal/mol for the 1:1, 1:2 and 1:3 Gibbs free energies respectively. Interestingly MBD gives higher enthalpies and Gibbs free energies also in the case of the 1:2 adducts, where MBD@rsSCS calculations, done using the COSMO implicit solvent model, have the opposite trend. The pure PBE without the dispersion interactions

Table 6.3: The enthalpies and Gibbs free energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ at 298.15 K computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the intermediate basis set in the vacuum and the partial explicit solvent.

method		PBE		D4		TS		MBD@ rsSCS		MBD-NL	
	occ. σ - hole	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG
in vacuum											
1:1	σ_1	-21	-12	-30	-17	-30	-17	-28	-15	-28	-14
	σ_2	-22	-11	-29	-16	-29	-16	-27	-14	-27	-13
	σ_3	-20	-10	-30	-17	-31	-15	-28	-15	-28	-14
1:2	$\sigma_{1,2}$	-	-	-56	-30	-57	-28	-52	-26	-51	-25
	$\sigma_{1,3}$	-	-	-56	-29	-56	-29	-53	-27	-51	-24
	$\sigma_{2,3}$	-	-	-57	-29	-57	-29	-53	-26	-52	-25
1:3	$\sigma_{1,2,3}$	-	-	-83	-42	-84	-47	-77	-37	-74	-34
partial explicit DCE (1 per free σ -hole)											
1:1	σ_1	-16	-14	-25	-21	-26	-22	-22	-18	-21	-17
	σ_2	-16	-14	-23	-19	-24	-20	-22	-18	-21	-17
	σ_3	-16	-13	-25	-20	-26	-20	-24	-19	-23	-19

underestimates the enthalpies and Gibbs free energies in all the calculated cases, due to the missing crucial dispersion attraction as shown in Sec. 6.3. Both enthalpies and Gibbs free energies are higher (less negative) in the presence of the partial explicit DCE solvent for the 1:1 adducts, having the correct trend when compared to the experimental values, which is not the case of the calculations done in the COSMO implicit solvent. MBD-NL is on average 2.5 kcal/mol higher (less negative) than the pairwise D4 dispersion method.

6.5 Additional analysis of the pairwise vs many-body dispersion methods

Table 6.4: The differences of the PBE-D4 and PBE-MBD-NL enthalpies, Gibbs free energies and single point binding energies of the association (in kcal/mol) of the OPPh₃ molecules with TeMePh₂⁺ at 298.15 K computed with the intermediate basis set in the vacuum.

method	occ. σ -hole	D4 - MBD-NL		
		$\Delta\Delta H$	$\Delta\Delta G$	$\Delta\Delta E_0$
1:1	σ_1	-2.4	-2.4	-2.4
	σ_2	-2.7	-2.4	-2.7
	σ_3	-2.4	-2.5	-2.4
	avg	-2.5	-2.4	-2.5
1:2	$\sigma_{1,2}$	-4.5	-4.4	-4.5
	$\sigma_{1,3}$	-5.0	-4.6	-5.1
	$\sigma_{2,3}$	-5.7	-4.7	-5.8
	avg	-5.1	-4.6	-5.1
1:3	$\sigma_{1,2,3}$	-8.5	-7.2	-8.6
fixed OPPh ₃ molecules without TeMePh ₂ ⁺				
1:2	$\sigma_{1,2}$	-	-	-0.5
	$\sigma_{1,3}$	-	-	-0.5
	$\sigma_{2,3}$	-	-	-1.2
	avg	-	-	-0.7
1:3	$\sigma_{1,2,3}$	-	-	-2.0

in Tab. 6.5 the Te-O distances of the optimized structures using all the dispersion methods mentioned above. All the results with additional dispersion for the 1:1 adducts in the vacuum are within 0.06 Å, with slightly overestimated distances for the pure PBE. With the partial explicit DCE solvent for the 1:1 adducts, we see large differ-

The differences between pairwise and many-body dispersion methods in the gas phase can be explained by comparing the differences of the single point binding energies ΔE_0 . The average differences $\Delta\Delta E_0 = \Delta E_0^{\text{D4}} - \Delta E_0^{\text{MBD-NL}}$ are of -2.5 kcal/mol, -5.1 kcal/mol and -8.6 kcal/mol for the 1:1, 1:2 and 1:3 adducts respectively (Tab. 6.4). These differences increase almost linearly with the number of OPPh₃, persisting also for $\Delta\Delta H$ and $\Delta\Delta G$. The small deviations come mostly from the $\Delta\Delta E_0$ between OPPh₃ molecules, which we found to be of -0.7 kcal/mol and -2.0 kcal/mol for the 1:2 and 1:3 adducts respectively (Tab. 6.4). We also find these results to be independent of the structural parameters, since almost the same energy differences $\Delta\Delta E_0$ are obtained using both the geometries obtained at fixed D4 or MBD-NL level.

As a proof of the weak dependency of the geometries of the systems on the type of the dispersion methods, we show

ences between the pairwise and many-body methods for σ_1 , which can be a sign of an unconverged structural minima. The distances with the COSMO implicit and hybrid solvation methods are always slightly larger than those in the vacuum and partial explicit solution.

In order to check the convergence of the thermochemical quantities with respect to the basis set we performed calculations of the 1:1 adducts in the vacuum and in the partial explicit DCE using the light basis set [69]. The results are summarized in Appendix A3 in Tab. A3.1, showing 1-2 kcal/mol difference when compared to the intermediate basis set. Interestingly the enthalpies are always less negative with the intermediate basis set in the vacuum, but the differences do not have a fixed sign for the case of explicit partial DCE. This difference is small compared to the DFT precision, so we do not expect further significant changes using the more accurate tight basis set, which is too large for thermochemistry calculations of the studied systems.

The error caused by the choice of the generalized gradient approximation (GGA) PBE functional was tested by recalculating the thermochemistry quantities using the hybrid PBE0 functional in the light basis set. The results of the 1:1 adducts in the vacuum and in the partial explicit DCE in Appendix A3 in Tab. A3.2 indicate a slight deviation of about 1-2 kcal/mol between the PBE and PBE0 results, with PBE being always higher (less negative). These discrepancies are within the accuracy of the DFT method, so we considered them to be indistinguishable and the use of the cheaper GGA PBE functional is sufficient.

In Tabs. A3.3, A3.4 and A3.5 we report the single point binding energies of PBE in the light and intermediate basis sets and of the PBE0 in the light basis set used for the calculations of the thermochemistry quantities discussed above. As we analyzed in the pairwise versus many-body methods paragraph, all the main trends observed in ΔH and ΔG can be seen already in the ΔE_0 .

The last analysis is the comparison of both functionals in all three basis sets using all the dispersion methods mentioned above for fixed geometries of the systems obtained at the PBE-D4-COSMO level of theory for 1:1 adducts in the vacuum and in the partial explicit solvent (Tabs. A3.6-A3.11). The differences between PBE and PBE0 for all the basis sets are up to 2 kcal/mol, with PBE being mostly higher (less negative). Both

Table 6.5: The Te-O distances (in Å) of $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the intermediate basis set in the vacuum and the partial explicit solvation and with PBE-D4 in the COSMO and hybrid solvent.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL	D4
	occ. σ - hole	Te-O	Te-O	Te-O	Te-O	Te-O	Te-O
in vacuum							COSMO solvent
1:1	σ_1	2.68	2.59	2.63	2.61	2.61	2.76
	σ_2	2.65	2.58	2.61	2.60	2.60	2.76
	σ_3	2.65	2.60	2.64	2.58	2.58	2.82
partial explicit DCE (1 per free σ -hole)							hybrid solvent
1:1	σ_1	2.74	2.63	2.66	2.77	2.77	2.87
	σ_2	2.73	2.69	2.70	2.72	2.72	2.82
	σ_3	2.69	2.63	2.64	2.65	2.65	2.85

functionals show very similar dependencies on the basis set, with differences between the light and the tight basis set being up to 2 kcal/mol, with the larger basis set always leading to higher (less negative) binding energies. The differences between the various dispersion methods in all six cases follow the trends observed for the thermochemistry quantities.

6.6 Conclusions

Even though the computed association ΔH and ΔG values lie far from the experimental values, they clearly show the extent of the solvent’s screening and entropic effects. Explicit solvation indeed outlines a significant solvent-dependent entropic contribution to the Gibbs free energy of association that is made visible using the hybrid solvation approach. Nonetheless the O-Te interaction in $[(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$ has a light covalent character that outlines that such chalcogen molecular complexes are nothing but a new class of Lewis-type donor-acceptor complexes in which charge density transfer may

be tuned by adjusting the electron-withdrawing properties of substituents either at the telluronium cation or at the binding base. Last, it is found that pair-wise and many-body methods for the dispersion interactions have similar performances in reproducing the association energies of small to medium-large molecular complexes from the 1:1 to the 1:2 complexes of TeMePh_2^+ with OPPh_3 . However, for large molecular systems such as 1:3 assemblies, the MBD approach yields a lower over-stabilization of molecular complexes. This result suggests that the systematic comparison of pair-wise versus many-body methods should be carried out for large assemblies containing extended or numerous π systems, particularly when the closest match with experimental data is sought. An improvement of the theoretical study, in order to reduce the discrepancy between the measured and calculated enthalpies, could be achieved via the inclusion of the counterion of the salt and via a better description of the solvent, beyond the hybrid approach used in this work.

Summary and Outlook

In the last decade, the interest of computational Quantum Chemistry has progressively shifted towards the accurate *ab initio* description of large chemical compounds, and from the study of intramolecular interactions to that of weaker intermolecular interactions. This was stimulated by the development of new generations of hardware in high-performance computing facilities and of new electronic-structure methods. Yet, *ab initio* approaches to solve the Schrödinger equation for large systems of interest are still unfeasible, especially considering the accuracy required to describe some of their physical and chemical properties. For example, to study molecules in a solvent, the correct description of the interactions between the solute and the environment, dominated by electrostatics, polarization and dispersion contributions, is necessary for the modelling of the structural properties, the excitation energies, the reactivity or the optical properties of the solute. In order to decrease the limiting computational cost, the separation of the energy and time scales in the systems has been successfully exploited. In the so-called embedding methods, the system of interest is split into fragments each treated using computational methods of different accuracy and thus also of different computational costs. Yet, the most common embedding methods still suffer from difficulties in capturing the correct mutual polarization of the main fragment and the environment, and from approximations of the many-body dispersion effects through pairwise functions that only depend on the distances between the nuclei.

