



UNIVERSITÀ DELL'AQUILA

DIPARTIMENTO DI INGEGNERIA E SCIENZE DELL'INFORMAZIONE E
MATEMATICA

"Mathematics and Models" PhD course

XXXIII cycle

Development of methods for polarizable Molecular Dynamics and electronic structure modelling

SSD: CHIM/02

Candidate

Francesco Cappelluti

Coordinator

Prof. Davide Gabrielli

Tutor

Prof. Leonardo Guidoni

AY 2019-2020

Contents

Abstract	1
1 Introduction	1
2 Methods	5
2.1 Electronic structure modelling	5
2.1.1 Density Functional Theory	7
2.1.2 Kohn-Sham method	7
2.1.2.1 DFT functionals	9
2.1.2.1.1 LDA	9
2.1.2.1.2 GGA	10
2.1.2.1.3 Meta-GGA	10
2.1.2.1.4 Hyper-GGA	10
2.1.2.1.5 Double-hybrid	10
2.1.2.1.6 Summary	11
2.1.3 The problem of static correlation	11
2.2 Vibrational Properties	13
2.2.1 Normal Mode Analysis	15
2.3 Extended Broken Symmetry	16
2.3.1 Estimation of J-coupling constant	17
2.3.2 Spin ladder	19
2.3.3 Ground state forces	20
2.4 Classical Molecular Dynamics	22
2.4.1 Force fields	25
2.4.1.1 Bonded terms	25
2.4.1.2 Non-bonded terms: van der Waals forces	27
2.4.1.2.1 Lennard-Jones potential	28
2.4.1.3 Non-bonded terms: electrostatic potential	29
2.4.2 Verlet lists	30

2.4.3	Ewald summation	31
2.4.3.1	Real space term	32
2.4.3.2	Fourier space term	33
2.4.3.3	Self-interaction correction	33
2.4.3.4	Derived methods	33
2.4.4	Calculation of partial charges	34
2.4.4.1	ESP and RESP charges	34
2.4.4.1.1	QM-MM/RESP charges	36
2.4.5	Polarizable force fields	36
2.4.5.1	Fluctuating charge model	37
2.4.5.2	Induced dipole model	38
2.4.5.3	Drude oscillator model	39
2.4.6	Ensembles	40
2.4.6.1	NVE ensemble	40
2.4.6.2	NVT Ensemble	40
2.4.6.3	NPT ensemble	41
2.4.6.4	μ VT ensemble	41
2.4.7	Trajectory analysis	42
2.4.7.1	Radial Distribution Function	42
2.4.7.2	Diffusion coefficient	43
3	Technical details	45
3.1	EBS	45
3.2	F-RESP	47
4	Results	49
4.1	Normal Mode Analysis within Extended Broken Symmetry	49
4.2	EBS simulation of Fe-S clusters	50
4.2.1	$\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ structure	52
4.2.2	$\text{Fe}_4\text{S}_4\text{Cl}_4$ structure	55
4.2.3	Vibrational frequencies	58
4.3	F-RESP	66
4.3.1	The polarization model	67
4.3.2	Fictitious electric field calculation	69
4.3.2.1	Ewald summation	69
4.3.2.2	Fennell potential	70

4.3.2.2.1	F-RESP virial contribution with the Fennell potential employed for the fictitious electric field calculation	71
4.3.2.3	Waibel potential	74
4.3.3	Damping for the electric field	74
4.3.4	Geometrical descriptors	76
4.3.5	Parameterization	76
4.4	F-RESP implementation and performances	81
4.5	F-RESP simulation of liquid water	84
4.5.1	Parameterization	84
4.5.2	Simulations	88
4.6	F-RESP simulation of ionic liquids	92
4.6.1	Parameterization	93
5	Conclusions	98
5.1	EBS	98
5.2	F-RESP	100
6	Supplementary material	103
6.1	EBS	103
6.1.1	Clebsch-Gordan coefficients calculation	104
6.1.2	Finite difference gradient test	104
6.1.3	EBS modes decomposition	105
6.1.4	Exchange coupling constants for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ and $\text{Fe}_4\text{S}_4\text{Cl}_4$	106
6.1.5	EBS vs BS normal modes frequencies comparison	108
6.1.6	VMARD analysis for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$	113
6.2	F-RESP	118
6.2.1	Electric field gradients using the Ewald summation	119
6.2.2	Electric field gradients using the Fennell potential	122
6.2.3	Electric field gradients using the Waibel potential	123
6.2.4	Damping function gradients	124
6.2.5	Geometrical descriptors gradients	124
6.3	Effects of the changes in the parameters of the damping function - Ewald summation	125
6.4	Effects of the changes in the parameters of the damping function - Fennell potential	126
7	Publication list	128

Acknowledgments	128
References	129

List of Figures

2.1	Example of superexchange interaction.	12
2.2	Theoretical IR spectra of water molecule obtained from NMA. . .	14
2.3	Graphic representation of the most commonly used parameters to model the internal geometric variations of a molecule.	26
2.4	The Lennard-Jones potential.	28
2.5	Graphic representation of the Verlet list method.	30
2.6	Ewald charge decomposition.	32
2.7	Example of molecular polarization for a water molecule.	37
3.1	Geometrical optimization procedure for a low-spin ground state using the Extended Broken Symmetry method.	46
4.1	The two investigated iron-sulfur clusters: $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ (left) and $\text{Fe}_4\text{S}_4\text{Cl}_4$ (right).	51
4.2	Example of spectrochemical series.	51
4.3	Crystal field splitting in tetrahedral and octahedral fields.	52
4.4	J-coupling constant and $Fe \cdots Fe$ distance vs spin description - $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$	54
4.5	J-coupling constant and $Fe \cdots Fe$ distance vs spin description - $\text{Fe}_4\text{S}_4\text{Cl}_4$	59
4.6	Comparison between Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$	60
4.7	Projection of the fictitious electric field along the bond axis.	68
4.8	Relative error with respect to the α and R_c parameters for TIP3P water.	72
4.9	Example of the two damping function forms: sinusoidal (left) and Gaussian (right).	75
4.10	The parameterization process.	79
4.12	Energy vs time in a NVE simulation.	84

4.13	Water HF/6-31G* QM/MM-RESP vs F-RESP atomic charges (with parameters obtained from charge matching) for oxygen (left) and hydrogen (right).	85
4.14	Water PBE-D3 vs F-RESP atomic forces (with parameters obtained from force matching) for oxygen (left) and hydrogen (right). . .	85
4.15	Comparison of the charge fluctuation for an oxygen atom for a F-RESP and an AIMD trajectory.	88
4.16	Oxygen-oxygen Radial Distribution Function comparison.	90
4.17	Changes in O-O RDF with respect to the damping (sinusoidal) function parameters - Fennell potential TIP3P charge matching. .	91
4.18	2-methoxyethylammonium (top), methylammonium (bottom left) and nitrate (bottom right) ions. For methylammonium and nitrate, the GAFF atom types are depicted.	92
4.19	Methylammonium nitrate HF/6-31G* QM/MM-RESP vs F-RESP atomic charges (with parameters obtained from charge matching) for the different GAFF atomic types.	93
4.20	Methylammonium nitrate PBE-D3 vs F-RESP atomic forces (with parameters obtained from force matching) for the different GAFF atomic types.	95
6.1	Changes in the O-O RDF with respect to the parameters of the damping (sinusoidal) function - Ewald summation TIP3P charge matching.	126

List of Tables

4.1	Geometrical comparison - $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$	53
4.2	Geometrical comparison - $\text{Fe}_4\text{S}_4\text{Cl}_4$	57
4.3	Extended Broken Symmetry VMARD analysis on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - M06/def2-TZVP.	63
4.4	Extended Broken Symmetry VMARD analysis on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B2PLYP/def2-TZVP.	64
4.5	Normal mode frequency comparison - $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ M06/def2-TZVP and B2PLYP/def2-TZVP.	65
4.6	F-RESP parameters for TIP3P water - charge matching.	86
4.7	F-RESP parameters for TIP3P water - force matching.	87
4.8	TIP3P water parameters.	87
4.9	Density and diffusion coefficient comparison.	89
6.1	Test of finite difference gradient.	105
6.2	Normal modes decomposition - $\text{Fe}_4\text{S}_4\text{Cl}_4$ OPBE.	106
6.3	J coupling constants - $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$	107
6.4	J coupling constants - $\text{Fe}_4\text{S}_4\text{Cl}_4$	107
6.5	Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B3LYP and BP.	108
6.6	Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - M06 and OPBE.	109
6.7	Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B2PLYP and SCS-MP2.	110
6.8	Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_4\text{S}_4\text{Cl}_4$ - B3LYP and BP.	111
6.9	Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_4\text{S}_4\text{Cl}_4$ - M06 and OPBE.	112

6.10	Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B3LYP.	114
6.11	Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - BP.	115
6.12	Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - OPBE.	116
6.13	Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - SCS-MP2.	117
6.14	Changes in the O-O RDF maxima and minima position and in the diffusion coefficient with respect to the parameters of the damping (sinusoidal) function - Ewald summation TIP3P charge matching.	126
6.15	Changes in the density and the diffusion coefficient with respect to the parameters of the damping (sinusoidal) function - Fennell potential TIP3P charge matching.	127

Abstract

The computational description of chemical systems is nowadays acknowledged of great importance for the widespread application of theoretical methods in microscopic structural modelling. The modelling of both simple and complex molecular systems is feasible by applying the first principles of quantum mechanics and such simulations may be accomplished with or without their evolution in time. These approaches have their respective advantages and disadvantages, depending on the application field.

In the present Thesis, two innovative methods for chemical modelling are presented, that is the Normal Mode Analysis within Extended Broken Symmetry (EBS) and the Fluctuating-RESP (F-RESP) method.

The Extended Broken Symmetry (EBS) approach is an electronic structure method that permits to correct the Density Functional Theory (DFT) (or, more generally mono-determinant based methods) when dealing with low spin open-shell molecular systems characterized by antiferromagnetic interactions. This method (relying on ORCA or CP2K program packages for *ab initio* calculations) is generalized with respect to the system size, meaning that it can deal with molecular species containing an arbitrary number of spin centres. The analytical expressions for the geometrical gradient and Hessian matrix have been derived, therefore permitting to perform geometrical optimizations and provide vibrational spectra.

The Fluctuating-RESP (F-RESP) approach is a framework for performing polarizable classical Molecular Dynamics simulations within the fluctuating charges formalism. The method (implemented in LAMMPS package) is able to perform, due to the remarkable parallelization and intrinsic theoretical feasibility, simulations within the limits of the most employed statistical ensembles.

The former method has been applied in my study on two iron-sulfur clusters, representing the most common structural patterns of this important class of compounds and for which high quality experimental data are available. The accuracy of the description that EBS was able to perform of these systems superseded the traditional BS-DFT approach both at the magnetic, structural and vibrational levels.

The latter method has been employed for the simulation of liquid water, a system of paramount importance whose knowledge is, despite of the apparent structural simplicity and huge research efforts made so far, still not thoroughly

complete. The water model obtained by applying the F-RESP approach on the pre-existing TIP3P model permitted to improve the accuracy of the description and derive physical observables that are closer to the experimentally measured ones.

Chapter 1

Introduction

"The fundamental laws necessary for the mathematical treatment of large parts of physics and the whole of chemistry are thus fully known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved." (P. A. M. Dirac, 1929)

Such quote of the Nobel laureate P. A. M. Dirac is today no more true: if the analytical solution of these equations still remains not possible, the use of numerical approaches is nowadays fully and deeply established[1, 2], the first attempts dating back to the Twenties of last century. In those years, in fact, Vladimir Fock and Douglas Hartree developed their own method, later known as the Hartree-Fock Method[3, 4, 5, 6, 7, 8] (henceforth reported in this Thesis as HF) able to approximately solve the Schrödinger equation:

$$\hat{H}\Psi(\vec{r}) = E\Psi(\vec{r}) \quad (1.1)$$

The Schrödinger equation (here displayed in its time-independent form) is the pillar of quantum mechanics, the branch of physics governing the world on atomic and molecular scale. It is well-known that this equation can be analytically solved only for hydrogen and hydrogen-like atoms and the H₂ molecule. The implementation of the Hartree-Fock method represents a real revolution, permitting to deal with the atomic and in particular the molecular world in an *ab initio* fashion (*i.e.* basing on first principles only, without introducing anything of empirical origin). Appearance of modern computers, starting from the Fifties, obviously enormously boosted the field of computational chemistry, which was until that time bounded within the narrow limits of handmade calculations.

An important innovation was introduced by Roothan and Hall[9, 10], who better "adapted" the Hartree-Fock method to the new world of computers by ex-

pressing the electronic wavefunctions as linear combinations of functions, therefore making easier to apply linear algebra to the calculations. In the same years, Boys[11] proposed to employ Gaussian functions in place of simple exponential ones (adopted first by Slater) for expressing atomic and molecular wavefunctions.

A few years later, Boys again first accomplished a Configuration Interaction calculation[12], namely a kind of technique overcoming the mean-field approximation of the Hartree-Fock method.

Configuration Interaction, or C.I. methods, is actually expected to provide much more accurate molecular properties.

Further, the discovery of an efficient way to calculate energy gradients as proposed by Pulay[13] definitively brought computational chemistry out of the pioneering era. The determination of an equilibrium geometry of a given structural model is the initial step of any investigation of this kind and is actually based on the energy gradient method. It is in fact adopted for geometry optimizations, that is the localization of a stable configuration on the Potential Energy Surface (PES) of a system. During the Seventies, the first computer codes with a broad and even commercial diffusion started to appear and the Gaussian package[14], which is the most used and known program for molecular computations, gained a particular success.

Almost in the same years, another revolution was broken out by Hohenberg, Kohn and Sham[15, 16], who gave life to the Density Functional Theory (DFT), first introduced in its orbital-free form by Thomas and Fermi[17, 18] in the field of solid state physics. This technique overcomes the limitations of Hartree-Fock approach in the description of the correlation of electronic motion by keeping computational cost comparable to the latter approach or anyway much lower than the already cited Configuration Interaction, the perturbative Møller-Plesset technique[19] or the more recent Coupled-Cluster[20, 21] method. For this reason, it is nowadays the most employed technique for the *ab initio* (even if the inclusion of DFT into this category is a bit debated) studies, having largely superseded the HF method (too inaccurate for the current needs) and the more accurate post-HF methods.

The importance reached by this field of Chemistry and Physics was testified more than twenty years ago by the Nobel Prize awarded

"to Walter Kohn for his development of the density-functional theory and to John Pople for his development of computational methods in quantum chemistry."
(Nobel prize press release, 10-13-1998)

In parallel with these *ab initio* (or nearly *ab initio*) methods, a completely different technique called classical Molecular Dynamics (MD) was developed. This approach is, at variance with the previously described ones, completely empiric. The electrons are ignored and atomic nuclei are described as classical particles interacting among them, using tailor-made potentials, derived by experiments or, more recently, accurate *ab initio* calculations. The use of these simple (with respect to quantum mechanics) interatomic potentials makes possible to study systems that are orders of magnitude larger than those accessible by *ab initio* approaches.

The very first computer simulation of this kind was performed by Metropolis in 1953[22], even though the simulation was accomplished for a very simplified system consisting of hard-spheres on a two-dimensional surface and the closely related Montecarlo Sampling method[23] was employed.

Just six years later, the same system was simulated again by Alder[24] in a real Molecular Dynamics computational experiment, solving the equations of the motion of the particles, and during the early Sixties, the first realistic simulations of matter were performed[25, 26].

These initial simulations were concerning systems relatable to the field of material science, although during the Seventies, Karplus, Levitt and Warshel started to apply the technique to biomolecular compounds, notably proteins[27, 28].

Their efforts were awarded in 2013 by the Nobel Prize having as motivation:

"for the development of multiscale models for complex chemical systems"

(Nobel prize press release, 10-09-2013)

In this way it was therefore fully recognized their contribution and indirectly the contribution of Molecular Dynamics to the study of systems having significant size.

Molecular Dynamics could enlarge its application field in 1985 when Car and Parrinello[29] developed a method for integrating DFT and MD much closely. Car-Parrinello Molecular Dynamics (CPMD) makes use of an extended Lagrangian approach and permitted to perform dynamical simulations at the high accuracy level delivered by Density Functional Theory in a fraction of the time required by the more traditional approaches (*e.g.* Born-Oppenheimer[30] Molecular Dynamics).

In the most recent years, other innovations such as GPU Computing[31] or Machine Learning[32] are making possible to further extend the capabilities of

Molecular Dynamics, both in the time (or size) scale and in accuracy.

In the current days, the accuracy of computer simulations has reached a remarkable level, permitting to reliably investigate systems whose experimental study would be too difficult, expensive or even dangerous (lets think about the high-throughput screening of compounds of medical interest[33] or the simulation of nuclear processes). For these reasons, the 2013 Nobel Committee, the same who awarded Karplus, Levitt and Warshel, could pronounce the historical quote:

"Today the computer is just as important a tool for chemists as the test tube."

In the current Thesis, I will describe the work performed during the three years of my PhD Course in " Mathematics and Models " at the University of L'Aquila. My research is mainly focused on two topics, that is the development of a polarizable force field (quoted as Fluctuating-Restrained Electrostatic Potential, F-RESP) for the classical Molecular Dynamics and the development of a method for dealing with systems affected by static correlation (quoted as Extended Broken Symmetry, EBS) within electronic structure calculations.

In section 2.1 and section 2.4, I will shortly describe the most important theoretical methods of computational chemistry, with particular consideration to the Density Functional Theory (DFT) and the Classical Molecular Dynamics. I am also going to outline the issues that prompted the development of the EBS and F-RESP approaches.

In section 2.3 and section 4.3, I will describe in detail the theoretical framework of these methods, while in chapter 3 I will take into consideration their technical implementation.