In this Thesis, we contributed to the embedding approaches by the development of

a new embedding method, *ie* the El-QDO method, in which the environment is modelled via quantum Drude oscillators, mimicking the response of real matter, while the main fragment is described at the fully electronic level through an *ab initio* method. In the El-QDO approach, we introduced a single many-body Hamiltonian for the mixed system of electrons and QDOs, proposed a correlated *ansatz* to represent the total variational wavefunction, and developed the quantum Monte Carlo algorithms for solving the problem, implementing them in the QMeCha quantum Monte Carlo package[12]. The applications of the El-QDO embedding method reported in Chap. 4 demonstrate the ability of the approach to capture the many-body mutual polarization and dispersion effects between the molecular electronic sub-system and the QDO environment. By comparing our results with the state-of-the-art DFT calculations and with QM:MM approaches we highlighted the limitations of the common pairwise approaches in describing the dispersion interaction contributions between the environment and the electronic subsystem. From these first results, it is clear that the El-QDO framework developed in this Thesis opens new possibilities for studying small to medium-sized molecular systems in a large explicit environment, built from point charges and QDOs. This will allow us to predict the effects of the environment on the structural and electronic properties of subsystems of interest that arise from electrostatics, mutual polarization and many-body dispersion interactions. All this can be achieved with a computational cost that is comparable to that of the subsystem in the vacuum, *ie* without the presence of the environment. Furthermore, the El-QDO approach can be implemented in other numerical methods such as Coupled Cluster and Configuration Interaction, with an appropriate generalization of the total wavefunction.

Despite the success of these first applications, modelling the short-range limit of the interactions between QDOs, and between QDOs and electrons in a systematic way still remains a challenge. One of the first steps in this direction has been taken in Chap. 5, where we have studied the potential energy surface of the Coulomb-coupled QDO dimer for all spatial separations, via quantum Monte Carlo methods, constructing a variational *ansatz* inspired by the physical properties of the model. We discussed the similarities and differences between the model and the H₂ molecule, and we explained the origin of the potential energy surface minima of the QDO dimer, which is due to the combination of the repulsion between the QDO centers and the drift of the drudonic particles towards the opposite Coulomb hole. Furthermore, within this Chapter, we

showed that even though the model is unable to form a covalent bond, it can be parametrized to reproduce the PES of covalently bonded atomic dimers, through a reparametrization of the Hamiltonian as a function of the distance. The generalization of the short-range regime in the QDO model, and of its interaction with the electrons, is also crucial for the construction of a more portable quantum Force-Field built out of interacting QDOs. Such a quantum Force-Field would naturally include the correct many-body dispersion effects, and at the same time would be able to at least effectively describe the short-range interactions between the environment and the subsystem of interest. Thus, the results presented in this Thesis will serve as a starting point in this broader direction that will inspire and guide further investigations.

Finally in Chap. 6 we showed the importance of the dispersion interactions and the effects of the environment on the association energies of biomolecular complexes. It is clear from the presented results, that a proper description of the many-body dispersion effects together with the inclusion of explicit solvent molecules are necessary to close the gap between the theoretical prediction and the experimental findings. Unfortunately, we were able to include only up to 3 explicit DCE solvent molecules at the full *ab initio* level with the currently available HPC facilities. Thus a reliable and more general embedding method with explicit treatment of electrostatics, polarization and dispersion, like the El-QDO method developed in this Thesis, will be of great interest in the fields of biochemistry, biophysics, biology and many others.

In conclusion, the developments, applications and analyses presented in this Thesis may become important tools for the investigation of a wider variety of solvated systems, expanding the landscape of the existing QM:MM and QM:QM methods, or of the implicit solvent models. Through the further generalization of the El-QDO embedding for all distances, the method will be computationally efficient and able to stand as an alternative for the QM:MM embedding approaches based on the classical point charges and classical polarizable dipoles for correlated methods of Quantum Chemistry, making a significant impact in the fields of drug development and in the simulations of biologically relevant systems.

Bibliography

- [1] a particle by particle size-extensive cut-off from private communication with M. Barborini.
- [2] Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Comm.*, 180(11):2175–2196, 2009.
- [3] Orlando Acevedo and William L. Jorgensen. Advances in quantum and molecular mechanical (qm/mm) simulations for organic and enzymatic reactions. *Acc. Chem. Res.*, 43(1):142–151, 2010.
- [4] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *J. Chem. Phys.*, 110(13):6158–6170, 1999.
- [5] Yasmine S. Al-Hamdani, Péter R. Nagy, Andrea Zen, Dennis Barton, Mihály Kállay, Jan Gerit Brandenburg, and Alexandre Tkatchenko. Interactions between large molecules pose a puzzle for reference quantum mechanical methods. *Nat. Commun.*, 12(1):3927, Jun 2021.
- [6] Ahmet Altun, Frank Neese, and Giovanni Bistoni. Local energy decomposition analysis of hydrogen-bonded dimers within a domain-based pair natural orbital coupled cluster study. *Beilstein J. Org. Chem.*, 14:919–929, 2018.
- [7] Cláudio Nahum Alves, José Rogério A. Silva, and Adrian E. Roitberg. Insights into the mechanism of oxidation of dihydroorotate to orotate catalysed by human

- class 2 dihydroorotate dehydrogenase: a qm/mm free energy study. *Phys. Chem. Chem. Phys.*, 17:17790–17796, 2015.
- [8] Alberto Ambrosetti, Dario Alfè, Robert A. Jr. DiStasio, and Alexandre Tkatchenko. Hard numbers for large molecules: Toward exact energetics for supramolecular systems. *J. Phys. Chem. Lett.*, 5(5):849–855, 2014.
- [9] Alberto Ambrosetti, Anthony M. Reilly, Robert A. DiStasio, and Alexandre Tkatchenko. Long-range correlation energy calculated from coupled atomic response functions. *J. Chem. Phys.*, 140(18):18A508, 2014.
- [10] Abdulgani Annaberdiyev, Guangming Wang, Cody A. Melton, M. Chandler Bennett, Luke Shulenburg, and Lubos Mitas. A new generation of effective core potentials from correlated calculations: 3d transition metal series. *J. Chem. Phys.*, 149(13):134108, 2018.
- [11] Lei Bao, Wen Liu, Yanwei Li, Xueyu Wang, Fei Xu, Zhongyue Yang, Yue Yue, Chenpeng Zuo, Qingzhu Zhang, and Wenxing Wang. Carcinogenic metabolic activation process of naphthalene by the cytochrome p450 enzyme 1b1: A computational study. *Chem. Res. Toxicol.*, 32(4):603–612, 2019.
- [12] Matteo Barborini. Quantum Mecha (**QMeCha**) package (private repository March 2023).
- [13] Giuseppe M. J. Barca, Colleen Bertoni, Laura Carrington, Dipayan Datta, Nuwan De Silva, J. Emiliano Deustua, Dmitri G. Fedorov, Jeffrey R. Gour, Anastasia O. Gunina, Emilie Guidez, Taylor Harville, Stephan Irle, Joe Ivanic, Karol Kowalski, Sarom S. Leang, Hui Li, Wei Li, Jesse J. Lutz, Ilias Magoulas, Joani Mato, Vladimir Mironov, Hiroya Nakata, Buu Q. Pham, Piotr Piecuch, David Poole, Spencer R. Pruitt, Alistair P. Rendell, Luke B. Roskop, Klaus Ruedenberg, Tosaporn Sattasathuchana, Michael W. Schmidt, Jun Shen, Lyudmila Slipchenko, Masha Sosonkina, Vaibhav Sundriyal, Ananta Tiwari, Jorge L. Galvez Vallejo, Bryce Westheimer, Marta Wloch, Peng Xu, Federico Zahariev, and Mark S. Gordon. Recent developments in the general atomic and molecular electronic structure system. *J. Chem. Phys.*, 152(15):154102, April 2020.
- [14] Federico Becca and Sandro Sorella. *Quantum Monte Carlo Approaches for Correlated Systems*. Cambridge University Press, 2017.

- [15] M. Chandler Bennett, Cody A. Melton, Abdulgani Annaberdiyev, Guangming Wang, Luke Shulenburger, and Lubos Mitas. A new generation of effective core potentials for correlated calculations. *J. Chem. Phys.*, 147(22):224106, 2017.
- [16] M. Chandler Bennett, Guangming Wang, Abdulgani Annaberdiyev, Cody A. Melton, Luke Shulenburger, and Lubos Mitas. A new generation of effective core potentials from correlated calculations: 2nd row elements. *J. Chem. Phys.*, 149(10):104108, 2018.
- [17] P. K. Biswas and V. Gogonea. A regularized and renormalized electrostatic coupling Hamiltonian for hybrid quantum-mechanical–molecular-mechanical calculations. *J. Chem. Phys.*, 123(16):164114, 10 2005.
- [18] Michel Bockstedte, Felix Schütz, Thomas Garratt, Viktor Ivády, and Adam Gali. Ab initio description of highly correlated states in defects for realizing quantum bits. *npj Quantum Mater.*, 3:31, 12 2018.
- [19] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Ann. Phys.*, 389(20):457–484, 1927.
- [20] S. F. Boys and Alfred Charles Egerton. Electronic wave functions - i. a general method of calculation for the stationary states of any molecular system. *Proc. R. Soc. A*, 200(1063):542–554, 1950.
- [21] Martin Brändle and Joachim Sauer. Acidity differences between inorganic solids induced by their framework structure. a combined quantum mechanics/molecular mechanics ab initio study on zeolites. *J. Am. Chem. Soc.*, 120(7):1556–1570, 1998.
- [22] M. Burkatzki, C. Filippi, and M. Dolg. Energy-consistent pseudopotentials for quantum monte carlo calculations. *J. Chem. Phys.*, 126(23):234105, 2007.
- [23] Rafel Cabot and Christopher A. Hunter. Molecular probes of solvation phenomena. *Chem. Soc. Rev.*, 41:3485–3492, 2012.
- [24] Matteo Calandra Buonauro and Sandro Sorella. Numerical study of the two-dimensional heisenberg model using a green function monte carlo technique with a fixed number of walkers. *Phys. Rev. B*, 57:11446–11456, May 1998.

- [25] Eike Caldeweyher, Sebastian Ehlert, Andreas Hansen, Hagen Neugebauer, Sebastian Spicher, Christoph Bannwarth, and Stefan Grimme. A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.*, 150(15):154122, 04 2019.
- [26] M. Casula, C. Attaccalite, and S. Sorella. Correlated geminal wave function for molecules: An efficient resonating valence bond approach. *J. Chem. Phys.*, 121:7110, 2004.
- [27] M. Casula, S. Moroni, S. Sorella, and C. Filippi. Size-consistent variational approaches to nonlocal pseudopotentials: standard and lattice regularized diffusion monte carlo. *J. Chem. Phys.*, 135:154113, 2010.
- [28] M. Casula and S. Sorella. Geminal wave function with jastrow correlation: A first application to atoms. *J. Chem. Phys.*, 119:6500, 2003.
- [29] Hanghui Chen, Alexander Hampel, Jonathan Karp, Frank Lechermann, and Andrew J. Millis. Dynamical mean field studies of infinite layer nickelates: Physics results and methodological implications. *Front. Phys.*, 10:835942, 2022.
- [30] Hyung Min Cho and William A. Lester, Jr. Explicit solvent model for quantum monte carlo. *J. Phys. Chem. Lett.*, 1(23):3376–3379, 2010.
- [31] F. S. Cipcigan, J. Crain, V. P. Sokhan, and G. J. Martyna. Electronic coarse graining: Predictive atomistic modeling of condensed matter. *Rev. Mod. Phys.*, 91:025003, May 2019.
- [32] Tim Clark. How deeply should we analyze non-covalent interactions? *J. Mol. Model.*, 29:66, 02 2023.
- [33] David J. C. Constable, Conchita Jimenez-Gonzalez, and Richard K. Henderson. Perspective on solvent use in the pharmaceutical industry. *Org. Process Res. Dev.*, 11(1):133–137, 2007.
- [34] David J. Coughtrie, Robin Giereth, Daniel Kats, Hans-Joachim Werner, and Andreas Köhn. Embedded multireference coupled cluster theory. *J. Chem. Theory Comput.*, 14(2):693–709, 2018.

- [35] C.J. Cramer. *Essentials of Computational Chemistry: Theories and Models*. Wiley, 2005.
- [36] Dieter Cremer. Møller-plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 1(4):509–530, 2011.
- [37] C. David Sherrill and Henry F. Schaefer. The configuration interaction method: Advances in highly correlated approaches. volume 34 of *Advances in Quantum Chemistry*, pages 143–269. Academic Press, 1999.
- [38] Matej Ditte, Matteo Barborini, Leonardo Medrano Sandonas, and Alexandre Tkatchenko. Molecules in environments: Toward systematic quantum embedding of electrons and drude oscillators. *Phys. Rev. Lett.*, 131:228001, Nov 2023.
- [39] P. Drude. Zur elektronentheorie der metalle. *Ann. Phys.*, 306(3):566–613, 1900.
- [40] Paul Drude. *The Theory of Optics*. Dover Publications, New York, 1959.
- [41] N. D. Drummond, M. D. Towler, and R. J. Needs. Jastrow correlation factor for atoms, molecules, and solids. *Phys. Rev. B*, 70:235119, 2004.
- [42] Drummond N. D. and Needs R. J. Variance-minimization scheme for optimizing jastrow factors. *Phys. Rev. B*, 72:85124, 2005.
- [43] Matúš Dubecký, Lubos Mitas, and Petr Jurečka. Noncovalent interactions by quantum monte carlo. *Chem. Rev.*, 116(9):5188–5215, 2016.
- [44] Pablo Echenique and J. L. Alonso. A mathematical and computational review of hartree–fock scf methods in quantum chemistry. *Mol. Phys.*, 105(23-24):3057–3098, 2007.
- [45] Thomas E. Exner, Andrea Frank, Ionut Onila, and Heiko M. Möller. Toward the quantum chemical calculation of nmr chemical shifts of proteins. 3. conformational sampling and explicit solvents model. *J. Chem. Theory Comput.*, 8(11):4818–4827, 2012.
- [46] Thomas E. Exner and Paul G. Mezey. The field-adapted adma approach: Introducing point charges. *J. Phys. Chem. A*, 108(19):4301–4309, 2004.

- [47] Dmitry V. Fedorov, Mainak Sadhukhan, Martin Stöhr, and Alexandre Tkatchenko. Quantum-mechanical relation between atomic dipole polarizability and the van der waals radius. *Phys. Rev. Lett.*, 121:183401, 2018.
- [48] Alan R. Fersht. Profile of martin karplus, michael levitt, and arieh warshel, 2013 nobel laureates in chemistry. *Proc. Natl. Acad. Sci.*, 110(49):19656–19657, 2013.
- [49] C. Filippi, X. Gonze, and C. J. Umrigar. *Generalized-gradient approximation to density-functional theory: comparison with exact results*. J. M. Seminario, Elsevier, Amsterdam, 1996.
- [50] C. Filippi and C. J. Umrigar. Correlated sampling in quantum monte carlo: A route to forces. *Phys. Rev. B Rapid Communications*, 61:16291, 2000.
- [51] V. Fock. Selfconsistent field mit austausch für natrium. *Z. Phys. A*, 62:795, 1930.
- [52] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33–83, 2001.
- [53] Joachim Friedrich, Michael Hanrath, and Michael Dolg. Fully automated implementation of the incremental scheme: Application to ccsd energies for hydrocarbons and transition metal compounds. *J. Chem. Phys.*, 126(15):154110, 2007.
- [54] Fye R. M. Study of trotter-like approximations. *J. Stat. Phys.*, 43:827–843, 1986.
- [55] Robert J. Gdanitz. An accurate interaction potential for neon dimer (ne₂). *Chem. Phys. Lett.*, 348(1):67–74, 2001.
- [56] Konstantinos Gkionis and James Platts. Qm/mm investigation into binding of square-planar platinum complexes to dna fragments. *J. Biol. Inorg. Chem.*, 14:1165–74, 08 2009.
- [57] Stefan Grimme. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.*, 18(32):9955–9964, 2012.
- [58] Stefan Grimme and Jean-Pierre Djukic. The crucial role of dispersion in the cohesion of nonbridged binuclear os → cr and os → w adducts. *Inorg. Chem.*, 49(6):2911–2919, 2010. PMID: 20143784.

- [59] Stefan Grimme and Jean-Pierre Djukic. Cation-cation “attraction”: When london dispersion attraction wins over coulomb repulsion. *Inorg. Chem.*, 50(6):2619–2628, 2011. PMID: 21309538.
- [60] Loïc Gros Lambert, Yann Cornaton, Matej Ditte, Emmanuel Aubert, Patrick Pale, Alexandre Tkatchenko, Jean-Pierre Djukic, and Victor Mamane. Affinity of telluronium chalcogen bond donors for lewis bases in solution: A critical experimental-theoretical joint study. *Chem. - Eur. J.*, n/a(n/a):e202302933.
- [61] Riccardo Gueschi, Habiburrahman Zulfikri, Csaba Daday, Franca Maria Floris, Claudio Amovilli, Benedetta Mennucci, and Claudia Filippi. Introducing qmc/mmpol: Quantum monte carlo in polarizable force fields for excited states. *J. Chem. Theory Comput.*, 12(4):1674–1683, 2016.
- [62] Szabolcs Góger, Almaz Khabibrakhmanov, Ornella Vaccarelli, Dmitry V. Fedorov, and Alexandre Tkatchenko. Optimized quantum drude oscillators for atomic and molecular response properties. *J. Phys. Chem. Lett.*, 14(27):6217–6223, 2023.
- [63] Hatice Gökcen, Eric Kratz, Thomas A. Darden, Jean-Philip Piquemal, and G. Andrés Cisneros. Qm/mm simulations with the gaussian electrostatic model: A density-based polarizable potential. *J. Phys. Chem. Lett.*, 9(11):3062–3067, 2018. PMID: 29775314.
- [64] Mustapha Hamdaoui, Marjolaine Ney, Vivien Sarda, Lydia Karmazin, Corinne Bailly, Nicolas Sieffert, Sebastian Dohm, Andreas Hansen, Stefan Grimme, and Jean-Pierre Djukic. Evidence of a donor-acceptor (ir-h)→sir3 interaction in a trapped ir(iii) silane catalytic intermediate. *Organometallics*, 35:2207–2223, 06 2016.
- [65] Andreas Hansen, Christoph Bannwarth, Stefan Grimme, Predrag Petrović, Christophe Werlé, and Jean-Pierre Djukic. The thermochemistry of london dispersion-driven transition metal reactions: Getting the ‘right answer for the right reason’. *ChemistryOpen*, 3(5):177–189, 2014.
- [66] A. Harju, B. Barbiellini, S. Siljamäki, R. M. Nieminen, and G. Ortiz. Stochastic gradient approximation: An efficient method to optimize many-body wave functions. *Phys. Rev. Lett.*, 79:1173–1177, 1997.

- [67] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods. *Mat. Proc. Cambridge Phil. Soc.*, 24:89, 1928.
- [68] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [69] V. Havu, V. Blum, P. Havu, and M. Scheffler. Efficient $\mathcal{O}(n)$ integration for all-electron electronic structure calculation using numeric basis functions. *J. Comput. Phys.*, 228(22):8367–8379, 2009.
- [70] W. Heitler and F. London. Wechselwirkung neutraler atome und homöopolare bindung nach der quantenmechanik. *Z. Phys.*, 44(6):455–472, 1927.
- [71] Robert Hellmann, Benjamin Jäger, and Eckard Bich. State-of-the-art ab initio potential energy curve for the xenon atom pair and related spectroscopic and thermophysical properties. *J. Chem. Phys.*, 147(3), 07 2017.
- [72] Jan Hermann, Robert A. DiStasio, and Alexandre Tkatchenko. First-principles models for van der waals interactions in molecules and materials: Concepts, theory, and applications. *Chem. Rev.*, 117(6):4714–4758, 2017.
- [73] Jan Hermann, Zeno Schätzle, and Frank Noé. Deep-neural-network solution of the electronic Schrödinger equation. *Nat. Chem.*, 12(10):891–897, 2020.
- [74] Jan Hermann, James Spencer, Kenny Choo, Antonio Mezzacapo, W. M. C. Foulkes, David Pfau, Giuseppe Carleo, and Frank Noé. Ab initio quantum chemistry with neural-network wavefunctions. *Nat. Rev. Chem.*, 7(10):692–709, Oct 2023.
- [75] Jan Hermann and Alexandre Tkatchenko. Density functional model for van der waals interactions: Unifying many-body atomic approaches with nonlocal functionals. *Phys. Rev. Lett.*, 124:146401, Apr 2020.
- [76] Joseelyne Hernández-Lima, Karla Ramírez-Gualito, Beatriz Quiroz-García, Ana Luisa Silva-Portillo, Ernesto Carrillo-Nava, and Fernando Cortés-Guzmán. How solvent determines the molecular reactive conformation and the selectivity: Solvation spheres and energy. *Front. Chem.*, 10, 2022.