At last, in chapter 4, I will illustrate the encouraging results obtained from the application of these methods on some test systems. These results have revealed themselves suitable for publication in valuable peer-reviewed scientific journals, demonstrating their novelty and importance for the scientific community. In chapter 5, I will summarize the results accomplished so far and outline the possible further advancements in my research.

Chapter 2

Methods

2.1 Electronic structure modelling

In quantum mechanics, any system is represented by a wavefunction $\Psi(\vec{r}, t)$, function of the position of each particle and of time, containing all the information concerning the system under consideration. The most diffused (but not just the only) interpretation of quantum mechanics, namely the Copenhagen interpretation elaborated by Niels Bohr and Werner Heisenberg[34], affirms that the square of the normalized wavefunction represents the probability density of finding the system in the configuration given by $\{\vec{r}, t\}$. In several practical applications however, at least in the typical chemical domain, the wavefunction is considered to be stationary, namely time-independent. The information contained in the wavefunction can be extracted under the action of specific operators (for example, in the well-known Schrödinger equation 1.1, the Hamiltonian operator \hat{H} permits to obtain the total energy of the system represented by the given wavefunction).

Such as in classical mechanics, the total energy can be splitted into two independent components, that is to say, kinetic energy and potential energy. Therefore, the Hamiltonian operator is written as sum of the operators T and V:

$$\hat{H} = \hat{T} + \hat{V} \quad (2.1)$$

where each operator can be further decomposed in terms referred to nuclei, electrons and their respective interactions:

$$\hat{T} = \hat{T}_N(\vec{R}) + \hat{T}_e(\vec{r}) \quad (2.2)$$

$$\hat{T}_N(\vec{R}) = \sum_{I=1}^N \left(-\frac{\hbar^2}{2M_I} \nabla_{\vec{R}_I}^2 \right) \quad (2.3)$$

$$\hat{T}_e(\vec{r}) = \sum_{i=1}^n \left(-\frac{\hbar^2}{2m_{e^-}} \nabla_{\vec{r}_i}^2 \right) \quad (2.4)$$

$$\hat{V} = \hat{V}_{NN} + \hat{V}_{Ne} + \hat{V}_{ee} \quad (2.5)$$

$$\hat{V}_{NN}(\vec{R}) = \sum_{I < J}^N \frac{Z_I Z_J e^{-2}}{|\vec{R}_I - \vec{R}_J|} \quad (2.6)$$

$$\hat{V}_{Ne}(\vec{R}, \vec{r}) = - \sum_{I,i}^{N,n} \frac{Z_I e^{-2}}{|\vec{R}_I - \vec{r}_i|} \quad (2.7)$$

$$\hat{V}_{ee}(\vec{r}) = \sum_{i < j}^n \frac{e^{-2}}{|\vec{r}_i - \vec{r}_j|} \quad (2.8)$$

Here N , n , \vec{R} , \vec{r} , M_I and m_{e^-} represent the number, the coordinates and the masses of nuclei and electrons, respectively and e^- represents the electric charge.

Due to the large mass ratio between atomic nuclei and electrons (≥ 1836), a very useful approximation for the energy calculation is to decouple their motion, that is to consider the nuclear positions fixed with respect to the motion of the electrons. In other words, that means that electrons instantly react to a perturbation in nuclear positions. The parametric dependence of electron energy on the nuclear position, together with the assumption that thermal energy is much lower than the electronic excitation energy, compose the so-called Born-Oppenheimer Approximation[30]. In this approximation, the nuclei are considered to move on a Potential Energy Surface which is solution of the electronic Schrödinger equation.

Unfortunately, even with Born-Oppenheimer approximation and as already written in chapter 1, an analytical solution to the equation is possible only for extremely simple systems lacking of any practical interest. Other approximations are therefore necessary for approaching larger systems and their features often depend on the purpose of the calculation. One of the most successful approaches both for small and large systems is the Density Functional Theory (DFT).

2.1.1 Density Functional Theory

The DFT is based on the theorem of Hohenberg and Kohn [15], according to which the ground state (the state lying at the lowest energy value for a given configuration) can be fully described by its electron density $\rho(\vec{r})$. In other words, it exists a direct correspondence between the energy of the system and the electron density of its fundamental state. Extensions of this theorem for the excited states have been also proposed[35], even though they are still debated[36].

The intuition of Hohenberg and Kohn [15] is based on simple considerations about $\rho(\vec{r})$ done by E. B. Wilson[37]:

- The integral of $\rho(\vec{r})$ over the entire space defines the number of electrons.
- The cusps of $\rho(\vec{r})$ correspond to the nuclear positions.
- The heights of the cusps define the charge of the nuclei.

A system composed of n electrons is described by a wavefunction containing $4n$ variables (three spatial and one spin variable for each particle). As one can easily imagine, the complexity of the function is quickly increasing with the size of the system, becoming extremely convoluted for particularly extended systems.

On the contrary, the electron density $\rho(\vec{r})$ is, regardless of the number of electrons of the system, a function of three spatial variables only. The simplification is evident and ingenious but still does not provide a practical solution since a functional (namely a prescription for producing numerical values from a function) connecting the electron density to the energy of the system is required. The design of functionals accurately connecting the electron density to the energy therefore represents the core of any DFT method.

2.1.2 Kohn-Sham method

The practical solution for the problem just mentioned above was provided by Kohn and Sham [16]. The method they proposed takes into account a fictitious system in which the electron-electron interaction is absent, allowing in this way an analytic solution of the Schrödinger equation. The electron density of the fictitious system is imposed to be equivalent to that of the real one:

$$\rho(\vec{r}) := \rho^s(\vec{r}) \tag{2.9}$$

thanks to the use of an external potential v_s .

The Hamiltonian operator of the fictitious system is therefore:

$$\hat{H}_s = \sum_{i=1}^N \left[-\frac{\nabla_i^2}{2} + v_s(\vec{r}) \right] = \sum_{i=1}^N \hat{h}_i^{KS} \quad (2.10)$$

where \hat{h}_i^{KS} is the Kohn-Sham (KS) single-electron Hamiltonian. The factorization shown in Equation 2.10 would in general be not possible, due to the lack of account of the mutual interactions between the electrons. Such interactions are however neglected in the virtual system previously defined, therefore permitting this operation and the interpretation of the eigenvalues obtained by the application on KS orbitals $\phi(\vec{r})$ as single electron energies:

$$\hat{h}_i^{KS} \phi_i^{KS}(\vec{r}_i) = \epsilon_i^{KS} \phi_i^{KS}(\vec{r}_i) \quad (2.11)$$

Being the kinetic energy of the fictitious system the sum of the single electron contributions:

$$\langle T_s \rangle = \sum_{i=1}^N \left\langle \phi_i^{KS}(\vec{r}_i) \left| -\frac{\nabla_i^2}{2} \right| \phi_i^{KS}(\vec{r}_i) \right\rangle \quad (2.12)$$

the total energy of the fictitious system can be expressed as:

$$E_s = \langle T_s \rangle + \int v_s(\vec{r}) \rho(\vec{r}) d\vec{r} \quad (2.13)$$

since each term is known and can be separately calculated for its independence from the others. Consequently, this equation can be used to describe the real system, adding a term taking into consideration the previously neglected interaction between electrons. The resulting equation can be arranged to highlight the unknown term as difference between real and fictitious terms:

$$\begin{aligned} E[\rho(\vec{r})] = & \sum_i \left\langle \phi_i^{KS} \left| -\frac{\nabla_i^2}{2} \right| \phi_i^{KS} \right\rangle + \int v(\vec{r}) \rho(\vec{r}) d\vec{r} + \frac{1}{2} \int \int \frac{\rho(\vec{r}) \rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' + \\ & \left(T[\rho(\vec{r})] - \sum_i \left\langle \phi_i^{KS} \left| -\frac{\nabla_i^2}{2} \right| \phi_i^{KS} \right\rangle \right) + \\ & \left(V_e[\rho(\vec{r})] - \frac{1}{2} \int \int \frac{\rho(\vec{r}) \rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' \right) \end{aligned} \quad (2.14)$$

The first three terms respectively represent the kinetic electronic energy, the nuclear-electron and the electron-electron interactions calculated on KS orbitals. The last two terms (those into the parentheses) are the differences between the

real and the fictitious system energies, but they still cannot be calculated since the functional form of $T[\rho(\vec{r})]$ and $V_e[\rho(\vec{r})]$ is unknown. In order to simplify the equation, these two terms are grouped into a single functional quoted as exchange and correlation functional $V_{xc}[\rho(\vec{r})]$, as it arises from these two typically quantum properties of the electrons of the system (neglected in the fictitious one). To overcome the problem of the unknown form of V_{xc} , Kohn and Sham introduced a variational calculation, within a self consistent field (SCF) approach, which is very similar to that employed in Hartree-Fock (HF) calculations.

$$\left[-\frac{\nabla_i^2}{2} - \sum_I \frac{Z_I}{r_{iI}} + \frac{1}{2} \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + V_{xc} \right] \phi_i^{KS}(\vec{r}_i) = \epsilon_i^{KS} \phi_i^{KS}(\vec{r}_i) \quad (2.15)$$

Starting from trial orbital wavefunctions ϕ_i^{KS} , a minimization is performed in order to obtain the minimum eigenvalue of energy for Equation 2.15. For each SCF step, a new set of KS orbitals is obtained and thus used in the next step to calculate the electron density and a lower energy value is achieved until convergence.

2.1.2.1 DFT functionals

The inclusion of all the unknown terms into the V_{xc} functional has however only shifted the problem, since a functional form for the exchange and correlation potential is still lacking. In the years, many different approaches have been proposed to model the contributions to kinetic and potential energies due to the electronic repulsions. The most used approximation for the exchange and correlation term in a strongly correlated system are the Local Density Approximation (LDA)[16], the Generalized Gradient Approximation (GGA)[38], the meta-GGA, the hyper-GGA and the double-hybrid approaches.

2.1.2.1.1 LDA Within the LDA, the electronic density is considered to change very slowly, and therefore, on a local scale, this quantity can be approximately considered to be homogeneous.

$$E_{xc}^{LDA}[\rho(\vec{r})] = \int \rho(\vec{r}) \epsilon_{xc}(\rho) d\vec{r} \quad (2.16)$$

Such an approximation is very useful, because for a homogeneous electron gas the exchange and correlation functional is well-known[39, 40, 41, 42, 43] and exact and can therefore act as a model for the functionals of this family. PW[43] and VWN[42] functionals notably belong to this category.

2.1.2.1.2 GGA A more generalized formalism is required when the electron density has large spatial fluctuations, as in common molecular systems. In GGA both the dependencies on the density and its gradient are included and, as shown in Equation 2.17, a different contribution based on the electronic spin moment (α and β or \uparrow and \downarrow) can be considered:

$$E_{XC}^{GGA} [\rho^\uparrow, \rho^\downarrow] = \int \rho(\vec{r}) \epsilon_{xc}(\rho^\uparrow, \rho^\downarrow, |\nabla\rho^\uparrow|, |\nabla\rho^\downarrow|) d\vec{r} \quad (2.17)$$

In some cases the correlation part is determined with Quantum Monte Carlo methods at different electronic densities. Two of the early but still largely employed functionals of this kind are, respectively for the exchange and correlation part, B88[44] and LYP[45] (from the acronym of its developers Lee, Yang and Parr).

2.1.2.1.3 Meta-GGA The formalism of meta-GGA is the logical extension of GGA methods, allowing the exchange and correlation functional to depend on the Laplacian ($\nabla_{\vec{r}}^2$) and on higher order derivatives of the electronic density. As could be easily imagined, the increase in the complexity of the V_{xc} expression enlarges the accuracy of the calculations but, unfortunately, the computational cost raises appreciably.

2.1.2.1.4 Hyper-GGA At variance, hyper-GGA (or, more commonly, hybrid) functionals are based on a slightly different assumption. Knowing in fact that the exchange term coming from Hartree-Fock theory is exact for a system of non-interacting electrons and relying on the Adiabatic Connection Formula (ACF)[46], these functionals are characterized by an exchange and correlation part that contains a fraction of exact HF exchange, a LDA formula and a gradient correction term.

The most used functional of this family (and probably one of the most diffused at all) is the B3LYP[45, 44] functional, shown in Equation 2.18.

$$E_{XC}^{B3LYP} = E_X^{LDA} + a_0 (E_X^{HF} - E_X^{LDA}) + a_x (E_X^{GGA} - E_X^{LDA}) + E_C^{LDA} + a_c (E_C^{GGA} - E_C^{LDA}) \quad (2.18)$$

2.1.2.1.5 Double-hybrid In the last years, new double-hybrid functionals[47] have been implemented. For such hybrid functionals (such as B3LYP), a fraction of the correlation energy from the second order Møller-Plesset perturbative

method[19] is added and in this way the dynamical correlation of the electronic motion is taken into account. Two of the most diffused functionals of this family, still scarcely adopted due to their high computational cost, are B2PLYP[48] and mPW2PLYP[49], both developed by Grimme.

2.1.2.1.6 Summary It is important to point out, at the end of this list, that in most cases a functional that is very efficient in the calculation of some properties is very poor, or completely wrong, in the calculation of the others. Even though the order in which functionals were presented retraces that of the so-called "Jacob's ladder", a classification of the functionals based on their accuracy introduced by Kurth, Perdew, and Blaha [50], the matter is unfortunately much less straightforward than it could appear. Nowadays, in fact, a universal functional which is able to finely describe every system and every property does still not exist.

2.1.3 The problem of static correlation

Both in DFT and in HF calculations the total wavefunction of the system expressed using the Slater determinant[51] formalism is:

$$\Psi(\vec{r}_1, \dots, \vec{r}_n, \vec{s}_1, \dots, \vec{s}_1) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\vec{r}_1) \cdot \chi_1(\vec{s}_1) & \cdots & \phi_n(\vec{r}_1) \cdot \chi_n(\vec{s}_1) \\ \vdots & \ddots & \vdots \\ \phi_1(\vec{r}_n) \cdot \chi_1(\vec{s}_n) & \cdots & \phi_n(\vec{r}_n) \cdot \chi_n(\vec{s}_n) \end{vmatrix} \quad (2.19)$$

where $\phi(\vec{r})$ and $\chi(\vec{s})$ are the spatial and spin part of the mono-electronic wavefunction, respectively.

Such a simple, but very ingenious idea, insures that the resulting wavefunction is antisymmetric with respect to the exchange of fermionic particles (electrons), as prescribed by Pauli's exclusion principle.

This kind of expression for the total wavefunction works well for closed-shell systems (namely molecules having an equal number of α and β electrons, all of them grouped in pairs having the same spatial wavefunction *i.e.* belonging to the same orbital) or for high spin open-shell systems (having unpaired electrons with the same spin).

It is however possible to have systems containing unpaired electrons which do not have the same spin component, thus having a total spin $(\frac{1}{2} \cdot (n^\alpha - n^\beta))$ smaller than the highest theoretically conceivable one $(\frac{1}{2} \cdot (n_{unpaired}^\alpha + n_{unpaired}^\beta))$. This phenomenon is due to the superexchange (known as Kramers-Anderson

exchange) process[52, 53] which, differently from the classical exchange interaction, induces spin moments over different atomic nuclei to couple among them in an antiferromagnetic (antiparallel) way in the ground state.

Physically, the phenomenon occurs when an atom having its external electronic shell fully occupied bridges the two nuclei possessing unpaired electrons. The orbitals occupancy and Pauli's exclusion principle impose the unpaired electrons to have antiparallel orientation, as shown in Figure 2.1.

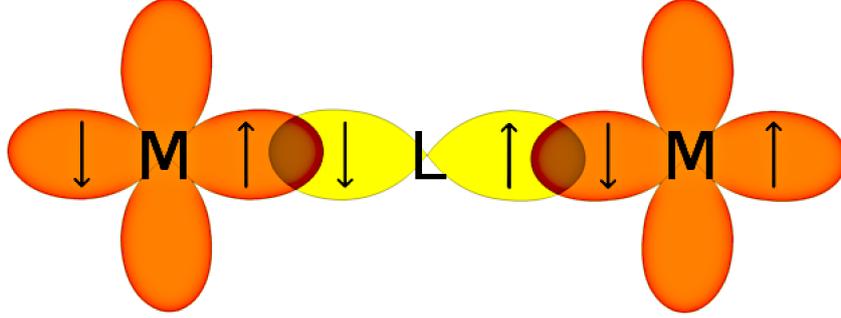


Figure 2.1: Example of superexchange interaction.

Such a situation can be phenomenologically described by employing the Heisenberg-Dirac-Van Vleck Hamiltonian[54, 55, 56]:

$$\hat{H}_{HDvV} = -2 \sum_{i,j < i} J_{ij} \hat{s}_i \hat{s}_j \quad (2.20)$$

Here \hat{s}_i and \hat{s}_j are the spin operators of magnetic momenta of the i and j centres and J_{ij} is the coupling constant between the metal centres.

The sign of J defines the kind of the energetically favored coupling: more precisely, if the sign is positive, the lowering of the potential energy is produced by the alignment of the spin moments on nuclei i and j (that is true for ferromagnetic coupling). If, on the contrary, J is smaller than zero, then the antiparallel arrangement of the spin moments (occurring for antiferromagnetic coupling) is preferred.

For these kinds of systems, the wavefunction expressed as in Equation 2.19 does no more provide an even qualitatively correct description. The wavefunction produced is in fact no more an eigenfunction of the spin squared operator \hat{S}^2 and, when such operator is applied on it, a linear combination of spin functions with different multiplicities is obtained. The wavefunction is thus defined "spin-contaminated" and, following from the principles of quantum mechanics, the associated energy is higher than for a completely analogous wavefunction possessing the correct spin symmetry. Another largely employed definition for

this kind of wavefunctions is "Broken Symmetry"[57, 58] (BS), owing to the lack of the correct spin symmetry.

The most rigorous way to deal with this particular electronic arrangement without incurring in the spin contamination is the use of multiconfigurational methods such as CASSCF[59]. This method, or more accurate ones such as MRCI[60] and NEVPT2[61], are correctly expected to treat the static correlation phenomenon that arises when dealing with degenerate (or nearly-degenerate) orbitals by adopting a trial wavefunction that is a sum of many Slater determinants:

$$\Psi^{CASSCF} = \sum_i c_i \Psi^{Slater} \quad (2.21)$$

These computational methods are however computationally expensive because the number of Slater determinants to be included in the linear combination quickly increases with the size of the studied system. The multideterminantal methods are therefore manageable for systems with a limited number of electrons.

For this reason, a number of procedures have been proposed with the aim of decontaminating, and thus correcting, the BS calculation results, producing the exchange coupling constants J (which can be experimentally measured through EPR spectroscopy) of the systems as side-product. The group of K. Yamaguchi, for instance, proposed the approximate spin projection method[62] (later adapted to be employed in geometrical optimizations[63, 64] and generalized to multi-centre systems[65]). M. Head-Gordon and A. Krylov, on the contrary, proposed the spin-flip technique[66, 67]. Geometrical optimizations using the former approach cannot however be at the moment accomplished. Many studies have been conducted by the group of Malrieu[68] in this field, but a generalization to systems having more than two magnetic centres or bearing a spin higher than $\frac{1}{2}$, although envisaged, has never been applied so far. Procedures based on the mapping of BS-DFT energy levels onto Ising Hamiltonian eigenstates have been further proposed[69] and broadly used in the study of magnetic systems, for example by Verma, Maurice, and Truhlar [70].

2.2 Vibrational Properties

The study of the vibrational states associated with the small displacements of the atomic nuclei around the equilibrium configuration of a molecule are the ba-

sis of the Infrared and Raman spectroscopies (spanning within the wavenumber range $10000 \div 10 \text{ cm}^{-1}$). The spectral range may at any rate be extended to higher wavenumbers because overtone spectroscopy is at the present time gaining a particular interest[71]. Vibrational spectroscopy is widely used in Physics and Chemistry for several purposes and is often flanked by complementary experimental and theoretical methods.

A particular interest is connected to the medium IR (MIR) range ($4000 - 400 \text{ cm}^{-1}$), because structural analysis can be usually carried out in this region of the spectrum. For simple compounds, the assignment of diagnostic absorption bands to a particular vibrational mode might reasonably be easy, although quick increase of molecular complexity often requires theoretical tools in order to correctly investigate the spectra.

The most used computational approach is the Normal Modes Analysis (NMA) allowing to decompose the complex Lissajous motion of a molecule (due to thermal excitation exerted by the environment) into molecular vibrations (also known as molecular modes) within the harmonic approximation. A common example of NMA may be represented by the theoretical infrared spectrum of a water molecule (a symmetrical non-linear triatomic molecule of C_{2v} point-group symmetry), shown in Figure 2.2. It must be pointed out that the width of the signals depends on a series of different phenomena, which would deserve to be considered apart. For this reason, they are imposed *a posteriori*, being based on simplified considerations.

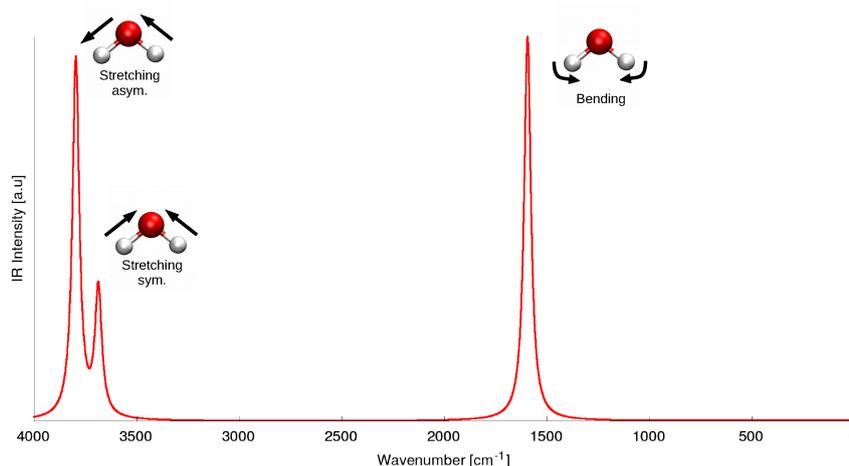


Figure 2.2: Theoretical IR spectra of water molecule obtained from NMA.

In systems including loosely bound species, harmonic approximation is expected to be very poor and it could therefore be necessary to employ algorithms

including anharmonic effects[72, 73, 74]. Such algorithms are however extremely expensive from a computational point of view and are scarcely implemented in common electronic structure simulation packages, their use being therefore restricted to small sized systems.

2.2.1 Normal Mode Analysis

For an oscillating system, a normal mode is a pattern of motion where all the parts of the molecule are moving harmonically at the same frequency, in phase and with different amplitudes. Such vibrations are fully independent (*e.g.* orthogonal). The fixed frequencies typical of the normal modes of a system are known as its resonant frequencies.

As previously written, Normal Mode Analysis is generally performed in the harmonic approximation. This means that the potential energy of the molecule is expressed around the equilibrium position \vec{r}_{eq} as a Taylor expansion truncated to the second order:

$$V(\vec{r}) \simeq V(\vec{r}_{eq}) + \left(\frac{\partial V(\vec{r}_{eq})}{\partial \vec{r}} \right) \vec{r} + \frac{1}{2!} \vec{r}^T \left(\frac{\partial^2 V(\vec{r}_{eq})}{\partial \vec{r}^2} \right) \vec{r} \quad (2.22)$$

The first term is a constant and can be arbitrarily set to zero. Since the expansion is around the equilibrium position, even the term containing the first derivative is null. The final equation therefore is:

$$V(\vec{r}) = \frac{1}{2} \vec{r}^T \mathcal{H}_{ij} \vec{r} \quad (2.23)$$

where \mathcal{H} is the $3N \times 3N$ Hessian matrix of the system, containing the second derivatives of the potential energy with respect to the atomic positions.

The use of the transformation matrix \mathbf{G} (having elements in the form $G_{ij} = \frac{1}{\sqrt{m_i m_j}}$) converts the Hessian matrix into its mass-weighted form.

Eventually, the diagonalization of the product matrix $\mathcal{H} \cdot \mathbf{G}$ permits to obtain eigenvalues $\{\varepsilon\}$ and eigenvectors $\{\vec{q}\}$. The formers are related to the vibrational frequencies by the relation:

$$\nu_i = \frac{1}{2\pi c} \sqrt{\varepsilon_i} \quad [cm^{-1}] \quad (2.24)$$

where c is the speed of light. The latters are instead the mass-weighted vibrational normal coordinates.

One should remind that six (or five, for a linear system) eigenvalues are exactly zero, being associated to the translational and rotational modes. This

means that the number of harmonic vibrational frequencies of a molecule is $3N - 6$ (or $3N - 5$ for linear shaped molecules). In the computational procedure, due to numerical issues that arises even when the vibrational frequencies are computed using an analytical method, the eigenvalues associated to rotations and translations can have values of up to 50 cm^{-1} and mix with the lowest actual vibrational modes. For this reason, the translational and rotational degrees of freedom are generally removed from the Hessian matrix before performing the diagonalization by using a projection procedure.

2.3 Extended Broken Symmetry

As previously reported, in the low-spin open-shell species the single-determinant character of the Kohn-Sham formulation of the DFT is lacking of the multi-determinant character that is necessary to correctly describe the ground state. Such an electronic arrangement is usually found in some transition metal-complexes; these elements in fact usually possess a large number of unpaired electrons, parallelly aligned thanks to the direct exchange interaction stabilizing electrons having the same sign of the spin moment. The presence of non-metallic atoms (such as sulphur and oxygen) acting as bridge between the metal centres induces, as explained in subsection 2.1.3, the antiparallel alignment of the total spin moments on each metallic atom. Among metallic atoms, iron widely shows this behaviour, due to the presence of five unpaired electrons in its ferric (Fe^{3+}) form. Manganese complexes are further examples of strongly magnetic species.

The basic idea behind many theoretical methods employed for dealing with these systems is to describe the energetic features of the ground state through the Heisenberg or Heisenberg-Dirac-van Vleck (HDvV) empirical Hamiltonian shown in Equation 2.20, estimating the magnetic J-coupling using independent total energy calculations of single-determinant unrestricted spin configurations. With such EBS (Extended Broken Symmetry) method, two-centre [75, 76, 77, 78, 79] and multi-centre [65] transition metal complexes have been investigated. In the EBS method, the ground state energy surface of a multi-centre complex is reconstructed using separated calculations performed for all the possible broken symmetry (BS) and high spin (HS) states through the evaluation of the J-coupling constants. The energetics of the ground and all the other spin states, the constituents of the "spin ladder", can therefore be calculated obtaining a very good compromise with respect to the post-Hartree-Fock correlated methods in terms of accuracy versus computational cost.

As it could be easily figured out by recalling the physics behind the superexchange process, briefly outlined in subsection 2.1.3, the values of the coupling constants strongly depend on the structure of the compound, as well as on the metal-ligand bond distances and on the metal-ligand-metal bond angles. For this reason, an important ingredient for the quantitative evaluation of the spin energetics is represented by the possibility of performing geometrical optimizations on the proper ground state potential energy surface. These calculations are of paramount importance since they permit to shed light on the structural variability[80, 81, 82] and whenever the experimental structures are unclear[83, 75].

However, the use of the gradient procedure for geometry optimizations within the EBS approximation has been for quite a long time limited to di-nuclear transition metal complexes, which represent just a small portion of the many interesting and useful prosthetic groups in metalloenzymes.

For the first time, in 2017, the EBS formalism for geometry optimizations was generalized for handling those systems containing an arbitrary number of spin centres, allowing to fully optimize multi-centre transition-metal complexes as large as desired[84]. Moreover, in the course of my PhD work in the group of Prof. L. Guidoni, I further extended[85] this technique to the calculation of the Hessian matrix, for normal mode analysis and harmonic vibrational frequency calculations (described in section 2.2) within the EBS method

2.3.1 Estimation of J-coupling constant

For the general case of N spin centres, the HDvV spin Hamiltonian, \hat{H}_{HDvV} can be expressed as the sum of the two spin centre operators:

$$\begin{aligned}\hat{H}_{HDvV} &= -2 \sum_{i < j}^N J_{ij} \hat{s}_i \cdot \hat{s}_j \\ &= - \sum_{i < j}^N J_{ij} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2),\end{aligned}\tag{2.25}$$

where \hat{s}_i is the spin operator on the spin site i , J_{ij} is the effective exchange coupling constant between the spin centres i and j and $\hat{S}_{ij} = \hat{s}_i + \hat{s}_j$.

The eigenvalues and eigenstates of this Hamiltonian generate a ladder of states, each one characterized by total spin S_{tot} , in the range $S_{tot} \in [S_{min}, S_{max}]$, where $S_{min} = \min|\sum_i \pm s_i|$ (the low spin state LS, often coinciding with the ground state GS) and $S_{max} = \sum_i |s_i|$ (high spin state HS).

The total number of independent BS spin states $|\text{BS}_k\rangle$, where k is the index of the BS spin state, equals to $2^N/2$. For the sake of simplicity, even the high spin state is considered to be a special instance of broken symmetry state. All these states, with the exception of the HS one as stated in subsection 2.1.3, are not eigenstates of the \hat{S}^2 operator. The spin contamination through DFT calculations[58] can be estimated by calculating the average value of the \hat{S}_{tot} operator as:

$$\begin{aligned}\langle \hat{S}_{tot}^2 \rangle_k &= \langle \left(\sum_{i=1}^N \hat{s}_i \right)^2 \rangle_k = \sum_i \langle \hat{s}_i^2 \rangle_k + 2 \sum_{i<j} \langle s_i \cdot s_j \rangle_k \\ &= \sum_i (s_i(s_i + 1))_k + 2 \sum_{i<j} \langle s_i \cdot s_j \rangle_k\end{aligned}\quad (2.26)$$

The average $\langle s_i \cdot s_j \rangle_k$ represents the spin correlation function on the BS spin state $|\text{BS}_k\rangle$ between the spin sites i and j . Under the essential assumption that the absolute value of $\langle s_i \cdot s_j \rangle_k$ remains the same under all the BS states, the spin correlation functions $\langle s_i \cdot s_j \rangle$ can be estimated by solving the system of linear equations 2.26 for all BS spin states $|\text{BS}_k\rangle$.

According to Equation 2.20, the spin interaction energy ε_k^{BS} of BS spin state $|\text{BS}_k\rangle$ can be expressed as:

$$\begin{aligned}\varepsilon_k^{BS} &= \langle \text{BS}_k | \hat{H} | \text{BS}_k \rangle + \text{constant} \\ &= -2 \sum_{i<j} J_{ij} \langle \text{BS}_k | \hat{s}_i \cdot \hat{s}_j | \text{BS}_k \rangle + \text{constant} \\ &= -2 \sum_{i<j} J_{ij} \langle s_i \cdot s_j \rangle_k + \text{constant} \\ &= -2 \sum_{i<j} J_{ij} A_{kp} + \text{constant}\end{aligned}\quad (2.27)$$

where p refers to a particular i, j spin pair and A_{kp} equals to $\langle s_i \cdot s_j \rangle_k$, namely the average of the spin correlation function between the centres i and j over the k -th BS state.

At this point, the following system of linear equations is defined:

$$\begin{pmatrix} \varepsilon_1^{BS} \\ \varepsilon_2^{BS} \\ \varepsilon_3^{BS} \\ \vdots \\ \varepsilon_{N_k}^{BS} \end{pmatrix} = -2 \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N_p} \\ A_{21} & A_{22} & \cdots & A_{2N_p} \\ A_{31} & A_{32} & \cdots & A_{3N_p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N_k1} & A_{N_k2} & \cdots & A_{N_kN_p} \end{pmatrix} \begin{pmatrix} J_1 \\ J_2 \\ J_3 \\ \vdots \\ J_{N_p} \end{pmatrix}\quad (2.28)$$

where $N_p = \frac{N(N-1)}{2}$ and ε_k^{BS} are the DFT energy differences between the BS state[86] k and the high spin state. By expressing the energies in this way, the constant term they contain and that independent on the spin description can be removed and ε_0^{BS} is zero by definition.

Using the matrix notation, where $\vec{\varepsilon}^{BS} = (\varepsilon_1^{BS}, \varepsilon_2^{BS}, \dots, \varepsilon_{N_p}^{BS})$ and $\vec{J} = (J_1, J_2, \dots, J_{N_p})$, Equation 2.28 can be rewritten as:

$$\vec{\varepsilon}^{BS} = -2\mathbf{A} \cdot \vec{J}. \quad (2.29)$$

and the the coupling constants \vec{J} of the systems are recovered as:

$$\vec{J} = -\frac{1}{2}\mathbf{A}^{-1} \cdot \vec{\varepsilon}^{BS} \quad (2.30)$$

Being the system overdetermined, the inverse of matrix \mathbf{A} can be evaluated by employing Singular Value Decomposition technique[87] (SVD).

2.3.2 Spin ladder

Based on the J-coupling constants introduced in subsection 2.3.1, it is possible to build the Hamiltonian matrix in an appropriate left- (right-) side basis set $\langle b_l |$ ($| b_r \rangle$) with index $1 \leq l \leq N_b$ ($1 \leq r \leq N_b$), where the total number of the spin states N_b is equal to $\prod_{i=1}^N (2s_i + 1)$.

The coupled state of the spin sites i and j is $| S_{ij} M_{ij} \rangle$, where S_{ij} is the total spin and M_{ij} is the total spin moment of spin i and j .

The detailed formula of the right-side basis set $| b_r \rangle$ of the N spin site systems is the following:

$$| b_r \rangle = | s_1 s_2 \dots s_N; m_1 m_2 \dots m_N \rangle \quad (2.31)$$

The elements of the Hamiltonian matrix can be written as:

$$\begin{aligned} \langle b_l | \hat{H} | b_r \rangle &= -2 \sum_{i < j} \langle b_l | J_{ij} \hat{s}_i \cdot \hat{s}_j | b_r \rangle \\ &= - \sum_{i < j} J_{ij} \langle b_l | \hat{S}_{ij}^2 | b_r \rangle + \sum_{i < j} J_{ij} \langle b_l | (\hat{s}_i^2 + \hat{s}_j^2) | b_r \rangle \\ &= - \sum_{i < j} J_{ij} \times CG_{ij}^l \times CG_{ij}^r \times S_{ij}(S_{ij} + 1) + \sum_{i < j} J_{ij} \times (s_i(s_i + 1) + s_j(s_j + 1)), \end{aligned} \quad (2.32)$$

where Clebsch-Gordan (CG) coefficients $CG_{ij}^l = \langle S_{ij}M_{ij} | s_i s_j; m_i m_j \rangle$ and $CG_{ij}^r = \langle s_i s_j; m_i m_j | S_{ij}M_{ij} \rangle$ were calculated following the scheme reported in subsection 6.1.1.

The diagonalization of the matrix provides the complete ladder of eigenstates and eigenvalues and from this list one can identify the ground state, which, as already reported earlier in this work, usually coincides with the state having the lowest total spin.