- [77] Joseph O. Hirschfelder, Charles F. Curtiss, and R. Byron Bird. *The Molecular Theory of Gases and Liquids*. John Wiley & Sons, New York, 1954.
- [78] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.
- [79] Edward G. Hohenstein and C. David Sherrill. Wavefunction methods for non-covalent interactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2(2):304–326, 2012.
- [80] Chien-Jung Huang, Claudia Filippi, and C. J. Umrigar. Spin contamination in quantum monte carlo wave functions. *J. Chem. Phys.*, 108(21):8838–8847, 1998.
- [81] Ming Huang, Thakshila Dissanayake, Erich Kuechler, Brian K. Radak, Tai-Sung Lee, Timothy J. Giese, and Darrin M. York. A multidimensional b-spline correction for accurate modeling sugar puckering in qm/mm simulations. *J. Chem. Theory Comput.*, 13(9):3975–3984, 2017.
- [82] Yafei Huang, Joachim Reimann, Laila M.R. Singh, and Pia Ädelroth. Substrate binding and the catalytic reactions in cbb3-type oxidases: The lipid membrane modulates ligand binding. *Biophys. Acta - Bioenerg.*, 1797(6):724–731, 2010. 16th European Bioenergetics Conference 2010.
- [83] Pengfei Huo, Christopher Uyeda, Jason D. Goodpaster, Jonas C. Peters, and Thomas F. Miller, III. Breaking the correlation between energy costs and kinetic barriers in hydrogen evolution via a cobalt pyridine-diimine-dioxime catalyst. *ACS Catal.*, 6(9):6114–6123, 2016.
- [84] Wissam Iali, Predrag Petrović, Michel Pfeffer, Stefan Grimme, and Jean-Pierre Djukic. The inhibition of iridium-promoted water oxidation catalysis (woc) by cucurbit[n]urils. *Dalton Trans.*, 41:12233–12243, 2012.
- [85] Christoph R. Jacob and Lucas Visscher. Calculation of nuclear magnetic resonance shieldings using frozen-density embedding. *J. Chem. Phys.*, 125(19):194104, 2006.
- [86] A. Jones, F. Cipcigan, V. P. Sokhan, J. Crain, and G. J. Martyna. Electronically coarse-grained model for water. *Phys. Rev. Lett.*, 110:227801, May 2013.

- [87] Andrew Jones, Andrew Thompson, Jason Crain, Martin H. Müser, and Glenn J. Martyna. Norm-conserving diffusion monte carlo method and diagrammatic expansion of interacting drude oscillators: Application to solid xenon. *Phys. Rev. B*, 79:144119, Apr 2009.
- [88] Andrew P. Jones, Jason Crain, Vlad P. Sokhan, Troy W. Whitfield, and Glenn J. Martyna. Quantum drude oscillator model of atoms and molecules: Many-body polarization and dispersion interactions for atomistic simulation. *Phys. Rev. B*, 87:144103, Apr 2013.
- [89] A.P. Jones, J. Crain, F.S. Cipcigan, V.P. Sokhan, M. Modani, and G.J. Martyna. Electronically coarse-grained molecular dynamics using quantum drude oscillators. *Mol. Phys.*, 111(22-23):3465–3477, 2013.
- [90] Leighton O. Jones, Martín A. Mosquera, George C. Schatz, and Mark A. Ratner. Embedding methods for quantum chemistry: Applications from materials to life sciences. *J. Am. Chem. Soc.*, 142(7):3281–3295, 2020.
- [91] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, 87:897–923, Aug 2015.
- [92] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2):926–935, 07 1983.
- [93] William L. Jorgensen and Daniel L. Severance. Aromatic-aromatic interactions: free energy profiles for the benzene dimer in water, chloroform, and liquid benzene. *J. Am. Chem. Soc.*, 112(12):4768–4774, 1990.
- [94] Martin Kabeláč, Pavel Hobza, and Vladimír Špirko. The structure and vibrational dynamics of the pyrrole dimer. *Phys. Chem. Chem. Phys.*, 11:3885–3891, 2009.
- [95] Malvin H. Kalos and Paula A. Whitlock. *Quantum Monte Carlo*, chapter 8, pages 159–178. John Wiley & Sons, Ltd, 2008.
- [96] Konstantin P. Katin, Vladimir S. Prudkovskiy, and Mikhail M. Maslov. Chemisorption of hydrogen atoms and hydroxyl groups on stretched graphene: A coupled qm/qm study. *Phys. Lett. A*, 381(33):2686–2690, 2017.

- [97] Tosio Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.*, 10(2):151–177, 1957.
- [98] Rick A. Kendall, Jr. Dunning, Thom H., and Robert J. Harrison. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.*, 96(9):6796–6806, 05 1992.
- [99] P. R. C. Kent, R. J. Needs, and G. Rajagopal. Monte carlo energy and variance-minimization techniques for optimizing many-body wave functions. *Phys. Rev. B*, 59:12344–12351, May 1999.
- [100] Almaz Khabibrakhmanov, Dmitry V. Fedorov, and Alexandre Tkatchenko. Universal pairwise interatomic van der waals potentials based on quantum drude oscillators. *J. Chem. Theory Comput.*, 0(0):null, 0. PMID: 37875419.
- [101] Roopam Khare, Steven L. Mielke, Jeffrey T. Paci, Sulin Zhang, Roberto Ballarini, George C. Schatz, and Ted Belytschko. Coupled quantum mechanical/molecular mechanical modeling of the fracture of defective carbon nanotubes and graphene sheets. *Phys. Rev. B*, 75:075412, Feb 2007.
- [102] Jeongnim Kim, Andrew D Baczewski, Todd D Beaudet, Anouar Benali, M Chandler Bennett, Mark A Berrill, Nick S Blunt, Edgar Josué Landinez Borda, Michele Casula, David M Ceperley, Simone Chiesa, Bryan K Clark, Raymond C Clay, Kris T Delaney, Mark Dewing, Kenneth P Esler, Hongxia Hao, Olle Heinonen, Paul R C Kent, Jaron T Krogel, Ilkka Kylänpää, Ying Wai Li, M Graham Lopez, Ye Luo, Fionn D Malone, Richard M Martin, Amrita Mathuriya, Jeremy McMinis, Cody A Melton, Lubos Mitas, Miguel A Morales, Eric Neuscamman, William D Parker, Sergio D Pineda Flores, Nichols A Romero, Brenda M Rubenstein, Jacqueline A R Shea, Hyeondeok Shin, Luke Shulenburger, Andreas F Tillack, Joshua P Townsend, Norm M Tubman, Brett Van Der Goetz, Jordan E Vincent, D ChangMo Yang, Yubo Yang, Shuai Zhang, and Luning Zhao. Qmcpack: an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids. *J. Phys.: Condens. Matter*, 30(19):195901, apr 2018.
- [103] A. Klamt and G. Schüürmann. Cosmo: a new approach to dielectric screening

- in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2*, pages 799–805, 1993.
- [104] Andreas Klamt and Volker Jonas. Treatment of the outlying charge in continuum solvation models. *J. Chem. Phys.*, 105(22):9972–9981, 12 1996.
- [105] Gerald Knizia and Garnet Kin-Lic Chan. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.*, 109:186404, 2012.
- [106] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.
- [107] Ioan Kosztin, Byron Faber, and Klaus Schulten. Introduction to the diffusion monte carlo method. *Am. J. Phys.*, 64(5):633–644, 1996.
- [108] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti. Electronic structure calculations with dynamical mean-field theory. *Rev. Mod. Phys.*, 78:865–951, Aug 2006.
- [109] R Kubo. The fluctuation-dissipation theorem. 29(1):255–284, jan 1966.
- [110] Stephan Kupfer, Linda Zedler, Julien Guthmuller, Stefan Bode, Martin D. Hager, Ulrich S. Schubert, Jürgen Popp, Stefanie Gräfe, and Benjamin Dietzek. Self-healing mechanism of metallopolymers investigated by qm/mm simulations and raman spectroscopy. *Phys. Chem. Chem. Phys.*, 16:12422–12432, 2014.
- [111] Tran Nguyen Lan, Alexei A. Kananenka, and Dominika Zgid. Communication: Towards ab initio self-energy embedding theory in quantum chemistry. *J. Chem. Phys.*, 143(24):241102, 2015.
- [112] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael

- Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a python library for working with atoms. *J. Phys.: Condens. Matter*, 29(27):273002, 2017.
- [113] Bryan T. G. Lau, Gerald Knizia, and Timothy C. Berkelbach. Regional embedding enables high-level quantum chemistry for surface science. *J. Phys. Chem. Lett.*, 12(3):1104–1109, 2021.
- [114] Sebastian J. R. Lee, Matthew Welborn, Frederick R. Manby, and Thomas F. Miller, III. Projection-based wavefunction-in-dft embedding. *Acc. Chem. Res.*, 52(5):1359–1368, 2019.
- [115] Justin A. Lemkul, Jing Huang, Benoît Roux, and Alexander D. Jr. MacKerell. An empirical polarizable force field based on the classical drude oscillator model: Development history and recent applications. *Chem. Rev.*, 116(9):4983–5013, 2016.
- [116] Amanda Li, Hari S. Muddana, and Michael K. Gilson. Quantum mechanical calculation of noncovalent interactions: A large-scale evaluation of pmx, dft, and sapt approaches. *J. Chem. Theory Comput.*, 10(4):1563–1575, 2014. PMID: 24803867.
- [117] Dimitrios G. Liakos, Manuel Sparta, Manoj K. Kesharwani, Jan M. L. Martin, and Frank Neese. Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. *J. Chem. Theory Comput.*, 11(4):1525–1539, 2015. PMID: 26889511.
- [118] Florian Libisch, Chen Huang, and Emily A. Carter. Embedded correlated wavefunction schemes: Theory and applications. *Acc. Chem. Res.*, 47(9):2768–2775, 2014.
- [119] F. London. The general theory of molecular forces. *Trans. Faraday Soc.*, 33:8b–26, 1937.
- [120] Dmitry I. Lyakh, Monika Musiał, Victor F. Lotrich, and Rodney J. Bartlett. Multireference nature of chemistry: The coupled-cluster view. *Chem. Rev.*, 112(1):182–243, 2012. PMID: 22220988.