2.3.3 Ground state forces

The energies of the high spin (HS) and ground states (GS) can be written as:

$$\begin{aligned}\varepsilon^{HS} &= \langle \Psi^{HS} | \hat{H} | \Psi^{HS} \rangle \\ \varepsilon^{GS} &= \langle \Psi^{GS} | \hat{H} | \Psi^{GS} \rangle\end{aligned}\tag{2.33}$$

where $|\Psi^{HS}\rangle$ and $|\Psi^{GS}\rangle$ are the eigenstates of the HS and GS state, respectively. Following the basic principles of quantum mechanics, they can be rewritten in terms of the basis set $|b_r\rangle$:

$$\begin{aligned}|\Psi^{HS}\rangle &= \sum_r c_r^{HS} |b_r\rangle \\ |\Psi^{GS}\rangle &= \sum_r c_r^{GS} |b_r\rangle\end{aligned}\tag{2.34}$$

using the appropriate coefficients $c_{r(l)}^{GS(HS)}$.

Defining:

$$\Delta\varepsilon = \varepsilon^{HS} - \varepsilon^{GS}\tag{2.35}$$

Equation 2.33 and Equation 2.34 can be combined for writing:

$$\begin{aligned}\Delta\varepsilon &= \sum_{lr} c_l^{HS} c_r^{HS} \langle b_l | \hat{H} | b_r \rangle - \sum_{lr} c_l^{GS} c_r^{GS} \langle b_l | \hat{H} | b_r \rangle \\ &= \sum_{lr} \left[c_l^{HS} c_r^{HS} H_{lr} - c_l^{GS} c_r^{GS} H_{lr} \right] \\ &= \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) H_{lr} \right]\end{aligned}\tag{2.36}$$

whose gradient is given as:

$$\begin{aligned}
\nabla_{\vec{R}}(\Delta\varepsilon) &= \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \nabla_{\vec{R}}(\langle b_l | \hat{H} | b_r \rangle) + \right. \\
&\quad \left. \nabla_{\vec{R}}(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \langle b_l | \hat{H} | b_r \rangle \right] \\
&= \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \langle b_l | \nabla_{\vec{R}} \hat{H} | b_r \rangle \right. \\
&\quad \left. + (\nabla_{\vec{R}} c_l^{HS} c_r^{HS} + c_l^{HS} \nabla_{\vec{R}} c_r^{HS} - \nabla_{\vec{R}} c_l^{GS} c_r^{GS} - c_l^{GS} \nabla_{\vec{R}} c_r^{GS}) \langle b_l | \hat{H} | b_r \rangle \right]
\end{aligned} \tag{2.37}$$

Following what already assumed for the binuclear *specimen* by Bovi and Guidoni [77], one may assume that the main contribution to the gradient comes from the derivative of the Hamiltonian, namely the first half of Equation 2.37. The contribution connected with the geometrical dependence of the basis set coefficients, reported in the second half of Equation 2.37, may be neglected:

$$\nabla_{\vec{R}} c_{l(r)}^{HS(GS)} := 0 \tag{2.38}$$

Considering the definition of the HDvV Hamiltonian 2.25, the gradient of the Hamiltonian with respect to the nuclear coordinates $\nabla_{\vec{R}} \hat{H}$ is written as:

$$\nabla_{\vec{R}} \hat{H} = -\nabla_{\vec{R}} \sum_{i<j} J_{ij} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) = -\sum_{i<j} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) \nabla_{\vec{R}} J_{ij}, \tag{2.39}$$

because the spin operators \hat{S}_{ij}^2 , \hat{s}_i^2 and \hat{s}_j^2 are independent on the spatial coordinates, and consequently $\nabla_{\vec{R}} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) = 0$.

Using Equation 2.30 and the derivative chain rule:

$$\begin{aligned}
\nabla_{\vec{R}} \hat{H} &= -\sum_{i<j} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) \sum_{k=1}^{N_k} \frac{\partial J_{ij}}{\partial E_k} \nabla_{\vec{R}} E_k^{BS} \\
&\quad - \sum_{i<j} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) \sum_{k=1}^{N_k} -A_{kp}^{-1} \nabla_{\vec{R}} E_k^{BS}
\end{aligned} \tag{2.40}$$

The gradient of the energy difference, $\nabla_{\bar{R}}(\Delta\varepsilon)$ is therefore the following one:

$$\begin{aligned}
\nabla_{\bar{R}}(\Delta\varepsilon) &= \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \langle b_l | \nabla_{\bar{R}} \hat{H} | b_r \rangle \right] \\
&= \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \langle b_l | - \sum_{i<j} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) \sum_k \frac{\partial J_{ij}}{\partial E_k} \nabla_{\bar{R}} E_k^{BS} | b_r \rangle \right] \\
&= - \sum_k \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \langle b_l | \sum_{i<j} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) \frac{\partial J_{ij}}{\partial E_k} | b_r \rangle \nabla_{\bar{R}} E_k^{BS} \right] \\
&= \sum_k -f_k \nabla_{\bar{R}} E_k^{BS}.
\end{aligned} \tag{2.41}$$

and the force F^{GS} can be formulated as:

$$F^{GS} = -\nabla_{\bar{R}} E^{GS} = -(-F^{HS} - \nabla_{\bar{R}}(\Delta\varepsilon)) = F^{HS} + \nabla_{\bar{R}}(\Delta\varepsilon), \tag{2.42}$$

where the force F^{HS} on the HS state can be obtained by DFT or another electronic structure method.

By putting Equation 2.41 into Equation 2.42, we can evaluate the force on the ground state.

$$F^{GS} = F^{HS} - \sum_k f_k \nabla_{\bar{R}} E_k^{BS} \tag{2.43}$$

where $f_k = \sum_{lr} \left[(c_l^{HS} c_r^{HS} - c_l^{GS} c_r^{GS}) \langle b_l | \sum_{i<j} (\hat{S}_{ij}^2 - \hat{s}_i^2 - \hat{s}_j^2) \frac{\partial J_{ij}}{\partial E_k} | b_r \rangle \right]$.

Any first-order minimization algorithm can then be used to optimize the molecular geometry until the desired convergence criteria are met.

2.4 Classical Molecular Dynamics

Classical Molecular Dynamics simulations (MD) are performed using an algorithm that computes the evolution of a many-body system described at microscopic level. Many flavors of MD simulations are known, going from fully atomistic description of the system to the coarse-graining[88] approach, in which chemical groups or even entire molecules are approximated as single units. In this context, the term "classical" means that the particles of the system follow the laws of classical mechanics and the forces are calculated in an empirical way, without going in depth into the proper quantum-mechanical nature of the

particles' interactions. The term is required to be able to distinguish this kind of computational method from the *Ab Initio* Molecular Dynamics (classified as Car-Parrinello[29] Molecular Dynamics and Born-Oppenheimer[30] Molecular Dynamics), in which the atomic forces are calculated by the first principles while the atomic positions are propagated in a classical way, and from Path-Integral Molecular Dynamics, in which the Path-Integral formalism introduced by Feynmann[89, 90] is employed for propagating the nuclear positions taking into account the peculiar effects produced by the quantic nature of the atoms (*e.g.* the tunnel effect[91]).

Any simulation of Classical Molecular Dynamics, in spite of the differences between the available algorithms for its implementation, is however based on the principle that the properties of interest for a system are sampled by simulating its evolution over a sufficiently long period of time to eliminate any statistical noise that can bias the results.

Simulate the evolution of a system means to numerically solve its differential equations of motion, namely to choose a time step and calculate the forces acting on each element of the system at each unit of this interval, consequently updating their positions (which likely are going to be changed).

At each step of the dynamics, the evolution of the system is calculated by reversing the cornerstone of classical mechanics, the Newton's equation:

$$\vec{F}_i(t) = m_i \ddot{x}_i(t) \quad (2.44)$$

where $\vec{F}(t)$ is the total force acting on the particle i , m_i is its mass and $x_i(t)$ is its position at the time t .

A generic MD algorithm can be summarized in a few lines of pseudocode:

```

1 init()
2 t = 0
3 while(t < tmax)
4     force(t)
5     integrate(t)
6     t += time_step
7     print_traj(t)
8 end
9 perform_analysis()
```

The first line is the initialization of the system. At this stage, the initial positions and velocities of each element of the system are assigned. The positions are usually given by the user, while the velocities can be given by the user or randomly assigned following a distribution compatible with the target temperature of the simulation. Eventually, a certain number of molecular properties of the system are chosen to be included in the simulation (temperature and/or pressure, etc.). Specific algorithms called thermostats[92] or barostats correct the atomic velocities and/or positions in order to match the imposed values of the observables, nevertheless allowing a certain amount of fluctuation in order to keep the simulation as realistic as possible.

After the calculation of the forces, the fifth line of the code computes the position of each particle by integrating the Newton's equation of motion. Several algorithms have been purposely proposed.

The most used algorithm is the velocity Verlet algorithm. It is an evolution of the former Verlet algorithm[93] that provides an explicit expression for the velocities, a lacking feature in the original Verlet algorithm and that could be recovered only at the price of lower accuracy (with an error that is $O(\Delta t)$). For velocity Verlet the accuracy for calculated velocities is instead naturally $O(\Delta t^3)$. The expressions for the position and velocity of a given particle at the time $t + \Delta t$ are:

$$\vec{x}(t + \Delta t) = \vec{x}(t) + \vec{v}(t) + \frac{1}{2}\vec{a}(t) \Delta t \quad (2.45)$$

and

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \frac{\vec{a}(t) + \vec{a}(t + \Delta t)}{2} \Delta t \quad (2.46)$$

Systems are usually simulated within a sufficiently large box, where the "box" means any geometrical three-dimensional region of the space (usually, but not always, of cubic shape). The use of a box is for example mandatory in *ab initio* MD that uses plane waves as basis set, because the electronic density has to be calculated over a well-defined spatial region. There are many instances where it is useful, or necessary, to simulate an infinite (at least with respect to the very small scale of a Molecular Dynamics simulation) system (for example, if a liquid or a solid has to be simulated). In this case, the box is considered to have periodic boundary conditions (PBC), namely an exact copy of the system is placed on each side of the box (and an exact copy of the system for each side of the box of every copy created, and so on). Positions and velocities dynamically change only for the particles within the main box, being all the remaining

particles copies of the former ones, but the influence of all the particles in the "infinite" system are considered in the force evaluation (this is particularly important when thinking about long-range electrostatic interactions). Moreover, in this way, a particle leaving the box along one side is expected to rejoin it from the other side, keeping the number of particles of the system constant.

2.4.1 Force fields

In the context of Molecular Dynamics, a force field is a set of functional forms and parameters used to calculate the potential energy of the system chosen for the simulation. The potential energy functional form includes bonded terms for the interactions of atoms linked by a few covalent bonds, and non-bonded terms that describe the interactions between distant atoms, which could even belong to different molecules.

It exists a very large number of different force fields[94], each of them characterized by its functional form, parameters and, consequently, level of accuracy. Many force fields are in fact tailored for a specific application or kind of system and their performance is scarce if they are applied on topics different from those foreseen during their design. In general, the potential energy function (with the notable exception of the recently developed "numerical" force fields that make use of Machine Learning techniques) is decomposed as in Equation 2.47, namely as a sum of terms connected with the geometrical variations inside each molecule (stretching of a bond, bending of an angle or torsion of a dihedral angle, often termed "bonded" terms), with the electrostatic field produced by the system and from the action of Van der Waals interactions (these latter two often grouped as "non-bonded" terms).

$$U_{tot} = (U_{angles} + U_{bonds} + U_{dihedrals}) + (U_{VdW} + U_{electrostatic}) \quad (2.47)$$

2.4.1.1 Bonded terms

The bonded potential terms involve 2-, 3-, and 4-body interactions of covalently bonded atoms. They are represented in Figure 2.3 and, starting from the top left and turning clockwise, the following terms are depicted: stretching of a bond, torsion of a dihedral angle, out-of-plane torsion of an improper angle and bending of an angle.

The 2-body potential is represented by the variations in energy caused by the stretching of covalent bonds and it is generally expressed as an harmonic

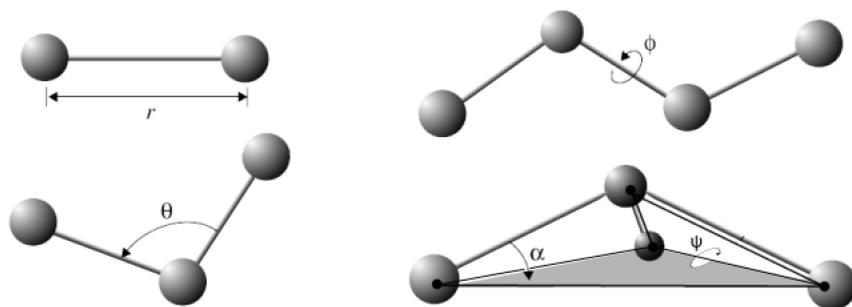


Figure 2.3: Graphic representation of the most commonly used parameters to model the internal geometric variations of a molecule.

potential:

$$U_{bond} = k_b(r_{ij} - r_0)^2 \quad (2.48)$$

where $r_{ij} = \|\vec{r}_j - \vec{r}_i\|$ gives the distance between the atoms, r_0 is the equilibrium distance, and k_b is the force constant of the harmonic spring.

The 3-body potential describes the variations of molecular angles with respect to equilibrium positions and, as previously, it is usually expressed with an harmonic function:

$$U_{angle} = k_\theta(\theta - \theta_0)^2 \quad (2.49)$$

where, in the first term, θ is the angle between vectors $\vec{r}_{ij} = \vec{r}_j - \vec{r}_i$ and $\vec{r}_{kj} = \vec{r}_j - \vec{r}_k$, θ_0 is the equilibrium angle, and k_θ is the angle constant.

Sometimes, a second term is added to Equation 2.49, namely the Urey-Bradley term:

$$U_{Urey-Bradley} = k_{ub}(r_{ik} - r_{ub})^2 \quad (2.50)$$

used to describe a (non-covalent) spring between the outer i and k atoms where, as for spring bond, $r_{ik} = \|\vec{r}_k - \vec{r}_i\|$ gives the distance between the pair of atoms and r_{ub} is the equilibrium distance. Such term permits to express coupling between bending and stretching of a group of atoms.

At last, the 4-body potential includes energy variations caused by proper and improper dihedral torsional terms. The term "improper" refers to dihedral angles produced by the intersection of the planes defined by the atoms connected to a common centre and whose equilibrium value is zero (such as the improper dihedral of the nitrate group NO_3^-). The role of this latter term is therefore that to preserve the planarity of the group of atoms under consideration.

At variance with 2- and 3-body potentials, for which the harmonic form largely superseded all the other existing formulas, several analytical forms for expressing 4-body potentials with a comparable diffusion could be examined. One of them, used for example in the well-known Amber[95] force field, uses a truncated Fourier series:

$$U_{\text{tors}} = \begin{cases} k(1 + \cos(n\psi + \phi)) & \text{if } n > 0, \\ k(\psi - \phi)^2 & \text{if } n = 0, \end{cases}$$

where ψ is the angle in radians between the (i, j, k) -plane and the (j, k, l) -plane. The integer constant n is non-negative and indicates the periodicity. For $n > 0$, ϕ is the phase shift angle and k is the multiplicative constant. For $n = 0$, ϕ acts as an equilibrium angle and the units of k change to $\text{potential}/\text{rad}^2$. A given (i, j, k, l) -quadruple of atoms might contribute multiple terms to the potential, each with its own parameterization, thus allowing for complex angular variation of the potential.

For all the examples just described, k coefficients often depend on the type of geometrical descriptor in order to quantify their different stiffness.

2.4.1.2 Non-bonded terms: van der Waals forces

Non-bonded atomic interactions, *i.e.* atomic interactions that are not depending on a physical connection as a bond, an angle or a dihedral angle, are not limited to the largely known electrostatic forces. A number of intermolecular forces of different kind are available, though they often have a smaller strength than the electrostatic ones.

These interatomic forces can be divided into four categories:

- Electrostatic interactions between two permanent multipoles. This electrostatic interaction is also quoted as the Keesom interaction.
- The attractive interaction between a permanent multipole on a molecule with an induced multipole on another. This interaction is reported as Debye force.
- The attractive interaction between the instantaneous multipoles produced by the motion of electrons. This interaction is the dispersion or London interaction, able to explain the condensation of noble gases.

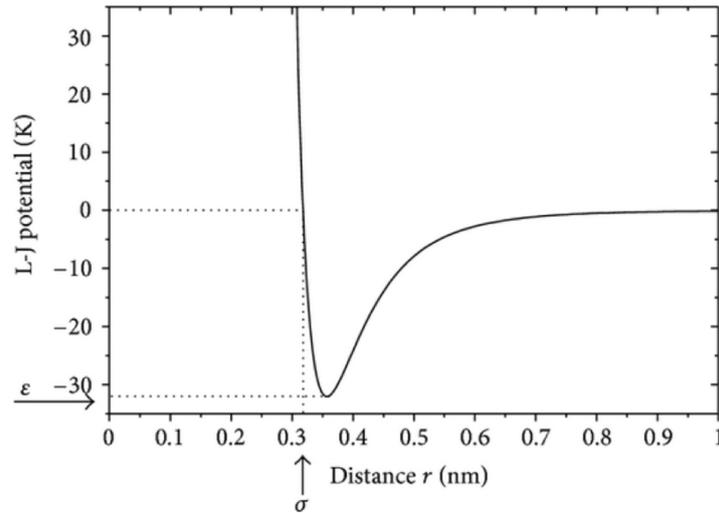


Figure 2.4: The Lennard-Jones potential.

- The Pauli's repulsion, a repulsive component resulting from the Pauli's exclusion principle preventing the superposition of electrons having the same spin.

Such interactions can be correctly explained only by taking into account the electronic structure of each atom and in some circumstance also the correlation of electronic motion, two features that classical Molecular Dynamics neglects by definition. For this reason, an approximate potential must be employed.

Experimentally, it has been observed that these forces tend to repel particles that are too close among them (with the force getting more and more intense as the particles approach) and to attract particles placed at a certain distance (relatively close but not too far).

The potential used to replicate this effect must therefore be strongly repulsive at short distances (to replicate the Pauli's forces) and must progressively aim at zero by increasing the distance among the particles (as already seen, the Van der Waals forces become imperceptible at long distances). In addition, negative values (attractive behaviour) must be present at an intermediate range.

The most used mathematical model to describe these intermolecular forces is the Lennard-Jones (LJ) potential, graphically depicted in Figure 2.4.

2.4.1.2.1 Lennard-Jones potential The expression of the Lennard-Jones potential is:

$$U_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] = \epsilon \left[\left(\frac{r_m}{r} \right)^{12} - 2 \left(\frac{r_m}{r} \right)^6 \right] \quad (2.51)$$

where r is the distance between the particles, r_m is the value of the distance corresponding to the minimum of the potential function, ϵ is the minimum absolute value of this potential (commonly known as depth of the potential well) and σ is the distance where the potential is zero, At r_m , the potential function has the value $-\epsilon$. These parameters can be fitted to reproduce experimental data or accurate quantum chemistry calculations.

The r^{-12} term represents the repulsive part of the potential, while the r^{-6} term describes the long range attraction. The great simplicity of the LJ potential, combined with its acceptable accuracy, is the reason of its wide diffusion, for instance with respect to the Buckingham potential [96], which has the disadvantage of containing an exponential term whose evaluation is computationally expensive.

An alternative way for expressing the LJ potential reported in Equation 2.51 is the "A-B expression":

$$U_{LJ}(r) = \frac{A}{r^{12}} - \frac{B}{r^6} \quad (2.52)$$

where, $A = 4\epsilon\sigma^{12}$ and $B = 4\epsilon\sigma^6$. Conversely, $\sigma = \sqrt[6]{\frac{A}{B}}$ and $\epsilon = \frac{B^2}{4A}$.

2.4.1.3 Non-bonded terms: electrostatic potential

The electrostatic forces are the most important component for the description of the atomic and molecular interactions of a system. In classical Molecular Dynamics simulations, the electrostatic forces of the system are usually calculated by considering each atom as a point charge, whose magnitude is calculated at the very beginning of the simulation and kept constant through the whole calculation. The electrostatic energy is determined according to the Coulomb's formula, having been described as the energy of a system with N point charges:

$$U_E(r) = \frac{1}{2}k_e \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{r_{ij}}, \quad (2.53)$$

where $k_e = \frac{1}{4\pi\epsilon_0}$ is the Coulomb's constant and r_{ij} is the distance between the point charges q_i and q_j . The screening effect of a medium can be included in the calculation by adding the relative electric permittivity ϵ_r to the denominator of the Coulomb's constant expression.

There are several numerical methods used to accomplish these calculations. Starting from the classic implementation of the Coulomb's equation, these methods are greatly affected by the need to streamline the computational times of the

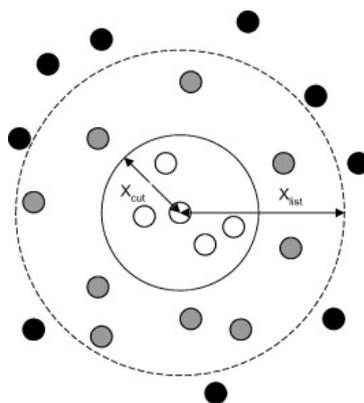


Figure 2.5: Graphic representation of the Verlet list method.

software.

Among the other computational procedures, two techniques have represented a great innovation to calculate these quantities as precisely and as fast as possible. These methods are the Verlet lists and Ewald summation.

2.4.2 Verlet lists

The Verlet list is a collection of data used to reduce the computational times required by a MD software without decreasing its accuracy. It is related to the cutoff algorithm that very simply affirms that pairwise long-range interactions have to be calculated only for the pairs of particles whose distance is smaller than a cutoff distance r_{cutoff} . So, beyond that distance, any interaction with the other particles is considered to be negligible and consequently ignored. In practice, this approach is implemented by creating a list of pairs of particles whose distance is shorter than the cutoff radius.

Obviously, this list cannot be updated at each step of the simulation, because in that instance it would lose all its usefulness, due to the need of evaluating all the $N \cdot (N - 1)$ distances as if this data structure would be absent. For this reason, the list is updated according to different criteria, specific to each software, but at any rate each two or more time-steps. In addition, for avoiding to neglect the interactions among particles that could decrease their distance under r_{cutoff} before the main list is updated, this algorithm also provides a small additional part, consisting of pairs whose distance is greater than the cutoff radius, but less than a second radius, the Verlet radius (often around 1 or 2 Å more than r_{cutoff}). This list of pairs of particles is the Verlet skin used by the algorithm to update the list faster (acting therefore like a cache memory). Actually, the distances of the particles in the Verlet skin are checked at each time-step to determine if

some of them have reduced in a way that makes necessary to include them in the main Verlet list, which is instead updated less often.

2.4.3 Ewald summation

The most intuitive way for evaluating the electrostatic interactions in a box for a Molecular Dynamics simulation is represented by the direct sum approach, namely to simply sum the contributions coming from the Coulomb's law ($\vec{F}_{ij} = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q_i q_j}{|\vec{r}_{ij}|^3} \vec{r}_{ij}$) up to a certain cutoff value. This approach can be used (such as done for the Verlet lists) both with or without periodic boundary conditions, which are however almost always employed due to the much better description they are able to provide for the vast majority of systems. The drawback of this extremely simple algorithm is that the average truncated potential energy (*i.e.* the potential energy ignored during the calculation) does not absolutely converge when increasing the distance (therefore the result depends on the way the summation is performed) and a very large value for the cutoff radius is required in order to get a low truncation error, consequently producing an increase of the simulation time.

In order to solve these problems, the Ewald summation technique[97] was proposed in 1921. The method was originally developed for the application to crystalline systems but, as soon as Molecular Dynamics was further developed, it started to be used also in this field. The success of the method was so large that it is still employed nowadays or, anyway, it represents the forefather of the currently employed techniques as the PME[98], SPME[99] or PPPM[100] approaches.

The main idea behind the Ewald summation algorithm is to consider that each particle i of the system (with charge q_i) is surrounded by two diffuse charge distributions of Gaussian shape, with overall charge magnitudes of $|q_i|$ and, respectively, the opposite and the same sign. The first set of charges is the "screening" distribution, the second one is the "compensating" distribution and, with this procedure, the system's total charge remains overall untouched.

The charge decomposition just outlined, depicted in Figure 2.6 produces therefore for each particle i , three contributions for the electrostatic potential. The first one is due to the point charge q_i , then the next one is due to the Gaussian screening charge cloud with charge $-q_i$ and, at last, the third one is due to the compensating charge cloud with charge q_i .

As affirmed in Figure 2.6, the electrostatic potential produced by the sets of

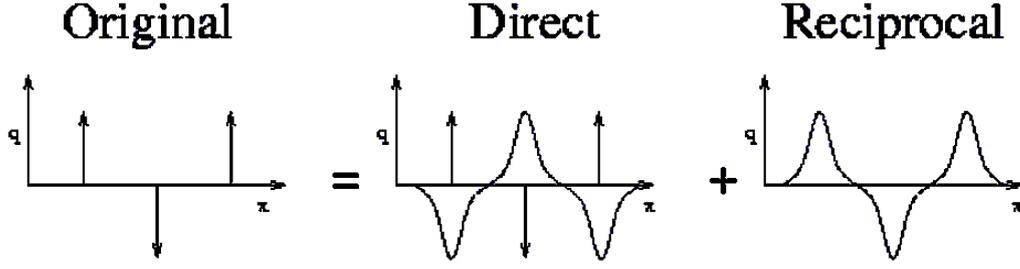


Figure 2.6: Ewald charge decomposition.

point charges and screening densities can be easily computed by direct summation using a cutoff because the electrostatic potential due to a screened charge is a rapidly decaying function of r . On the contrary, the interaction energy of the original point charges with the latter set of compensating diffuse charges requires the electrostatic potential of the diffuse distribution to be evaluated in the inverse Fourier space, where it quickly converges.

The total electrostatic potential energy is therefore calculated as:

$$U_{el} = U_R + U_F - U_S \quad (2.54)$$

where the first two terms U_R and U_F are the electrostatic potential energies generated by the interactions with the two sets of charges just described. The latter term U_S is a small and computationally cheap corrective moiety needed for canceling the spurious terms produced by the previous sum with the sake of keeping the algorithm computationally efficient.

2.4.3.1 Real space term

The term U_R is easily evaluated in the real space by a truncated direct sum approach because, as written elsewhere in the Thesis, the charge compensation makes the sum for calculating the electrostatic potential quickly and absolutely converging.

The potential energy of this interaction is therefore given by:

$$U_R = \sum_{j>i}^{i=N} \frac{q_i q_j \operatorname{erfc}(\alpha |\vec{r}_{ij}|)}{|\vec{r}_{ij}|} \quad (2.55)$$

where the parameter α is connected to the width of the employed Gaussian distributions (considered to be $\propto e^{-(\alpha x)^2}$) and only the pairs of atoms whose distance is smaller than the cutoff radius are considered.

2.4.3.2 Fourier space term

By applying the Poisson's equation $-\nabla^2\phi(\vec{r}) = 4\pi\rho(\vec{r})$, connecting a charge distribution $\rho(\vec{r})$ to its generated electrostatic potential, $\phi(\vec{r})$ it is possible to express the electrostatic potential energy due to the compensating set of charges as:

$$U_F = \frac{1}{2V} \sum_{\vec{k} \neq 0} \frac{4\pi}{|\vec{k}|^2} |\rho(\vec{k})|^2 e^{-\frac{|\vec{k}|^2}{4\alpha^2}} \quad (2.56)$$

where

$$\rho(\vec{k}) = \sum_{j=1}^N q_j e^{i\vec{k} \cdot \vec{r}_j} \quad (2.57)$$

2.4.3.3 Self-interaction correction

This last term is necessary because the Fourier part of the computation just described includes, for the sake of efficiency, the spurious interaction between a point charge and the same-sign, compensating diffuse distribution surrounding it. To correct it, the following correcting term, evaluated in real space, has to be calculated and subtracted to the terms reported in Equation 2.55 and Equation 2.56.

$$U_S = \sqrt{\frac{\alpha^2}{\pi}} \sum_{i=1}^N q_i^2 \quad (2.58)$$

Additional corrective terms could be present in order to account for the so-called "bonded" interactions, namely the interactions between atoms separated by no more than four bonds. The vast majority of force fields, in fact, partially or totally include these interactions in the bonded terms outlined in subsection 2.4.1.1.

2.4.3.4 Derived methods

As previously pointed out, Ewald summation technique is the basis for the development of more advanced protocols with better computational scaling. If, in fact, a fully optimized Ewald summation scales as $O(N^{\frac{3}{2}})$ (but it can reach, in the worst circumstances, $O(N^2)$) with respect to the number of particles N of the system, these new techniques permits to reach an $O(N \log(N))$ scaling. All of them have in common the use of a mesh on which the atomic

charges are decomposed, therefore allowing to have a discretized Poisson equation (in lieu of that shown in subsection 2.4.3.2), whose solution is much faster than for the undiscretized version, thanks to the use of the Fast Fourier Transform (FFT)[101] algorithm and of the many highly optimized numerical libraries[102, 103] developed for this task, mainly originating from the signal engineering field.

2.4.4 Calculation of partial charges

The description of the electrostatic interactions plays an important role in the calculation of the energies in classical Molecular Dynamics, especially in the presence of polarizable molecules. A detailed description of the electrostatic properties of the molecules is therefore necessary to get reliable results during a simulation. There are several methods to calculate these values, based on *ab initio* or semiempirical calculations. Starting from a known molecular geometry, the first step is to achieve optimization of the wavefunction (and also of the molecular geometry), using usually the Density Functional Theory.

Once the electronic density $\rho(\vec{r})$ of the system (*i.e.*, the integral of the wavefunction $\Psi(\vec{r}_1, \dots, \vec{r}_N, \vec{s}_1, \dots, \vec{s}_N)$ over all but three space coordinates) is obtained, a large number of information can be extracted, including the atomic point charges. The most employed method for this aim is RESP[104], straightforwardly derived from the ESP[105] approach.

2.4.4.1 ESP and RESP charges

The intermolecular electrostatic interactions can be thought in terms of interactions between the electrostatic potential generated by a molecule (or a fraction of it) and one charge of another. The general way to get the partial atomic charges is to choose a set of parameters for generating a molecular electrostatic potential as similar as possible to that calculated using the quantum mechanical wavefunction. The electrostatic potential in a generic position \vec{r} is generated by the sum of the charge contribution of the nuclei and of the electronic density $\rho(\vec{\xi})$:

$$V(\vec{r}) = \sum_{i=1}^N \frac{Z_i}{|\vec{r} - \vec{r}_i|} - \int \frac{|\rho(\vec{\xi})|}{|\vec{r} - \vec{\xi}|} d\vec{\xi} \quad (2.59)$$

where N is the number of the atomic nuclei of the system.

Once the electronic density is obtained, the potential is calculated on a grid of points centred on each nucleus, at a distance ranging from just outside the Van der Waals radius[106] to the double of the distance, typically sampling a few hundred points per atom. The partial charges are therefore determined as the best parameters to reproduce the potential on the grid points.

By considering V_2 (reported in Equation 2.60) as the electrostatic potential due to the set of charges q_j ,

$$V_2(\vec{r}) = \sum_{j=1}^N \frac{q_j}{|\vec{r}_j - \vec{r}|} \quad (2.60)$$

it is possible to define the mean squared deviation χ^2 between Equation 2.59 and Equation 2.60 as:

$$\chi^2 = \sum_{i \in \substack{\text{grid} \\ \text{points}}} (V(\vec{r}_i) - V_2(\vec{r}_i)) \quad (2.61)$$

The charges that minimize the mean squared deviation function under the constraint of molecular charge conservation are the ESP charges.

RESP (Restrained Electrostatic Potential) charges are obtained by following the same method but inserting a restraint in the fitting procedure. The function to be minimized for obtaining the RESP charges is:

$$\chi^2 = \sum_{i \in \substack{\text{grid} \\ \text{points}}} (V(\vec{r}_i) - V_2(\vec{r}_i)) + a \sum_{j=1}^N \left(\sqrt{q_j^2 + b^2} - b \right) \quad (2.62)$$

where a and b are the typical parameters of this method. The restraint is needed in order to prevent that the atoms deep buried in the molecular structure (let's only think about the central carbon atom of a neo-pentane), whose influence on the left-hand summatory is small due to their distance from the sampling points (that are usually located outside the molecule), could assume unphysically large values that minimize χ^2 but do not reflect the real electronic distribution of the system.

Many different levels of theory can be employed for the wavefunction optimization, as well as several basis sets, meaning that the calculated charges are expected to have a strong dependence on the employed computational level. Moreover, it is very important to emphasize that the calculation of these charges is accomplished for isolated molecules (*in vacuo* calculations), without any interaction with the external environment. The polarization effects due to the presence of surrounding molecules are thus not considered within this fitting proce-

dure, strongly limiting its accuracy when the simulation concerns a condensed phase system.

2.4.4.1.1 QM-MM/RESP charges The most simple way for including polarization effects in the fitting is to generate a set of RESP charges referred to a condensed phase system. To do this, a frame from an equilibrated Molecular Dynamics trajectory[107] is usually employed. Such a choice is preferred with respect to a random configuration because in the former instance the molecules have a chemically meaningful arrangement, namely a significant point in the phase space is employed for the parameterization.

Once defined the system, the target molecule is treated at quantum mechanical level while all the other molecules are described as simple point charges (the same set-up of a QM/MM simulation[108, 109, 110] is therefore employed). Periodic boundary conditions are strongly suggested to be applied. The wavefunction of the first molecule is then optimized in the presence of the external electrical field generated by the surrounding atoms, which are described at classical level. This optimization takes therefore into account the molecular polarization produced by the surrounding environment. The partial charges of the molecule of interest are eventually fitted to the electrostatic potential using the restrained electrostatic potential (RESP) method. It is important to repeat the described process several times in order to average the statistical noise coming from the distribution of the molecules in the classical box surrounding the considered moiety.

2.4.5 Polarizable force fields

Even though QM/MM-RESP approach permits to obtain the atomic charges that better represent a molecule in its real environment, the value of the partial charge on each atom and its position (coincident with the atomic nucleus), remain nonetheless fixed throughout all the simulation. This is still a strongly limiting approximation because it is well known that the charge distribution of a molecule can almost instantly rearrange in response to the external perturbation coming from the surrounding system or even to its own geometrical alterations; such phenomenon is defined as polarization.

A further step towards the accurate description of polarization effects is represented by the use of force fields that go beyond the simple representation of electrostatic interactions as produced by point charges. They are called polarizable force fields.

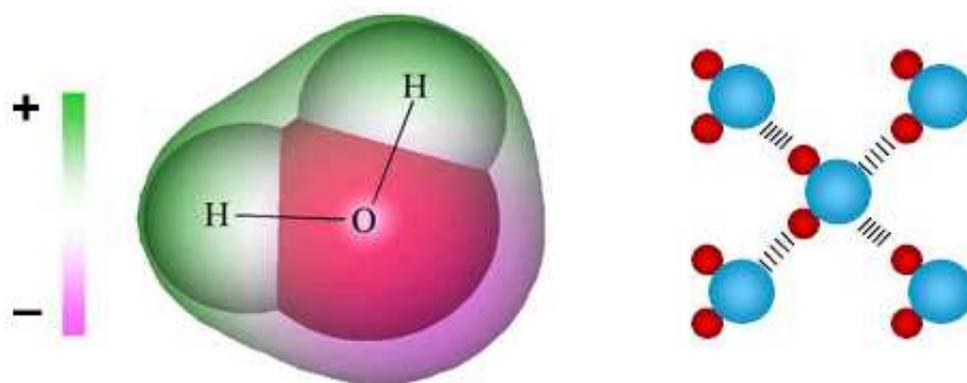


Figure 2.7: Example of molecular polarization for a water molecule.