- [121] He Ma, Nan Sheng, Marco Govoni, and Giulia Galli. Quantum embedding theory for strongly correlated states in materials. *J. Chem. Theory Comput.*, 17(4):2116–2125, 2021.
- [122] Frederick R. Manby, Martina Stella, Jason D. Goodpaster, and Thomas F Miller, III. A simple, exact density-functional-theory embedding scheme. *J. Chem. Theory Comput.*, 8(8):2564–2568, 2012.
- [123] M. Marchi, S. Azadi, C. Casula, and S. Sorella. Resonating valence bond wave function with molecular orbitals: Application to first-row molecules. *J. Chem. Phys.*, 131:154116, 2009.
- [124] María del Carmen Marín, Luca De Vico, Sijia S. Dong, Laura Gagliardi, Donald G. Truhlar, and Massimo Olivucci. Assessment of mc-pdft excitation energies for a set of qm/mm models of rhodopsins. *J. Chem. Theory Comput.*, 15(3):1915–1923, 2019.
- [125] Nobuyuki Matubayasi. Solvation energetics of proteins and their aggregates analyzed by all-atom molecular dynamics simulations and the energy-representation theory of solvation. *Chem. Commun.*, 57:9968–9978, 2021.
- [126] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.
- [127] Michael P. Metz, Krzysztof Szalewicz, János Sarka, Roland Tóbiás, Attila G. Császár, and Edit Mátyus. Molecular dimers of methane clathrates: ab initio potential energy surfaces and variational vibrational states. *Phys. Chem. Chem. Phys.*, 21:13504–13525, 2019.
- [128] Milan R. Milovanović, Mélanie Boucher, Yann Cornaton, Snežana D. Zarić, Michel Pfeffer, and Jean-Pierre Djukic. The thermochemistry of alkyne insertion into a palladacycle outlines the solvation conundrum in dft. *Eur. J. Inorg. Chem.*, 2021(45):4690–4699, 2021.
- [129] Milan R. Milovanović, Quentin Dherbassy, Joanna Wencel-Delord, Françoise Colobert, Snežana D. Zarić, and Jean-Pierre Djukic. The affinity of some lewis

- bases for hexafluoroisopropanol as a reference lewis acid: An itc/dft study. *ChemPhysChem*, 21(18):2136–2142, 2020.
- [130] Milan R. Milovanović, Snežana D. Zarić, Yann Cornaton, and Jean-Pierre Djukic. Joint isotherm calorimetric titration-dft investigation of the demethoxyamination of fischer carbenes. *J. Organomet. Chem.*, 929:121582, 2020.
- [131] Ellen Miseo. 2023 review of spectroscopic instrumentation. *Spectroscopy*, 38(5):29–33, Sep 2023.
- [132] L. Mitas. Diffusion monte carlo. In M. P. Nightingale and C. J. Umrigar, editors, *Quantum Monte Carlo Methods in Physics and Chemistry*, volume C-525 of *NATO ASI Series, Series C, Math. & Phys. Sciences*. Kluwer Academic Publishers, Boston, 1999.
- [133] Luboš Mitáš, Eric L. Shirley, and David M. Ceperley. Nonlocal pseudopotentials and diffusion monte carlo. *J. Chem. Phys.*, 95(5):3467–3475, 1991.
- [134] Vladimir A. Nasluzov, Elena A. Ivanova, Alexei M. Shor, Georgi N. Vayssilov, Uwe Birkenheuer, and Notker Rösch. Elastic polarizable environment cluster embedding approach for covalent oxides and zeolites based on a density functional method. *J. Phys. Chem. B*, 107(10):2228–2241, 2003.
- [135] Vladimir A. Nasluzov, Vladimir V. Rivanenkov, Alexey B. Gordienko, Konstantin M. Neyman, Uwe Birkenheuer, and Notker Rösch. Cluster embedding in an elastic polarizable environment: Density functional study of pd atoms adsorbed at oxygen vacancies of mgo(001). *J. Chem. Phys.*, 115(17):8157–8171, 2001.
- [136] Frank Neese, Frank Wennmohs, Ute Becker, and Christoph Riplinger. The orca quantum chemistry program package. *J. Chem. Phys.*, 152(22):224108, 2020.
- [137] Johannes Neugebauer, Manuel J. Louwarse, Evert Jan Baerends, and Tomasz A. Wesolowski. The merits of the frozen-density embedding scheme to model solvatochromic shifts. *J. Chem. Phys.*, 122(9):094115, 2005.
- [138] E. A. Nowadnick, J. P. Ruf, H. Park, P. D. C. King, D. G. Schlom, K. M. Shen, and A. J. Millis. Quantifying electronic correlation strength in a complex oxide: A combined dmft and arpes study of lanio₃. *Phys. Rev. B*, 92:245109, Dec 2015.

- [139] Max Nusspickel and George H. Booth. Systematic improvability in quantum embedding for real materials. *Phys. Rev. X*, 12:011046, Mar 2022.
- [140] Shinsuke Osoegawa, Risako Miyoshi, Kouhei Watanabe, Yuu Hirose, Tomotsumi Fujisawa, Masahiko Ikeuchi, and Masashi Unno. Identification of the deprotonated pyrrole nitrogen of the bilin-based photoreceptor by raman spectroscopy with an advanced computational analysis. *J. Phys. Chem. B*, 123(15):3242–3247, 2019.
- [141] H. Padé. Sur la représentation approchée d’une fonction par des fractions rationnelles. *Ann. Sci. Éc. Norm. Supér.*, 9:3–93, 1892.
- [142] Konrad Patkowski, Garold Murdachaew, Cheng-Ming Fou, and Krzysztof Szalewicz. Accurate ab initio potential for argon dimer including highly repulsive region. *Mol. Phys.*, 103(15-16):2031–2045, 2005.
- [143] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, Oct 1996.
- [144] Francesco Petocchi, Fredrik Nilsson, Ferdi Aryasetiawan, and Philipp Werner. Screening from e_g states and antiferromagnetic correlations in $d^{(1,2,3)}$ perovskites: A gw + EDMFT investigation. *Phys. Rev. Res.*, 2:013191, Feb 2020.
- [145] Predrag V. Petrović, Stefan Grimme, Snežana D. Zarić, Michel Pfeffer, and Jean-Pierre Djukic. Experimental and theoretical investigations of the self-association of oxaliplatin. *Phys. Chem. Chem. Phys.*, 16:14688–14698, 2014.
- [146] Alastair J. A. Price, Kyle R. Bryenton, and Erin R. Johnson. Requirements for an accurate dispersion-corrected density functional. *J. Chem. Phys.*, 154(23):230902, 06 2021.
- [147] Reynolds P. J., Ceperley D. M., Alder J. B., and Lester W. A. Fixed-node quantum monte carlo for molecules. *J. Chem. Phys.*, 77:5593–5603, 1982.
- [148] Christoph Riplinger and Frank Neese. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.*, 138(3):034106, 01 2013.

- [149] Christoph Riplinger, Peter Pinski, Ute Becker, Edward F. Valeev, and Frank Neese. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.*, 144(2):024109, 01 2016.
- [150] Christoph Riplinger, Barbara Sandhoefer, Andreas Hansen, and Frank Neese. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.*, 139(13):134101, 10 2013.
- [151] Arturo Robertazzi and James A. Platts. Gas-phase dna oligonucleotide structures. a qm/mm and atoms in molecules study. *J. Phys. Chem. A*, 110(11):3992–4000, 2006.
- [152] M. Sadhukhan and Frederick R. Manby. Quantum mechanics of drude oscillators with full coulomb interaction. *Phys. Rev. B*, 94:115106, Sep 2016.
- [153] Wolfgang Schneider, Giovanni Bistoni, Manuel Sparta, Masaaki Saitow, Christoph Riplinger, Alexander Auer, and Frank Neese. Decomposition of intermolecular interaction energies within the local pair natural orbital coupled cluster framework. *J. Chem. Theory Comput.*, 12, 08 2016.
- [154] E. Schrödinger. Quantisierung als eigenwertproblem. *Ann. der Phys.*, 385(13):437–490, 1926.
- [155] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, 1926.
- [156] Tobias Schwabe, Stefan Grimme, and Jean-Pierre Djukic. Noncovalent metal-metal interactions: The crucial role of london dispersion in a bimetallic indenyl system. *J. Am. Chem. Soc.*, 131(40):14156–14157, 2009. PMID: 19769356.
- [157] Tobias Schäfer, Alejandro Gallo, Andreas Irmeler, Felix Hummel, and Andreas Grüneis. Surface science using coupled cluster theory via local wannier functions and in-rpa-embedding: The case of water on graphitic carbon nitride. *J. Chem. Phys.*, 155(24):244103, 2021.
- [158] Hans Martin Senn and Walter Thiel. Qm/mm methods for biomolecular systems. *Angew. Chem., Int. Ed. Engl.*, 48(7):1198–1229, 2009.

- [159] Michael Seth, Peter M. Margl, and Tom Ziegler. A density functional embedded cluster study of proposed active sites in heterogeneous ziegler-natta catalysts. *Macromolecules*, 35(20):7815–7829, 2002.
- [160] James Shee, Evan J. Arthur, Shiwei Zhang, David R. Reichman, and Richard A. Friesner. Singlet–triplet energy gaps of organic biradicals and polyacenes with auxiliary-field quantum monte carlo. *J. Chem. Theory Comput.*, 15(9):4924–4932, 2019.
- [161] Leonid Shirkov and Vladimir Sladek. Benchmark ccsd-sapt study of rare gas dimers with comparison to mp-sapt and dft-sapt. *J. Chem. Phys.*, 147(17):174103, 2017.
- [162] José Rogério A. Silva, Adrian E. Roitberg, and Cláudio Nahum Alves. A qm/mm free energy study of the oxidation mechanism of dihydroorotate dehydrogenase (class 1a) from *Lactococcus lactis*. *J. Phys. Chem. B*, 119(4):1468–1473, 2015.
- [163] Dominic A. Sirianni, Xiao Zhu, Doree F. Sitkoff, Daniel L. Cheney, and C. David Sherrill. The influence of a solvent environment on direct non-covalent interactions between two molecules: A symmetry-adapted perturbation theory study of polarization tuning of π - π interactions by water. *J. Chem. Phys.*, 156(19):194306, 2022.
- [164] J. C. Slater. The theory of complex spectra. *Phys. Rev.*, 34:1293–1322, Nov 1929.
- [165] J. C. Slater. Note on hartree’s method. *Phys. Rev.*, 35:210, 1930.
- [166] Vlad P. Sokhan, Andrew P. Jones, Flaviu S. Cipcigan, Jason Crain, and Glenn J. Martyna. Signature properties of water: Their molecular electronic origins. *Proc. Natl. Acad. Sci.*, 112(20):6341–6346, 2015.
- [167] Thomas Sommerfeld and Kenneth D. Jordan. Quantum drude oscillator model for describing the interaction of excess electrons with water clusters: An application to (h₂o)₁₃⁻. *J. Phys. Chem. A*, 109(50):11531–11538, 2005.
- [168] S. Sorella, M. Casula, and D. Rocca. Weak binding between two aromatic rings: feeling the van der Waals attraction by quantum monte carlo methods. *J. Chem Phys.*, 127:14105, 2007.

- [169] Sandro Sorella. Generalized lanczos algorithm for variational quantum monte carlo. *Phys. Rev. B*, 64:024512, Jun 2001.
- [170] Sandro Sorella. Wave function optimization in the variational monte carlo method. *Phys. Rev. B*, 71:241103, Jun 2005.
- [171] Sebastian Spicher and Stefan Grimme. Single-point hessian calculations for improved vibrational frequencies and rigid-rotor-harmonic-oscillator thermodynamics. *J. Chem. Theor. Comput.*, 17(3):1701–1714, 2021. PMID: 33554604.
- [172] Fernand Spiegelman, Nathalie Tarrat, Jérôme Cuny, Leo Dontot, Evgeny Posenitskiy, Carles Martí, Aude Simon, and Mathias Rapacioli. Density-functional tight-binding: basic concepts and applications to molecules and clusters. *Adv. Phys.: X*, 5(1):1710252, 2020. PMID: 33154977.
- [173] Anthony Stone. *The Theory of Intermolecular Forces*. Oxford University Press, 2016.
- [174] Jeffrey L. Streator. Solid-liquid-solid interfaces. In Mahmood Aliofkhazraei, editor, *Surface Energy*, chapter 3. IntechOpen, Rijeka, 2015.
- [175] Martin Stöhr, Leonardo Medrano Sandonas, and Alexandre Tkatchenko. Accurate many-body repulsive potentials for density-functional tight binding from deep tensor neural networks. *J. Phys. Chem. Lett.*, 11(16):6835–6843, 2020.
- [176] Martin Stöhr, Troy Van Voorhis, and Alexandre Tkatchenko. Theory and practice of modeling van der waals interactions in electronic-structure calculations. *Chem. Soc. Rev.*, 48:4118–4154, 2019.
- [177] Qiming Sun and Garnet Kin-Lic Chan. Quantum embedding theories. *Acc. Chem. Res.*, 49(12):2705–2712, 2016.
- [178] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler. Chemistry with adf. *J. Comput. Chem.*, 22(9):931–967, 2001.
- [179] Alexandre Tkatchenko, Robert A. DiStasio, Roberto Car, and Matthias Scheffler. Accurate and efficient method for many-body van der waals interactions. *Phys. Rev. Lett.*, 108:236402, Jun 2012.

- [180] Alexandre Tkatchenko and Matthias Scheffler. Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.*, 102:073005, Feb 2009.
- [181] Juan Torras, Francisco Rodríguez-Ropero, Oscar Bertran, and Carlos Alemán. Controlled isomerization of a light-driven molecular motor: A theoretical study. *J. Phys. Chem. C*, 113(9):3574–3580, 2009.
- [182] Julien Toulouse and C. J. Umrigar. Optimization of quantum monte carlo wave functions by energy minimization. *J. Chem Phys.*, 126(8):084102, 2007.
- [183] Gregory S. Tschumper. Multicentered integrated qm:qm methods for weakly bound clusters: An efficient and accurate 2-body:many-body treatment of hydrogen bonding and van der waals interactions. *Chem. Phys. Lett.*, 427(1):185–191, 2006.
- [184] C. J. Umrigar and Claudia Filippi. Energy and variance optimization of many-body wave functions. *Phys. Rev. Lett.*, 94:150201, 2005.
- [185] C. J. Umrigar, M. P. Nightingale, and K. J. Runge. A diffusion monte carlo algorithm with very small time-step errors. *J. Chem. Phys.*, 99(4):2865–2890, 1993.
- [186] C. J. Umrigar, Julien Toulouse, Claudia Filippi, S. Sorella, and R. G. Hennig. Alleviation of the fermion-sign problem by optimization of many-body wave functions. *Phys. Rev. Lett.*, 98:110201, 2007.
- [187] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins. Optimized trial wave functions for quantum monte carlo calculations. *Phys. Rev. Lett.*, 60:1719–1722, 1988.
- [188] Oliver Unke, Stefan Chmiela, Michael Gastegger, Kristof Schütt, Huziel E. Sauceda, and Klaus-Robert Müller. Spookynet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nat. Commun.*, 12:7273, 12 2021.
- [189] Oliver T. Unke, Stefan Chmiela, Huziel E. Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T. Schütt, Alexandre Tkatchenko, and Klaus-Robert Müller. Machine learning force fields. *Chem. Rev.*, 121(16):10142–10186, 2021.

- [190] Ornella Vaccarelli, Dmitry V. Fedorov, Martin Stöhr, and Alexandre Tkatchenko. Quantum-mechanical force balance between multipolar dispersion and pauli repulsion in atomic van der waals dimers. *Phys. Rev. Res.*, 3:033181, Aug 2021.
- [191] E. Van Lenthe and E. J. Baerends. Optimized slater-type basis sets for the elements 1-118. *J. Comput. Chem.*, 24(9):1142–1156, 2003.
- [192] Erik van Lenthe, Andreas Ehlers, and Evert-Jan Baerends. Geometry optimizations in the zero order regular approximation for relativistic effects. *J. Chem. Phys.*, 110(18):8943–8953, 05 1999.
- [193] Daniele Varsano, Matteo Barborini, and Leonardo Guidoni. Kohn-sham orbitals and potentials from quantum monte carlo molecular densities. *J. Chem. Phys.*, 140(5):054102, 2014.
- [194] Konstantinos D. Vogiatzis, Wim Klopper, and Joachim Friedrich. Non-covalent interactions of co₂ with functional groups of metal–organic frameworks from a ccsd(t) scheme applicable to large systems. *J. Chem. Theory Comput.*, 11(4):1574–1584, 2015.
- [195] Lucas K. Wagner, Michal Bajdich, and Lubos Mitas. Qwalk: A quantum monte carlo program for electronic structure. *J. Comput. Phys.*, 228(9):3390–3404, 2009.
- [196] Jonathan M. Waldrop, Bo Song, Konrad Patkowski, and Xiaopo Wang. Accurate ab initio potential for the krypton dimer and transport properties of the low-density krypton gas. *J. Chem. Phys.*, 142(20), 05 2015.
- [197] F. Wang and K. D. Jordan. A drude-model approach to dispersion interactions in dipole-bound anions. *J. Chem. Phys.*, 114(24):10717–10724, 2001.
- [198] Guangming Wang, Abdulgani Annaberdiyev, Cody A. Melton, M. Chandler Bennett, Luke Shulenburger, and Lubos Mitas. A new generation of effective core potentials from correlated calculations: 4s and 4p main group elements and first row additions. *J. Chem. Phys.*, 151(14):144110, 2019.
- [199] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *J. Comput. Chem.*, 25(9):1157–1174, 2004.

- [200] Arieh Warshel, Mitsunori Kato, and Andrei V. Pisliakov. Polarizable force fields: History, test cases, and prospects. *J. Chem. Theory Comput.*, 3(6):2034–2045, 2007. PMID: 26636199.
- [201] Florian Weigend. Hartree-fock exchange fitting basis sets for h to rn. *J. Comput. Chem.*, 29:167–75, 01 2008.
- [202] Paul G. Wenthold, Robert R. Squires, and W. C. Lineberger. Ultraviolet photoelectron spectroscopy of the o-, m-, and p-benzyne negative ions. electron affinities and singlet-triplet splittings for o-, m-, and p-benzyne. *J. Am. Chem. Soc.*, 120(21):5279–5290, 1998.
- [203] Tomasz A. Wesolowski, Sapana Shedge, and Xiuwen Zhou. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.*, 115(12):5891–5928, 2015.
- [204] Troy W. Whitfield and Glenn J. Martyna. A unified formalism for many-body polarization and dispersion: The quantum drude model applied to fluid xenon. *Chem. Phys. Lett.*, 424(4):409 – 413, 2006.
- [205] David E. Woon and Jr. Dunning, Thom H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.*, 98(2):1358–1371, 01 1993.
- [206] Chia-Nan Yeh, Sergei Isakov, Dominika Zgid, and Emanuel Gull. Electron correlations in the cubic paramagnetic perovskite $\text{Sr}(\text{V}, \text{Mn})\text{O}_3$: Results from fully self-consistent self-energy embedding calculations. *Phys. Rev. B*, 103:195149, May 2021.
- [207] Andrea Zen, Jan Gerit Brandenburg, Jiří Klimeš, Alexandre Tkatchenko, Dario Alfè, and Angelos Michaelides. Fast and accurate quantum monte carlo for molecular crystals. *Proc. Natl. Acad. Sci.*, 115(8):1724–1729, 2018.
- [208] Andrea Zen, Jan Gerit Brandenburg, Angelos Michaelides, and Dario Alfè. A new scheme for fixed node diffusion quantum monte carlo with pseudopotentials: Improving reproducibility and reducing the trial-wave-function bias. *J. Chem. Phys.*, 151(13):134105, 2019.

- [209] Andrea Zen, Sandro Sorella, Michael J. Gillan, Angelos Michaelides, and Dario Alfè. Boosting the accuracy and speed of quantum monte carlo: Size consistency and time step. *Phys. Rev. B*, 93:241118, 2016.
- [210] Jingjing Zheng, Xuefei Xu, and Donald Truhlar. Minimally augmented karlsruhe basis sets. *Theor. Chem. Acc.*, 128:295–305, 02 2011.
- [211] Tom Ziegler and Arvi Rauk. A theoretical study of the ethylene-metal bond in complexes between copper(1+), silver(1+), gold(1+), platinum(0) or platinum(2+) and ethylene, based on the hartree-fock-slater transition-state method. *Inorg. Chem.*, 18(6):1558–1565, 1979.
- [212] Paul M. Zimmerman, Martin Head-Gordon, and Alexis T. Bell. Selection and validation of charge and lennard-jones parameters for qm/mm simulations of hydrocarbon interactions with zeolites. *J. Chem. Theory Comput.*, 7(6):1695–1703, 2011.

Appendices

A1 Parameterization of El-FF, QDO and El-QDO models, and supporting results

Parameterization of the El-FF embedding, the QDOs used in the Thesis, the short-range repulsion for QDO and El-QDO models and supporting results for the El-QDO applications.

Table A1.1: Parameterization of the pairwise dispersion part of the El-FF embedding used in this work [93].