During the last few decades many polarizable force fields have been proposed [111, 112, 113, 114, 115, 116]. They exploit different ideas to reproduce the effects of atomic polarization during a classical dynamic. Even though the accuracy of the methods may be sometimes very high, all the polarizable force fields proposed so far are characterized by computational times that are around one order of magnitude larger than the classical non-polarizable algorithms.

Currently, all the algorithms reported in literature can be divided into the following three main categories: fluctuating charge models [117, 118, 119], induced dipole moment models [120, 121, 122] and Drude oscillator [123, 124, 125] models.

2.4.5.1 Fluctuating charge model

The force fields belonging to this category are characterized by the fact that the magnitude of the partial point charges localized on each atomic site is not constant during the simulation, but changes in response to the electric field generated by the environment and the structural variations of the molecule.

These fluctuations are evaluated by solving a set of self-consistent calculations, stemming from the fact that the quantities to be calculated (the atomic partial charges) also appear in the expression of the electric field affecting their variation.

The electronegativity equalization principle stated by Sanderson [117] is at the basis of this approach. According to this algorithm, at each step the optimum set of charges is derived from the minimization of the total electrostatic energy

of the system:

$$\begin{aligned}
U = & \sum_i \sum_{\alpha} \left[\chi_{i\alpha}^0 q_{i\alpha} + \frac{1}{2} J_{i\alpha i\alpha}^0 q_{i\alpha}^2 \right] \\
& + \sum_i \sum_{\alpha < \beta} J_{i\alpha i\beta} (r_{i\alpha i\beta}) q_{i\alpha} q_{i\beta} \\
& + \sum_{i < j} \sum_{\alpha < \beta} J_{i\alpha j\beta} (r_{i\alpha j\beta}) q_{i\alpha} q_{j\beta}
\end{aligned} \tag{2.63}$$

where i and j stand for molecules, α and β represent atoms, $\chi_{i\alpha}^0$ and $J_{i\alpha i\alpha}^0$ are the Mulliken electronegativity[126] and the absolute hardness of atom α , and $J_{i\alpha j\beta}$ accounts for the Coulomb interaction between atoms α and β .

In addition, constraints and restraints can also be imposed to preserve the initial net charge of each molecule, thus preventing the unphysical long-range charge transfer process between different molecules that could be produced using this approach.

A drawback of this method is represented by the impossibility of representing an out-of-plane polarization of planar molecules (for example, those containing aromatic rings) without introducing additional virtual charge sites.

2.4.5.2 Induced dipole model

This model uses point polarizable dipoles (or even higher order multipoles) placed on the atomic nuclei along with the partial fixed charges. The placement of these dipoles in different positions, such as half bond points, is also executable. The theory of the linear response approximation states that the induced dipole moment at site i , $\vec{\mu}_i$, is proportional to the local electric field acting on the same site, \vec{E}_i . The proportionality constant is the polarizability tensor, α_i^0 .

$$\vec{\mu}_i = \alpha_i^0 \cdot \vec{E}_i. \tag{2.64}$$

Within this approach, the total electric field can be expressed as the sum of the permanent electric field \vec{E}^0 (generated by the fixed point charges of the atoms of the system) and of the electric field generated by the induced dipoles:

$$\vec{E}_i = \vec{E}_i^0 - \sum_{j \neq 0} T_{ij} \vec{\mu}_j \tag{2.65}$$

where T_{ij} is the ij -th element of the dipole field tensor[120].

A possible approach for solving this problem is the "explicitly interacting distributed polarizability". In this model, the Equation 2.65 and the equation:

$$U_{ind} = -\frac{1}{2} \sum_i \vec{\mu}_i \cdot \vec{E}_i^0 = -\frac{1}{2} \sum_i \alpha_i^0 \vec{E}_i \cdot \vec{E}_i^0 \quad (2.66)$$

are iteratively solved until self-consistency in the induced dipoles is achieved[121].

The most widely diffused polarizable force field, AMOEBA[122], is based on this approach.

2.4.5.3 Drude oscillator model

This model, also known as the core-shell model[123], is based on the introduction of massless particles connected to each atom of the system. These particles possess a partial charge of magnitude q_d , and are attached to the atom through an harmonic spring characterized by a force constant k_d . The net atomic charge of the pair is preserved by assigning to the atomic site a core charge of $(q - q_d)$. The pair between the atom and its associated Drude particle forms the Drude oscillator.

The presence of an external electric field will displace the Drude particle from its equilibrium position (coincident with the nucleus whom it is connected), therefore producing an induced dipole.

The potential energy of the Drude oscillator is composed by:

- a self energy term, coming from the harmonic spring
- the electrostatic potential energy due to interaction of the atomic charge $q_i - q_D$ and of the Drude charge q_D with the external electric field

By employing Hooke's law ($\vec{F} = -k(\vec{x} - \vec{x}_0)$), the induced dipole can be expressed as:

$$\vec{\mu}_i = \frac{q_{D_i}^2}{k_{D_i}} \vec{E}_i. \quad (2.67)$$

As one can easily see by comparing Equation 2.67 with Equation 2.64, the ratio $\frac{q_{D_i}^2}{k_{D_i}}$ expresses the polarizability of atom i .

At variance with the dipoles employed within the induced dipole model, those defined by the Drude oscillator are real dipoles, namely formed by real particles but, except for this slight theoretical difference, both the models provide an almost equivalent description. An advantage of Drude particle model over the induced dipole one is represented by the fact that the dynamics of the

Drude particles can be managed by the same routines employed for the other particles of the system, thus making the former method easier to implement in the simulation packages.

2.4.6 Ensembles

The classical Molecular Dynamics simulations are not only limited to the trivial integration of Newton's equations of motion. In many instances it can be necessary to keep temperature, pressure as well as other state variables of the system at constant values, in order to reproduce the desired experimental conditions. The most used simulation packages provide several methods for controlling temperature and pressure, said thermostats and barostats. Depending on what state variables (for example the energy E , volume V , temperature T , pressure P , number of particles N or even chemical potential μ) are kept fixed, different statistical ensembles are generated. A variety of structural, energetic and dynamic properties can in fact be calculated in connection to the averages or the fluctuations of these quantities over the generated ensembles.

2.4.6.1 NVE ensemble

The ensemble in which the particles' number, volume and energy are constant (NVE) is said microcanonical ensemble and it is recovered by solving the Newton's equations without any temperature and pressure control. The total energy is in fact naturally conserved by the employed symplectic integrators[127] and the fluctuations, due to numerical issues, should disappear as the time-step value tends to zero. This ensemble is unrealistic, being impossible to experimentally measure (and consequently monitor) the energy of a system undergoing a chemical experiment or process. However, it is sometime employed for performing productive simulations, being absent the artifacts inevitably introduced in the dynamics by the use of thermostats and/or barostats.

2.4.6.2 NVT Ensemble

The constant temperature, constant volume (and constant number of particles) ensemble (NVT), also referred as the canonical ensemble, is the default ensemble provided in almost every simulation software. As previously asserted, the temperature is kept constant by the use of algorithms said thermostats that perform direct scaling of the particle velocities or coupling with a constant-temperature external bath. This is the appropriate choice for simulations of free

molecules (computationally speaking, *in vacuo* system experimentally corresponding to gas phase molecules) without periodic boundary conditions, though in that instance it is not possible to correctly use the concept of volume. Even using periodic boundary conditions, if the pressure is not a crucial factor to be monitored, the constant-temperature /constant-volume ensemble provides the advantage of a smaller perturbation of the trajectory, owing to the absence of coupling with a pressure bath. For this reason, this ensemble is often employed for the productive simulation of condensed phase systems, previously equilibrated at the desired temperature in the NPT ensemble.

2.4.6.3 NPT ensemble

The constant temperature, constant pressure ensemble (NPT) allows control over both temperature and pressure. The unit cell vectors are allowed to change, and the pressure is adjusted by varying the volume in a way completely analogous to the velocity scaling performed by thermostats. This analogy is so strong that, in some circumstance, the same algorithm is disposable for both temperature and pressure control (*e.g* Nosé-Hoover[128, 129]). This is the ensemble to be chosen when pressure is of basic importance for the simulation but, as already asserted, the coupling with the external pressure bath can produce artifacts even when the most sophisticated algorithms are employed. For this reason, its employ is often limited to the equilibration phase, while no data for the final elaboration are collected.

2.4.6.4 μ VT ensemble

The constant temperature, constant volume and constant chemical potential ensemble (μ VT), also termed grand-canonical ensemble, permits to perform the most accurate simulation of a typical chemical system, that is continuously in contact with the surroundings (let's think, for example, to a chemical laboratory tube). This kind of ensemble is very useful when an adsorption process has to be studied because the condensed phase of the compound undergoing adsorption is simulated by allowing it to exchange particles with a reservoir of the same compound in the gas phase. Exchange (or removal) of particles is supervised by random numbers generator, in a fashion resembling Montecarlo simulations.

2.4.7 Trajectory analysis

The sequence of frames obtained at the end of a Molecular Dynamics calculation is a real trajectory in time, at variance with the configurations obtained from Monte Carlo (MC) simulation, that simply are the result of the Markovian process produced by the Metropolis algorithm mentioned in chapter 1. This makes possible to perform analysis focusing time-dependent properties.

The number of the potential investigations is huge and continuously growing, going from the most simple to the highly sophisticated ones, such as the calculation of the infrared spectrum of the simulated system by employing the dipole moment autocorrelation function[130] or the decomposition of the simulation box into Voronoi domains.

In the following section, two rather simple analytical techniques, that is the calculation of the Radial Distribution Function (RDF) and the determination of the diffusion coefficient D , are briefly outlined due to the consideration they will be given in section 4.5.

2.4.7.1 Radial Distribution Function

The Radial Distribution Function (RDF), often denoted as $g(r)$, expresses the probability of finding a particle at a given distance from another particle. It is a very useful quantity providing an immediate qualitative understanding of the structure of a system. For an ideal gas, in fact, the RDF has constant unitary value while on the opposite side, for a crystal, it is an infinite series of sharp peaks (with null width in absence of thermal excitation) separated by deep valleys where the function has zero value.

The expression[131] for $g_{ij}(r)$ coming from statistical mechanics in the canonical ensemble NVT is:

$$g_{ij}(r) = V^2 \left(1 - \frac{\delta_{ij}}{N_j}\right) \int \frac{e^{-\frac{\Phi(d\vec{r}_1, \dots, d\vec{r}_n)}{kT}} d\vec{r}_3, \dots, d\vec{r}_n}{Z(T, V, N)} \quad (2.68)$$

where δ_{ij} is the Kroenecker delta, $\Phi(d\vec{r}_1, \dots, d\vec{r}_n)$ is the potential energy of the system and $Z(T, V, N)$ is its partition function.

This expression would however require, for being analytically solved, the exact knowledge of the system potential and it is therefore useless except for the most trivial cases. For this reason, the function is commonly calculated from a Molecular Dynamics trajectory (or a Monte Carlo sequence of configurations)

as follows:

$$g_{ij}(r) = \frac{1}{N_{snapshots}} \sum_{m=1}^{N_{snapshots}} \frac{1}{N_i} \sum_{n=1}^{N_i} \frac{dn_r^{i,j,m}}{4\pi r^2 dr \cdot \rho^{bulk}} \quad (2.69)$$

where $dn_r^{i,j,m}$ is the number of particles j contained in a shell of thickness dr and distance r centred around a particle i of the snapshot m and ρ is the average (bulk) density of the considered system.

An interesting property of the Radial Distribution Function is that it permits to connect the bulk to the local density. The latter can in fact be calculated as:

$$\rho(r) = \rho^{bulk} g(r) \quad (2.70)$$

By reverting the previous equation, it can also be easily understood why, when r increases, the Radial Distribution Function tends asymptotically to one for the fluid (or, in general, not crystalline) systems. The size of the spherical shell where local density is calculated increases quickly ($V \propto r^3$), averaging thus the densities of the local environments towards the bulk density of the system.

2.4.7.2 Diffusion coefficient

The first Fick's law [132] gives an expression for the flux, namely the quantity of matter that crosses a surface unit in a time unit, in connection with the spatial gradient of the system's density:

$$\vec{J} = -D \nabla_{\vec{r}} \rho \quad (2.71)$$

where D is the diffusion coefficient acting as proportionality coefficient between the former quantities and the minus sign expresses, as intuitive, that the flux goes towards the less dense zones.

During their studies on Brownian motions, Einstein [133] and Von Smoluchowski [134] were independently able to demonstrate that the diffusion coefficient can be calculated from the mean squared displacement of the particles of the system, namely:

$$MSD(t) = \frac{1}{N_{particles}} \sum_{i=1}^{N_{particles}} [\vec{r}_i(t) - \vec{r}_i(0)]^2 \quad (2.72)$$

This expression can easily be recovered from a Molecular Dynamics trajectory (but not from the output of a Monte Carlo simulation).

The equation named after them has, in three dimensions, the form:

$$D = \lim_{t \rightarrow +\infty} \frac{MSD(t)}{6t} \quad (2.73)$$

In practice, the Fick's diffusion coefficient is directly proportional to the slope of the MSD vs t plot.

Attention has to be paid to the meaning of $+\infty$. As it is not possible to perform an infinite simulation, the meaning of the limit is to wait a sufficient time before starting to store the data for the calculation of the slope, in order for the system to "lose memory" of the starting configuration and to perform a true random-walk motion. The interval of time that has to be waited crucially depends on the studied system. For the slowly diffusing ionic liquids, tenths of ns have shown to be necessary in order to achieve diffusive motion[135].

A possible check suggested by Maginn [135] is to compute the quantity:

$$\beta(t) = \frac{\partial \log(MSD(t))}{\partial \log(t)} \quad (2.74)$$

Until $\beta(t) < 1$, a sub-diffusive regime is still present and a longer equilibration is needed.

Chapter 3

Technical details

3.1 EBS

The DFT calculations have been performed using the ORCA[86] package (in its 4.0.1, 4.1.2 and 4.2.1 versions) and the all-electron Ahlrichs' def2-TZVP[136] basis set has been employed.

The OPBE[137, 138], BP[44, 139], B3LYP[140, 45] and M06[141] functionals have been chosen to test the two main families of DFT functionals (respectively GGA and hybrid, whose characteristics have been outlined in subsection 2.1.1). SCS-MP2[142, 19] and B2PLYP[48] calculations have also been performed for one of the clusters, with the aim of respectively testing a functional-independent approach and one of the highly accurate recently developed double-hybrid functionals[143].

All the calculations have been speeded up by applying the Resolution of Identity[144, 145] methods, as implemented in ORCA. These methods required additional basis sets, namely def2-TZVP/c[146] and def2/j[147]; `Gridx4` precision level was imposed for the auxiliary integration grid.

The criterion for the SCF convergence has been set as `TightSCF` along with a `Grid5` precision level for the integration grids. `TightOpt` keyword (tighter than the default) was used for performing the geometrical optimizations preceding normal mode analysis (described in subsection 2.2.1), to subsequently calculate the vibrational frequencies over accurately optimized structures. Numerical evaluation of the Hessian matrix (using the keyword `NumFreq`) was sometimes required due to the unavailability of the corresponding analytical expression in the current ORCA package. In these cases, atomic displacement for the numerical calculation of the derivative of the force was set to 0.001 *bohr* (smaller than the default value and recommended in the manual of the software for highly

accurate calculations).

The geometrical optimizations within the EBS approach were produced by interfacing the external optimizer feature available in ORCA with an in-house made Fortran[148] code. The computational procedure is sketched in Figure 3.1.

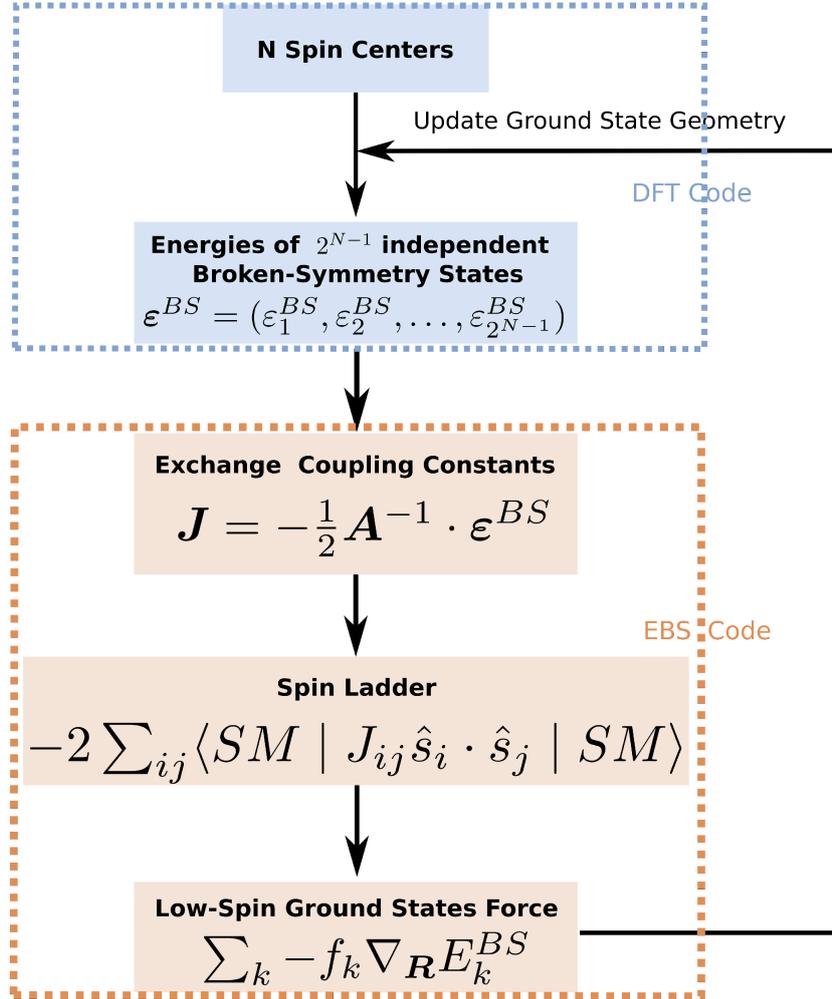


Figure 3.1: Geometrical optimization procedure for a low-spin ground state using the Extended Broken Symmetry method.