	ϵ [kcal/mol]	σ [Å]
C _{C₆H₆}	0.0700	3.5500
H _{C₆H₆}	0.0300	2.4200
O _{H₂O}	0.1521	3.1507
H _{H₂O}	0.0460	0.4000

Table A1.2: The parametrization (in Hartree atomic units) of the quantum Drude oscillators for Ne, Ar, Kr, Xe atoms and for the water molecule used in this work.

	q	ω	μ
Ne [88]	1.2494	1.2965	0.3491
Ar [88]	1.3314	0.7272	0.3020
Kr [88]	1.3741	0.6359	0.2796
Xe [88]	1.3570	0.5152	0.2541
H ₂ O [31]	1.1973	0.6287	0.3656

Table A1.3: The parameters of the external short-range repulsion from Eq. 2.66 with $N=2$ for the systems used in this work. C1 and C2 denote the energetically lowest and the second lowest geometry of the water dimer respectively, R and U denote refit and universal respectively and QDO=D and QDO=A denote the QDO representing the donor and the acceptor in the EI-QDO model respectively.

	a_1 [mHa]	a_2 [mHa]	b_1 [a_0^{-1}]	b_2 [a_0^{-1}]
QDO-QDO				
Ne	$1.518 \cdot 10^5$	$3.270 \cdot 10^0$	2.362	7.451
Ar	$4.030 \cdot 10^5$	$5.925 \cdot 10^{-2}$	1.891	33.380
Kr	$4.867 \cdot 10^5$	$3.376 \cdot 10^0$	1.753	5.957
Xe	$5.049 \cdot 10^5$	$3.466 \cdot 10^0$	1.567	5.020
H ₂ O (R, C1)	$1.851 \cdot 10^5$	$1.491 \cdot 10^2$	1.927	5.176
H ₂ O (R, C2)	$1.573 \cdot 10^5$	$1.496 \cdot 10^2$	1.956	5.064
H ₂ O [31] (U)	$6.133 \cdot 10^5$	$1.056 \cdot 10^3$	2.324	1.514
EI-QDO				
Ne	$6.085 \cdot 10^5$	$3.377 \cdot 10^0$	2.611	5.997
Ar	$3.112 \cdot 10^6$	$3.462 \cdot 10^0$	2.129	5.140
Kr	$5.788 \cdot 10^6$	$3.436 \cdot 10^0$	2.034	4.714
Xe	$2.041 \cdot 10^7$	$3.652 \cdot 10^0$	1.932	4.073
H ₂ O (C1, QDO=D)	$9.074 \cdot 10^5$	$1.503 \cdot 10^2$	2.042	52.388
H ₂ O (C2, QDO=D)	$1.153 \cdot 10^6$	$1.503 \cdot 10^8$	2.080	3.858
H ₂ O (C1, QDO=A)	$1.198 \cdot 10^6$	$1.492 \cdot 10^2$	2.136	5.163
H ₂ O (C2, QDO=A)	$5.887 \cdot 10^5$	$1.501 \cdot 10^2$	2.091	4.909

Table A1.4: The all-electron vs ECP calculations of the singlet-triplet gap (in kcal/mol) of ortho-benzyne with the aug-cc-pVTZ basis set at the UPBE0 level of theory. The differences are 2.87 kcal/mol and 2.86 kcal/mol in the vacuum and in the 4W cage respectively, thus we have chosen the ECP error for the DMC calculation to be 2.85 kcal/mol.

all-electron	
vacuum	28.12
4w cage	28.35
ECP	
vacuum	25.25
4w cage	25.49

A2 Adducts of TeMePh_2^+ with OPPh_3

Figures containing the rest of the adducts of TeMePh_2^+ with OPPh_3 studied in the vacuum and the partial explicit solvent.

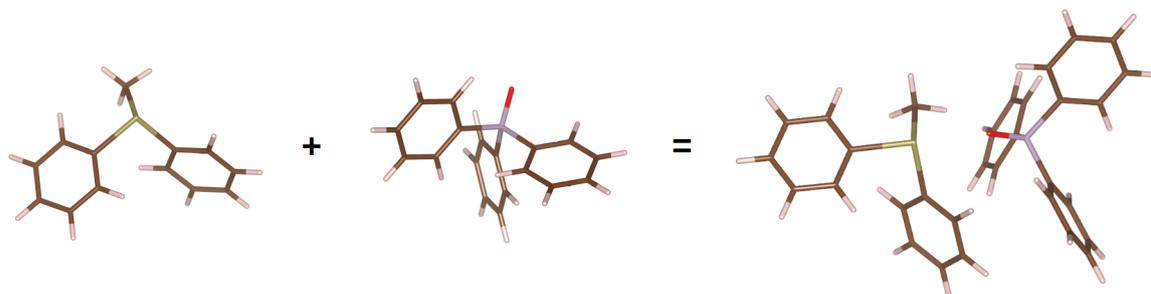


Figure A2.1: 1:1 adduct for σ_1 : $\text{TeMePh}_2^+ + \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$

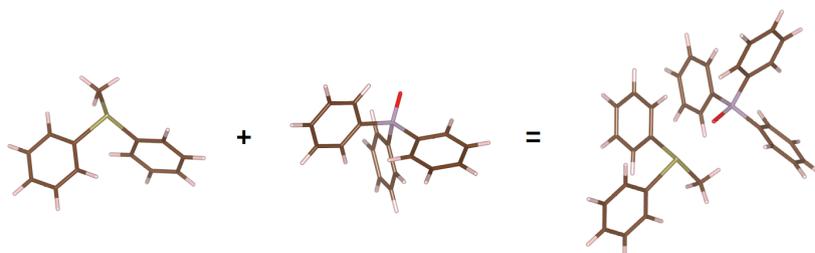


Figure A2.2: 1:1 adduct for σ_2 : $\text{TeMePh}_2^+ + \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$

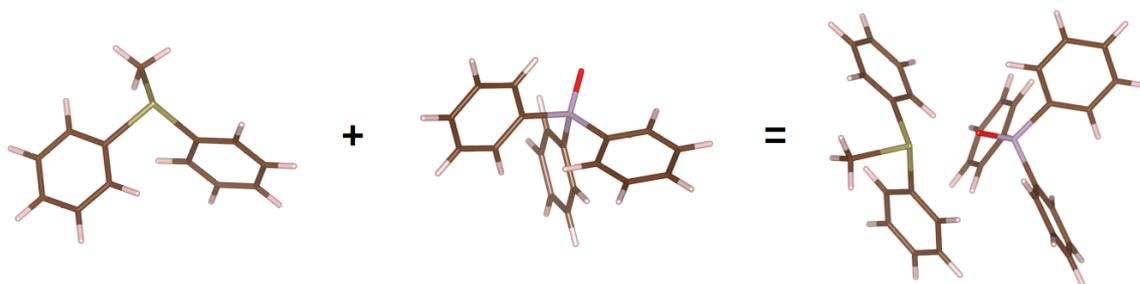


Figure A2.3: 1:1 adduct for σ_3 : $\text{TeMePh}_2^+ + \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})\cdot\text{TeMePh}_2]^+$

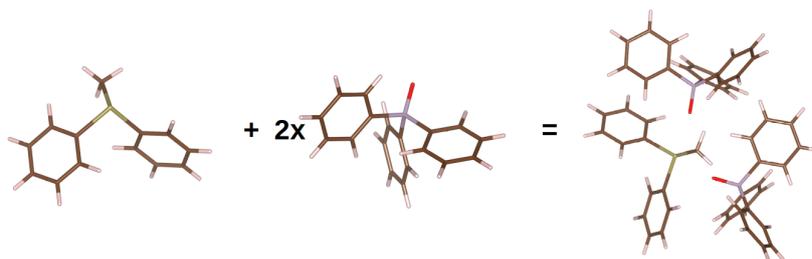


Figure A2.4: 1:2 adduct for $\sigma_{1,2}$: $\text{TeMePh}_2^+ + 2 \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})_2 \cdot \text{TeMePh}_2]^+$



Figure A2.5: 1:2 adduct for $\sigma_{1,3}$: $\text{TeMePh}_2^+ + 2 \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})_2 \cdot \text{TeMePh}_2]^+$

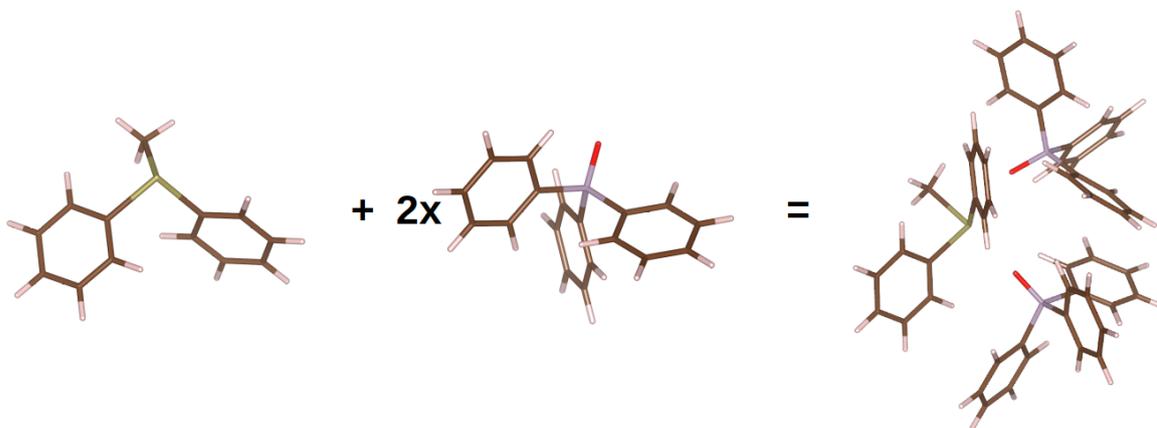


Figure A2.6: 1:2 adduct for $\sigma_{2,3}$: $\text{TeMePh}_2^+ + 2 \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})_2 \cdot \text{TeMePh}_2]^+$

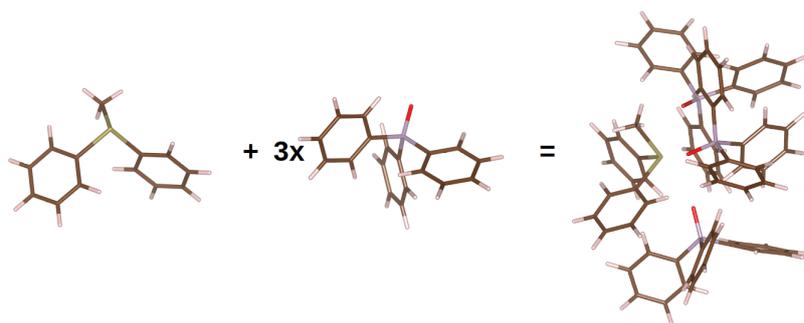


Figure A2.7: 1:3 adduct for $\sigma_{1,2,3}$: $\text{TeMePh}_2^+ + 3 \text{OPPh}_3 \rightarrow [(\text{Ph}_3\text{PO})_3 \cdot \text{TeMePh}_2]^+$