The Fortran program we use, along with the Python[149] script needed to automatize the whole optimization process and some commented input examples can be found in github.com/FrancescoCappelluti/EBS.

In order to assure that the desired broken symmetry state is obtained (during both the BS and EBS optimizations), the Mulliken atomic spin populations[126] for the iron atoms have been calculated and checked.

In order to check that Equation 2.37 is correct and above all to quantify the impact of the approximations undertaken for its derivation, the numerical gradient was calculated for the selected case of the $[\text{Fe}_4\text{S}_4(\text{SCH}_3)_4]^{2-}$ cluster. The

contribution of the neglected term was indeed about 1% of the total gradient, as one can see by inspecting the Table reported in subsection 6.1.2. The assumption we did (Equation 2.38) is therefore fully reasonable.

Vibrational frequencies are eventually calculated (within the harmonic approximation) using the `orca_vib` utility from ORCA package. To do this, the previously obtained ground state Hessian has to be embedded into an ORCA-like `.hess` file that can be correctly read by `orca_vib`.

For the complex $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$, in order to compare its theoretical frequencies to the experimental ones (measured using the NRVS[150] technique on ^{57}Fe isotope), the mass of the Fe atom has been set to 56.93539[151] a.m.u. (instead of using the default 55.847 value given by the natural isotopic distribution of iron).

VibAnalysis[152] tool has been employed for rationalizing calculated normal modes. This recently developed tool[153] uses Vibrational Mode Automatic Relevance Determination (VMARD) analysis for performing the decomposition of calculated normal modes into chemically meaningful internal coordinates variations such as bond stretchings, angle bendings or dihedral angles torsions. It is reported to return results comparable to the more traditional Potential Energy Distribution (PED) technique implemented in the VEDA[154] software but, at variance with the latter, it possesses an excellent integration with ORCA package.

3.2 F-RESP

The parameterization has been performed using homemade Python 3[149] codes, which heavily rely on the numerical library Scipy[155]. Due to the complex analytical form possessed by the cost function reported in Equation 4.24, its optimization was performed in a numerical way, employing the SLSQP-[156] or trust-constr-[157] Basin-Hopping[158] global optimizer. The use of the Basin-Hopping algorithm would ensure that a global minimum of the cost function surface is attained, instead of a local one.

Parameterization has been performed, concerning the Ewald summation over cubic boxes (with a 24.87 Å side) containing 512 water molecules. An 8 Å long cutoff was chosen for fictitious electric field calculation and α was set as 0.4375 Å⁻¹.

When Fennell potential was used, instead, due to the longer cutoff required, boxes with shape (43.69, 45.40, 42.81) Å and containing 2880 water molecules

were used. The QM/MM-RESP charges have been calculated with the well known CP2K[159] package at the HF[9]/6-31G*[160, 161] level of theory. A cut-off for plane-waves of 400 Ry was employed along with Goedecker-Teter-Hutter (GTH) pseudopotentials[162].

The parameterization through Machine Learning techniques has been attempted by using Tensorflow[163] and scikit-learn[164] Python packages.

Chapter 4

Results

4.1 Normal Mode Analysis within Extended Broken Symmetry

As previously written, the generalized Extended Broken Symmetry approach was presented for the first time in an article[84] appeared slightly before the beginning of my PhD and containing the work performed during the internship of my Master Thesis. The method permits to obtain a description for the low-spin ground state of multi-centre transition metal complexes (containing an arbitrary number of spin centres) at the DFT level or, anyway, using a single-determinant approach without the spin contamination phenomenon that would affect the results using a Broken Symmetry wavefunction, as generally done. Such a method is much faster and less computationally expensive than the more theoretically rigorous multi-configurational techniques. The greatest innovation introduced by this article is represented by the derivation of the EBS geometrical gradient; the knowledge of this expression permits therefore to perform geometrical optimizations on the correct, spin-uncontaminated ground state Potential Energy Surface.

The algorithm underlying the method, already discussed in section 2.3, can be shortly summarized as:

1. The single-point energies and gradients of all the BS states (including HS) are calculated using an electronic structure method.
2. The expectation values for the product of two one-atom spin operators are recovered from the values of the total spin square and the spin populations of the spin centres, thus permitting to calculate the exchange coupling constants.

3. The HDvV Hamiltonian is constructed in its matrix form and diagonalized, providing the energy spectrum.
4. The total gradient of the ground state with respect to the coordinates of nuclei is obtained as linear combination of Broken Symmetry gradients.
5. Once the force for the ground state is known, the geometry is updated by a first-order minimization algorithm.

The vibrational frequencies for the EBS-optimized ground state structure can be calculated by estimating (and then diagonalizing) the Hessian matrix obtained by applying the following expression:

$$\mathcal{H}_{GS} = \mathcal{H}_{HS} + \sum_{k=1}^{N_{BSstates}} f_k \mathcal{H}_k \quad (4.1)$$

where the coefficients f_k are the same used in the linear combination of the BS gradients leading to the ground state one. As it can be seen, the expression is closely related to Equation 2.43, having been derived directly from it in the approximation of $\nabla f_k = 0$.

It is worthy of remark that in Equation 4.1 (such as in Equation 2.43) the high spin state (that we consider, only for the sake of clarity, to be the first Broken Symmetry state) is also included in the summation. It appears therefore twice, the first time with coefficient 1 and the second time with coefficient c_1 . Such a particular behaviour is due to the use of the high spin state gradient as a "zero point" for the expression of the GS and to the inclusion of the high spin state in Equation 2.28 for the J coefficients determination. The sum of these two coefficients usually turns out to be close to zero. The contribution to the GS gradient is therefore usually the smallest one, being usually HS the state with the highest energy value and the most different from the ground state.

4.2 EBS simulation of Fe-S clusters

In an article published recently[85], I applied the Extended Broken Symmetry method (EBS) for the geometrical optimization and normal mode analysis of ferredoxin-like $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ and cubane-like $\text{Fe}_4\text{S}_4\text{Cl}_4$ clusters. The structures were chosen with the aim of representing two of the main important families[82] of iron-sulfur clusters.

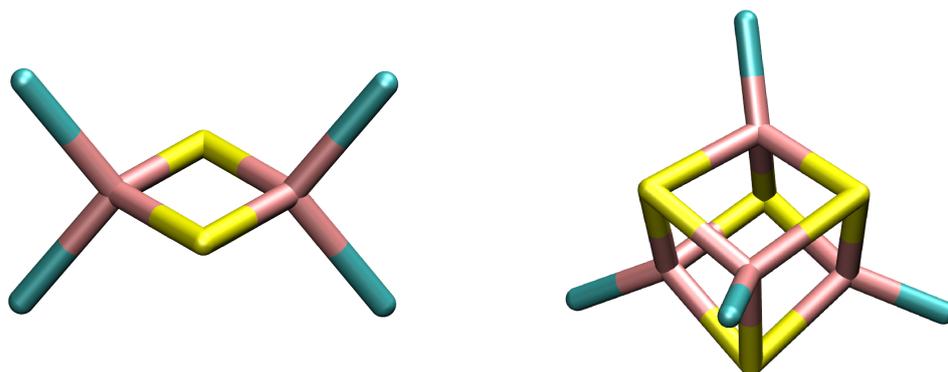
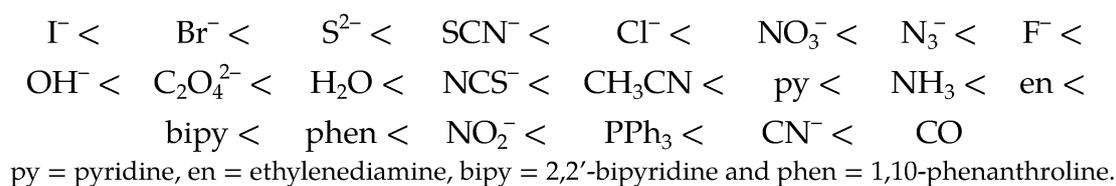


Figure 4.1: The two investigated iron-sulfur clusters: $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ (left) and $\text{Fe}_4\text{S}_4\text{Cl}_4$ (right).

Both of them are characterized by an open-shell low spin ground state, thus containing high spin iron atoms interacting among themselves in an antiferromagnetic fashion and having null total spin (due to the fact that all the iron atoms are in the same oxidation state and thus possess the same valence).

Figure 4.2: Example of spectrochemical series.



The high spin nature of the ferric iron atoms is guaranteed by the very low position that the chloride ligand holds in the spectrochemical series[165] (see Figure 4.2). Such a low position produces a small energy separation between the t_2 and e levels (represented in the left part of Figure 4.3) and the electrons they contain are induced, to maximize the direct exchange interactions, to singly occupy all the available orbitals and assume parallel alignment.

Antiferromagnetic interactions are (instead) experimentally testified by magnetic susceptibility measurements at different temperatures[166]. The positive temperature coefficients of the curve of the magnetic susceptibility with respect to temperature are in fact a clear hint about the presence of this kind of intra-cluster interactions.

The experimental structure of the first cluster, in particular for the crystalline specie $[\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}][\text{N}(\text{Et})_4^+]_2$, is known by X-ray diffraction experiments and this permits to validate our theoretical results.

Concerning, at variance, the normal mode analysis side, the frequency calculations are of particular importance for the largest $\text{Fe}_4\text{S}_4\text{Cl}_4$ cluster, being the first

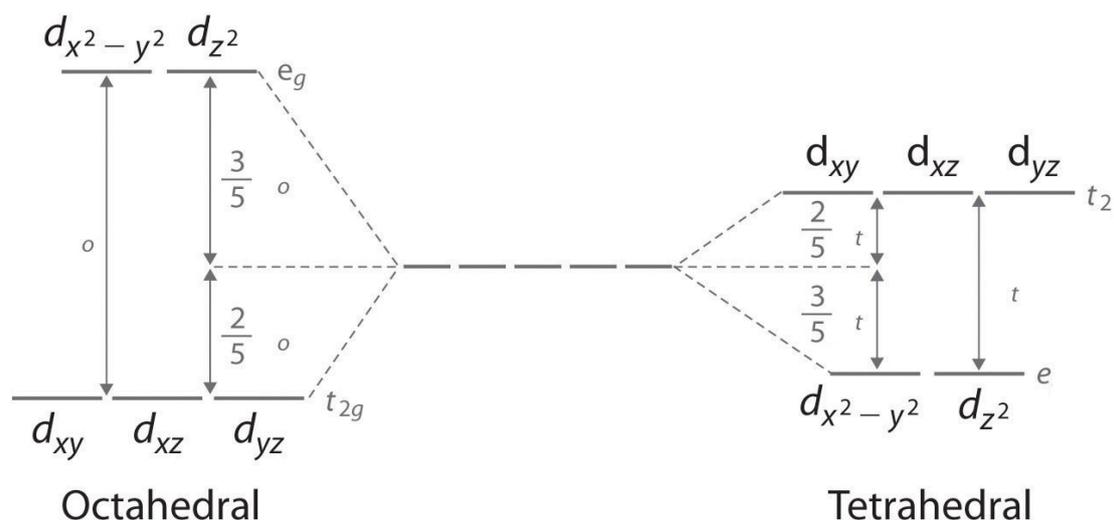


Figure 4.3: Crystal field splitting in tetrahedral and octahedral fields (courtesy of chem.libretexts.org).

time that, thanks to the introduced generalization[85], the vibrational spectrum of a so large compound is simulated within the EBS approach. The vibrational frequencies were in fact calculated within the EBS approach only in the simplest case of a di-iron cluster by Bovi and Guidoni [77].

In addition, for the $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ cluster, experimental vibrational frequencies (measured using NRVS technique) are known, allowing the comparison of our results with the band assignments proposed by Smith *et al.* [167] and to assess the quality of the frequencies calculated with our method.

4.2.1 $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ structure

The structures obtained from the optimizations in the high spin state are very different from the experimental ones and this can be appreciated by inspecting the values of $Fe \cdots Fe$ and $Fe - S$ distances and angles reported in Table 4.1.

The general trend that can be observed ranging from HS to BS and to EBS spin description is that the iron-sulfur cluster get compressed, as one can see by the values of the $Fe \cdots Fe$ distances and $Fe - S$ bond lengths. This behaviour can be interpreted as a consequence of the better theoretical description given to the antiferromagnetic interaction between the two spin centres by the introduction of the static correlation treatment. As a first approximation, the decrease of $Fe - S - Fe$ angle widths would prevent superexchange interactions between the spin centres due to the worse superposition between the half-filled d orbitals of the iron atoms and the completely filled p orbital of the central sulphur atom. In reality, this drawback is probably overcompensated by the contemporary de-

crease of the $Fe - S$ bond lengths, making the nuclei closer and improving the superposition of the orbitals involved in the bonding even in presence of a unfavorable angular arrangement.

In most of the cases (when the hybrid M06 and B3LYP, the double-hybrid B2PLYP functional and the SCS-MP2 method are employed) this compression makes the EBS optimized structure closer to the experimental geometry with respect to the BS and HS ones. In the remaining cases, the considered DFT functionals (namely the GGA-type BP and OPBE) overestimate the coupling between the spin centres, consequently returning optimized structures that are found to be particularly compressed along the $Fe \cdots Fe$ axis.

Table 4.1: Geometrical comparison - $Fe_2S_2Cl_4^{2-}$.

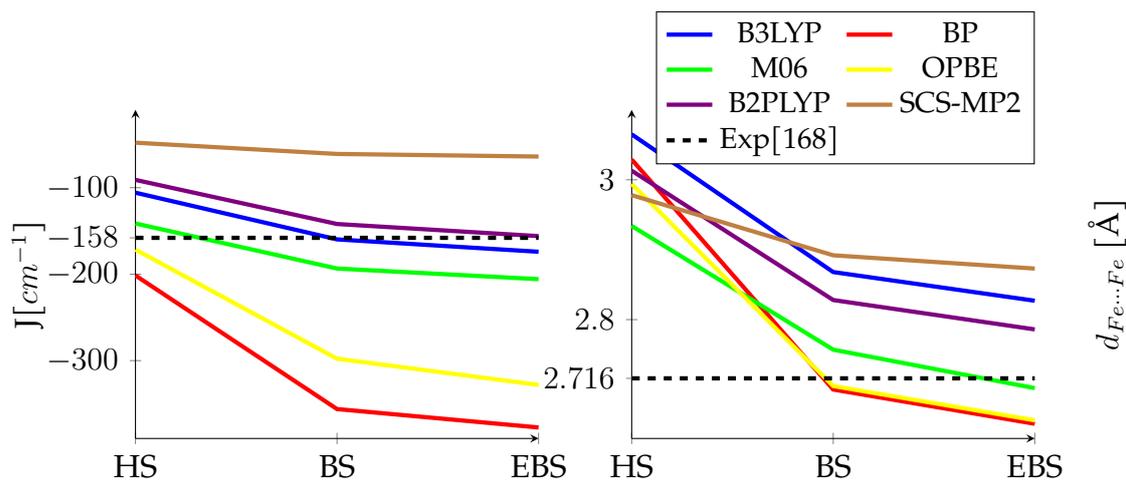
		$Fe \cdots Fe$	$Fe - S$	$Fe - S - Fe$	$S - Fe - S$
Exp[168]		2.716	2.200	76.2	103.8
B3LYP	HS	3.065	2.307	83.2	98.6
	BS	2.869	2.255	79.0	101.0
	EBS	2.827	2.243	78.1	101.9
BP	HS	3.029	2.294	82.6	97.4
	BS	2.700	2.202	75.6	104.5
	EBS	2.651	2.187	74.6	105.4
M06	HS	2.934	2.281	80.1	99.9
	BS	2.758	2.237	76.1	103.9
	EBS	2.702	2.227	75.4	104.6
OPBE	HS	2.994	2.276	82.2	97.8
	BS	2.705	2.196	76.0	104.0
	EBS	2.656	2.178	75.1	104.9
B2PLYP	HS	3.013	2.285	82.5	97.5
	BS	2.828	2.240	78.3	101.7
	EBS	2.786	2.229	77.4	102.6
SCS-MP2	HS	2.978	2.724	81.8	98.2
	BS	2.892	2.254	79.8	100.2
	EBS	2.873	2.249	79.4	100.6

Distances are expressed in Å, angles are expressed in °.

Along with the compression of the structures, a decrease of the exchange coupling constant J (which in a more intuitive view can also be viewed as an increase in the absolute value but keeping the negative sign, connected with antiferromagnetism) can be observed. The antiparallel alignment between the electronic spin moments on the couple of iron atoms becomes therefore more stable. The change of the J constant of the cluster for the structures optimized within the different spin descriptions (HS, BS and EBS) and employing differ-

ent levels of theory is represented in Figure 4.4 along with the change of the highly significant $Fe \cdots Fe$ distance. The exchange coupling constants are more exhaustively listed in Table 6.3 too. It is worth to note that the exchange coupling constants and the distances between the spin centres are deeply connected: the stronger are the magnetic interactions, the smaller are the interatomic distances.