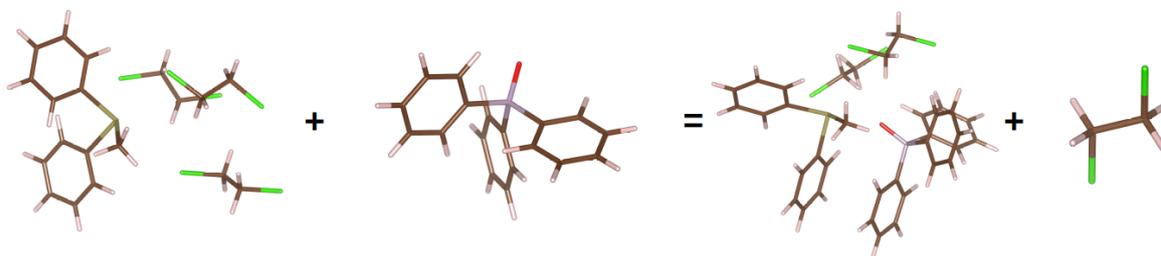


Figure A2.8: 1:1 adduct for σ_2 with 1 explicit DCE molecule per free σ -hole: $(\text{C}_2\text{H}_4\text{Cl}_2)_3 \cdot \text{TeMePh}_2^+ + \text{OPPh}_3 \rightarrow [(\text{C}_2\text{H}_4\text{Cl}_2)_2(\text{Ph}_3\text{PO}) \cdot \text{TeMePh}_2]^+ + \text{C}_2\text{H}_4\text{Cl}_2$

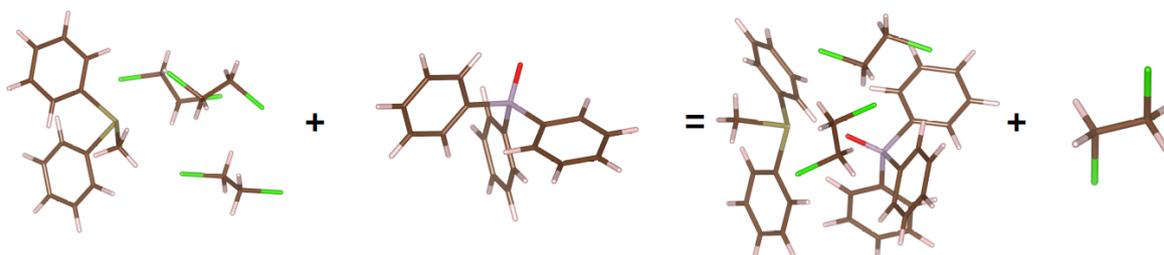


Figure A2.9: 1:1 adduct for σ_3 with 1 explicit DCE molecule per free σ -hole: $(\text{C}_2\text{H}_4\text{Cl}_2)_3 \cdot \text{TeMePh}_2^+ + \text{OPPh}_3 \rightarrow [(\text{C}_2\text{H}_4\text{Cl}_2)_2(\text{Ph}_3\text{PO}) \cdot \text{TeMePh}_2]^+ + \text{C}_2\text{H}_4\text{Cl}_2$

A3 Additional results of TeMePh_2^+

Addition results of the computational study of TeMePh_2^+ interacting with OPPh_3 in DCE solvent.

Table A3.1: The enthalpies and Gibbs free energies of association (in kcal/mol) of the OPPh₃ molecules with TeMePh₂⁺ at 298.15 K computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the light basis set in the vacuum and the partial explicit solvation.

method		PBE		D4		TS		MBD@ rsSCS		MBD-NL	
	occ. σ - hole	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG	ΔH	ΔG
in vacuum											
1:1	σ_1	-21	-11	-31	-18	-30	-18	-29	-16	-28	-16
	σ_2	-22	-10	-30	-16	-29	-17	-28	-14	-27	-14
	σ_3	-20	-9	-31	-17	-31	-17	-29	-15	-28	-14
1:2	$\sigma_{1,2}$	-	-	-56	-29	-57	-31	-53	-27	-52	-26
	$\sigma_{1,3}$	-	-	-57	-30	-57	-31	-54	-28	-52	-26
	$\sigma_{2,3}$	-	-	-58	-30	-58	-32	-54	-28	-52	-26
1:3	$\sigma_{1,2,3}$	-	-	-84	-42	-87	-47	-79	-39	-76	-36
partial explicit DCE (1 per free σ -hole)											
1:1	σ_1	-17	-9	-23	-20	-24	-25	-22	-19	-22	-18
	σ_2	-16	-9	-23	-19	-24	-20	-22	-18	-22	-18
	σ_3	-14	-8	-25	-21	-27	-22	-24	-20	-23	-19

Table A3.2: The enthalpies and Gibbs free energies of association (in kcal/mol) of OPPh_3 molecules with TeMePh_2^+ at 298.15 K computed at the PBE0, PBE0-D4, PBE0-TS, PBE0-MBD@rsSCS and PBE0-MBD-NL level of theory with the light basis set in the vacuum and a partial explicit solvation.

method		PBE0		D4		TS		MBD@rsSCS		MBD-NL	
	occ. σ -hole	ΔH	ΔG								
in vacuum											
1:1	σ_1	-22	-10	-32	-19	-31	-19	-30	-17	-30	-17
	σ_2	-23	-11	-31	-17	-30	-18	-28	-15	-28	-14
	σ_3	-21	-10	-32	-18	-32	-18	-30	-16	-29	-15
partial explicit DCE (1 per free σ -hole)											
1:1	σ_1	-18	-9	-25	-23	-25	-22	-24	-20	-23	-20
	σ_2	-18	-9	-25	-21	-26	-22	-24	-20	-23	-20
	σ_3	-16	-8	-27	-22	-28	-23	-26	-21	-25	-21

Table A3.3: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the light basis set in the vacuum and the partial explicit solvation.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-23	-32	-32	-30	-30
	σ_2	-24	-31	-31	-29	-29
	σ_3	-22	-32	-32	-30	-30
1:2	$\sigma_{1,2}$	-	-59	-60	-56	-55
	$\sigma_{1,3}$	-	-60	-60	-57	-55
	$\sigma_{2,3}$	-	-61	-61	-57	-55
1:3	$\sigma_{1,2,3}$	-	-89	-92	-83	-80
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-17	-23	-24	-23	-22
	σ_2	-17	-24	-25	-23	-22
	σ_3	-14	-25	-27	-24	-23

Table A3.4: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the intermediate basis set in the vacuum and the partial explicit solvation.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-22	-32	-31	-30	-29
	σ_2	-23	-31	-30	-28	-28
	σ_3	-21	-32	-32	-30	-29
1:2	$\sigma_{1,2}$	-	-58	-60	-55	-54
	$\sigma_{1,3}$	-	-59	-59	-56	-54
	$\sigma_{2,3}$	-	-60	-60	-56	-54
1:3	$\sigma_{1,2,3}$	-	-88	-89	-81	-79
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-16	-25	-26	-23	-22
	σ_2	-16	-23	-24	-22	-22
	σ_3	-16	-25	-26	-24	-23

Table A3.5: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE0, PBE0-D4, PBE0-TS, PBE0-MBD@rsSCS and PBE0-MBD-NL level of theory with the light basis set in the vacuum and the partial explicit solvation.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-23	-34	-32	-31	-31
	σ_2	-24	-32	-32	-30	-29
	σ_3	-22	-33	-33	-31	-30
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-18	-25	-26	-25	-24
	σ_2	-18	-25	-26	-24	-23
	σ_3	-16	-27	-28	-26	-25

Table A3.6: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the light basis set in the vacuum and the partial explicit solvation calculating at the geometries obtained at the PBE-D4-COSMO level of theory.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-21	-32	-32	-30	-30
	σ_2	-20	-31	-30	-29	-28
	σ_3	-21	-32	-31	-30	-29
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-14	-24	-25	-23	-23
	σ_2	-15	-24	-25	-23	-22
	σ_3	-14	-25	-27	-24	-23

Table A3.7: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the intermediate basis set in the vacuum and the partial explicit solvation calculating at the geometries obtained at the PBE-D4-COSMO level of theory.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-20	-32	-31	-30	-29
	σ_2	-20	-31	-30	-28	-28
	σ_3	-20	-31	-31	-29	-29
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-14	-24	-24	-23	-22
	σ_2	-15	-24	-25	-23	-22
	σ_3	-14	-25	-26	-24	-23

Table A3.8: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE, PBE-D4, PBE-TS, PBE-MBD@rsSCS and PBE-MBD-NL level of theory with the tight basis set in the vacuum and the partial explicit solvation calculating at the geometries obtained at the PBE-D4-COSMO level of theory.

method		PBE	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-20	-31	-30	-29	-28
	σ_2	-19	-30	-29	-28	-27
	σ_3	-19	-30	-30	-28	-28
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-13	-23	-23	-22	-21
	σ_2	-14	-23	-24	-22	-21
	σ_3	-13	-24	-25	-23	-22

Table A3.9: The single point binding energies of the association (in kcal/mol) of the OPPh_3 molecules with TeMePh_2^+ computed at the PBE0, PBE0-D4, PBE0-TS, PBE0-MBD@rsSCS and PBE0-MBD-NL level of theory with the light basis set in the vacuum and the partial explicit solvation calculating at the geometries obtained at the PBE-D4-COSMO level of theory.

method		PBE0	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-22	-33	-33	-31	-31
	σ_2	-22	-32	-31	-30	-29
	σ_3	-22	-33	-32	-31	-30
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-16	-26	-26	-25	-24
	σ_2	-16	-25	-26	-24	-23
	σ_3	-16	-27	-28	-25	-25

Table A3.10: The single point binding energies of the association (in kcal/mol) of the OPh₃ molecules with TeMePh₂⁺ computed at the PBE0, PBE0-D4, PBE0-TS, PBE0-MBD@rsSCS and PBE0-MBD-NL level of theory with the intermediate basis set in the vacuum and the partial explicit solvation calculating at the geometries obtained at the PBE-D4-COSMO level of theory.

method		PBE0	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-22	-33	-32	-31	-31
	σ_2	-21	-32	-31	-29	-29
	σ_3	-22	-32	-32	-30	-30
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-16	-25	-26	-25	-24
	σ_2	-16	-25	-26	-24	-23
	σ_3	-16	-27	-28	-25	-25

Table A3.11: The single point binding energies of the association (in kcal/mol) of the OPh₃ molecules with TeMePh₂⁺ computed at the PBE0, PBE0-D4, PBE0-TS, PBE0-MBD@rsSCS and PBE0-MBD-NL level of theory with the tight basis set in the vacuum and the partial explicit solvation calculating at the geometries obtained at the PBE-D4-COSMO level of theory.

method		PBE0	D4	TS	MBD@ rsSCS	MBD-NL
	occ. σ - hole	ΔE_0	ΔE_0	ΔE_0	ΔE_0	ΔE_0
in vacuum						
1:1	σ_1	-21	-32	-31	-30	-28
	σ_2	-20	-30	-30	-28	-27
	σ_3	-20	-31	-30	-29	-28
partial explicit DCE (1 per free σ -hole)						
1:1	σ_1	-14	-24	-24	-23	-21
	σ_2	-15	-24	-24	-22	-21
	σ_3	-14	-25	-26	-24	-22