Figure 4.4: J-coupling constant and $Fe \cdots Fe$ distance vs spin description - $Fe_2S_2Cl_4^{2-}$.



As one could expect, due to the presence of a percentage of exact Hartree-Fock exchange in their expressions, the hybrid functionals B3LYP and M06 and the double-hybrid B2PLYP functional provide higher spin populations on the Fe atoms with respect to the GGA ones (with an average difference around $0.8 e^-$). Nevertheless, the formal theoretical value per atom (5 for Fe(III) in its high spin state) was not at any rate achieved. Among the two GGA functionals employed, BP most underestimates the presence of unpaired electrons on iron ions

In accordance with the results of Chu *et al.* [84], the calculated exchange coupling constants are stronger when the GGA functionals are used. As for in that case, the experimentally measured value [168] ($J = -158 \pm 8 \text{ cm}^{-1}$) confirms that the GGA functionals much overestimate the strength of the coupling. Double-hybrid functional B2PLYP, on the contrary, permits to obtain a remarkable agreement with the experiments, by perfectly matching the measured value within the $\pm 8 \text{ cm}^{-1}$ experimental uncertainty. Intermediate performances can eventually be observed for the hybrid functionals B3LYP and M06, because the shifts from the experimental J constant are, in fact, respectively of 20 and 50 cm^{-1} .

A quite interesting observation is that, notwithstanding the perfect agreement concerning the coupling constant value, the B2PLYP functional does not estimate the $Fe \cdots Fe$ distance with the same remarkable accuracy: a discrep-

ancy of 0.07 Å can in fact be observed between the experimental values and the calculated ones. Using the M06 functional, instead, a much better matching (2.702 vs 2.716 Å) can be obtained.

Eventually, the wavefunction-based SCS-MP2 method strongly underestimates the strength of the coupling between the iron atoms (as it can be appreciated by inspecting the narrow J-coupling constant and the huge distance between the spin centres). It is also interesting to observe that, when using this method, the geometrical differences between the HS, BS and EBS states are very limited, as a consequence of the little strength predicted for the exchange interaction between the magnetic nuclei.

It is important to stress that, within the Extended Broken Symmetry formalism, the J-coupling constants and the structural parameters are both computed consistently. The geometry with the minimal energy indeed corresponds to the lowest energy of the same Hamiltonian used to estimate the J-coupling parameters, at variance with other currently employed approaches. The obtained results therefore indicate that a functional which is fully able to estimate with high precision both the geometrical and the magnetic properties is still lacking and even the highly sophisticated double-hybrid B2PLYP one, in fact, unfortunately overestimates in a non-negligible way the distance between the iron spin centres.

As a final point, the results shown for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ are highly consistent with those obtained for the cluster $\text{Fe}_2\text{S}_2(\text{SH})_4^{2-}$ in a previous study[84] carried out in my research group, both with respect to the overestimation of the exchange coupling constant and the underestimation of the iron-iron distance when the functionals of GGA type are adopted. Even the experimental $Fe \cdots Fe$ distances of the two clusters are quite similar (2.716 Å compared to 2.75 Å, being the uncertainty of the latter quantity larger due to the experimental technique employed, namely Mössbauer spectroscopy on a protein prosthetic group). Such an observation therefore demonstrates that, at least for ligands that are so close along the spectrochemical series such as chloride and sulphide, the influence of the side chains on the $[2\text{Fe}2\text{S}]$ core structure is quite reduced.

4.2.2 $\text{Fe}_4\text{S}_4\text{Cl}_4$ structure

At variance with the previous system, any experimental information is available for this cluster, at least in its neutral charge state. The lack is likely connected to its instability. The chloride ligands are in fact not able to stabilize in a suf-

sufficient amount the highly oxidized Fe(III) ions of the core, at variance to what happens for the $\text{Fe}_4\text{S}_4(\text{N}(\text{Si}(\text{CH}_3)_3)_2)_4$ cluster [169] studied by Chu *et al.* [84]. Experimental data are only available for the charge state -2, in which this system has probably mixed-valence spin centres [166] (because of the formal presence of two Fe(II) and two Fe(III) atoms). A hint about the presence of this particular electronic arrangement comes from the D_{2d} point-group symmetry of the core, as observed by X-ray diffraction measures [168]. The presence of two non-equivalent groups of iron atoms would indeed have pushed the compound to assume a lower symmetry point-group. Unfortunately, we could not perform EBS calculations for the doubly negative charge state (to perform the comparison with measures) because of the presence of mixed valence centres and of the consequently arising double-exchange process, whose correct description would require a further extension of the method for the appearance of an additional term in the Heisenberg-Dirac-van Vleck Hamiltonian. This term has the form $\sum_{\substack{i<j, \\ v_i \neq v_j}} \pm B_{ij} \cdot (S_{ij} + 1)$, where the sum runs over the mixed-valence pairs (in which, therefore, the two atoms have a different valence v), B_{ij} is quoted as resonance parameter and S_{ij} is the spin quantum number for the coupled pair of sites ij , and accounts for the energy of delocalization of the itinerant electron of a mixed-valence pair in a spin-coupled system [170, 171, 172, 173].

Some representative structural parameters of the optimized structures are reported in Table 4.2 for the high spin, Broken Symmetry and Extended Broken Symmetry spin descriptions. As observed for the $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ cluster, when moving from the HS to BS and eventually to the EBS description, the $Fe-S$ bonds get compressed irrespective of the DFT functionals employed. In connection with this squeezing, the $Fe-S-Fe$ and $S-Fe-S$ angles increase and decrease their widths respectively.

In general, it can be observed that the employ of BS and EBS approaches causes a lowering of the symmetry of the cluster and a removal of degeneracy between the geometrical parameters. While with high spin description the values of all the bond and $Fe \cdots Fe$ distances and of all the angle widths are almost coincident, producing the highly symmetric O_h structure, when BS or EBS treatments are employed, two well separate groups of geometrical parameters are observed. The symmetry lowering is therefore produced and the resulting structure is very similar to the $\text{Fe}_4\text{S}_4(\text{N}(\text{Si}(\text{CH}_3)_3)_2)_4$ cluster studied by Chu *et al.* [84]. In both the cases, the two tetrahedra composed by, respectively, the four iron and the four sulphur atoms (contained in the core of the cluster) are the one compressed and the other elongated. Therefore, a structure having rhomboidal

[2Fe2S] faces, where the main diagonal connects the sulphur atoms, is defined.

The same removal of degeneracy (which produces eight longer and four shorter $Fe - S$ bonds) could also be observed in the $Fe_4S_4Cl_4^{2-}$ cluster, studied by Wong, Bobrik, and Holm [168] using X-ray crystallography. The authors excluded the presence of any valence-trapping phenomenon, therefore the observed removal of degeneracy is fully analogous to what happening for the currently studied all-ferric $Fe_4S_4Cl_4$ specie.

Table 4.2: Geometrical comparison - $Fe_4S_4Cl_4$.

		$Fe \cdots Fe$ long (2)	$Fe \cdots Fe$ short (4)	$Fe - S$ long (8)	$Fe - S$ short (4)
B3LYP	HS		3.169		2.394
	BS	3.144	2.960	2.370	2.286
	EBS	3.138	2.938	2.366	2.273
BP	HS		3.084		2.378
	BS	2.978	2.702	2.315	2.186
	EBS	2.962	2.672	2.305	2.172
M06	HS		3.060		2.365
	BS	3.054	2.878	2.346	2.270
	EBS	3.049	2.860	2.345	2.260
OPBE	HS		3.040		2.355
	BS	2.956	2.707	2.301	2.186
	EBS	2.942	2.678	2.292	2.172
		$Fe - S - Fe$ big (4)	$Fe - S - Fe$ small (8)	$S - Fe - S$ big (8)	$S - Fe - S$ small (4)
B3LYP	HS		82.8		96.7
	BS	83.2	79.0	100.7	94.7
	EBS	83.1	78.5	101.2	94.5
BP	HS		80.8		98.5
	BS	80.1	73.7	105.7	94.9
	EBS	79.9	73.3	106.2	94.7
M06	HS		80.6		98.7
	BS	81.2	77.1	102.4	96.1
	EBS	81.2	76.8	102.7	95.5
OPBE	HS		80.4		98.8
	BS	79.9	74.2	105.2	95.5
	EBS	79.8	73.7	105.8	95.2

Distances are Å, angles are °. Numbers inside the parentheses in the heading rows indicate the number of geometrical parameters in the structure.

The non-ideal tetrahedral geometry of the cluster produces, as written above and as one can clearly see in the first row of Table 4.2, two distinct groups of Fe

atom pairs, one with shorter iron-iron distances and smaller $Fe - S - Fe$ angles and the other one with longer distances and wider angles.

As observed for the $Fe_4S_4(N(Si(CH_3)_3)_2)_4$ cluster by Chu *et al.* [84], the most intense structural changes passing from the BS to EBS description are observed for the functionals BP and OPBE.

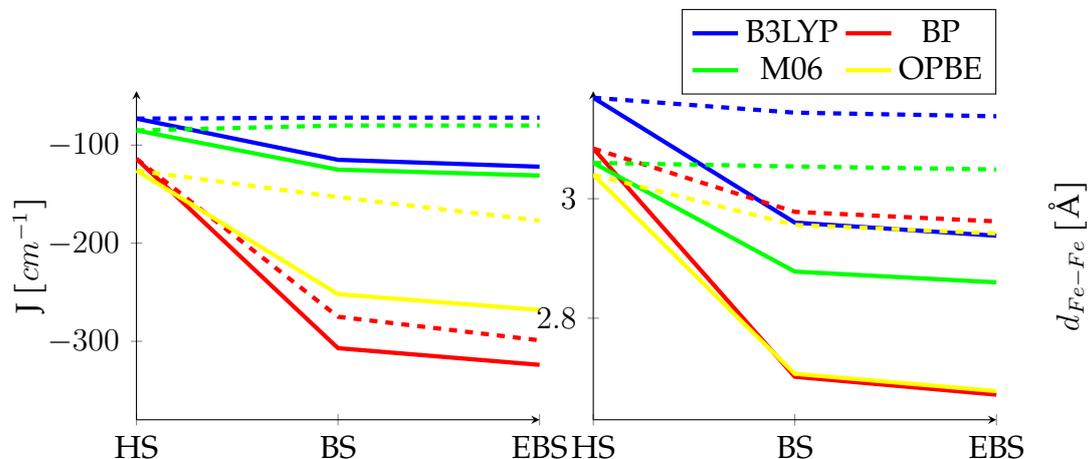
The application of the most general and rigorous form of the Heisenberg-Dirac-van Vleck Hamiltonian performed within the Extended Broken Symmetry method (in contrast with the empirically approximated techniques used elsewhere[174] when studying similar clusters), permits to obtain all the possible $\frac{N_{sc}(N_{sc}-1)}{2}$ J-coupling constants values (six, in the current case, considering that the number of the spin centres N_{sc} is four). Nevertheless, it is possible to identify, among the values, two well-distinct groups, in agreement with the D_{2d} symmetry point-group of the iron-sulfur core and with the previously highlighted presence of two well-distinct groups of spin centres. The averages of the J values over the two groups are shown in Figure 4.5, along with the corresponding average iron-iron distances. All the J-coupling constants are listed in detail in Table 6.4. The values of these constants for two high spin and low-spin Broken Symmetry state optimized structures have been calculated by applying single point EBS calculations on them.

It is worth to note that when the hybrid DFT functionals such as B3LYP and M06 are employed (therefore producing a strong localization of the unpaired electrons), the Extended Broken Symmetry description still enhances, thanks to the more rigorous description of the static correlation phenomenon, the divergence between the two groups of iron atoms with respect to the Broken Symmetry approach. For example, when the M06 functional is employed, the gap between J^{strong} and J^{weak} goes from 45 cm^{-1} at the BS level to 51 cm^{-1} at the EBS level.

4.2.3 Vibrational frequencies

The extension of the generalized EBS approach to the calculation of the vibrational frequencies is the most interesting advancement accomplished during my PhD research concerning this matter. To understand the effects produced on the vibrational pattern, with particular reference to the computed frequencies of the vibrations of these transition metal complexes, due to the inclusion of the correct treatment for static correlation, the comparison between the EBS and BS results is purposely reported in Figure 4.6 for $Fe_2S_2Cl_4^{-2}$. The frequency comparisons

Figure 4.5: J-coupling constant and $Fe \cdots Fe$ distance vs spin description - $Fe_4S_4Cl_4$.



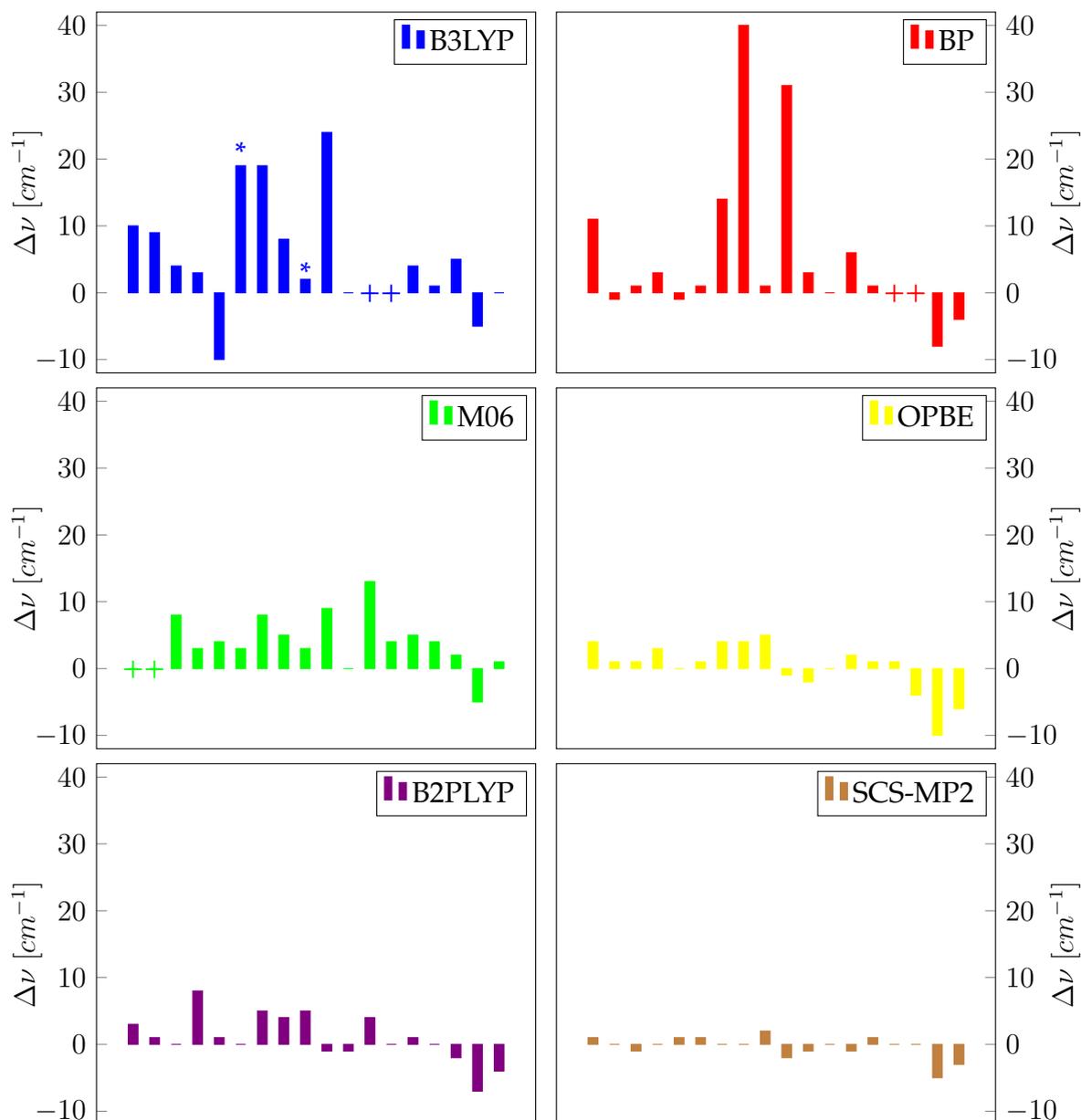
Solid lines are for the first group (short $Fe \cdots Fe$ distances, small $Fe - S - Fe$ angles, J^{strong}), dashed lines are for the second group (long $Fe \cdots Fe$ distances, large $Fe - S - Fe$ angles, J^{weak}).

are reported in a more detailed way in Table 6.5 and Table 6.6.

The problem of identifying the correspondence between the normal modes calculated using the EBS and BS descriptions was faced by decomposing each EBS mode into the space spanned by the orthonormal modes calculated within the BS formalism. More precisely, for each EBS normal mode, the dot products between its 3N-dimensional representation (where 3N is the total number of the Cartesian displacements of the N-atom cluster) and the corresponding representations of the BS normal modes were considered for this kind of calculation. Following this procedure, the values we are concerned with were provided and normalized and their squares were interpreted as percentages quantifying the similarity between each EBS mode with all the other BS ones. A detailed example of this decomposition is provided in Table 6.2.

The results of these decompositions are rather evident for a wide number of normal vibrations but, nevertheless, it was not always straightforward to establish a direct and unambiguous correspondence between vibrational modes, at least for some of them characterized by an high vibrational coupling. In these circumstances, in fact, a given EBS mode can be expressed as a linear combination of several BS modes having comparable coefficients (and thus weights expressing the relative amount of participation to the whole normal mode). This situation is represented in Figure 4.6 by a cross sign (+). Further, in these cases, the frequency shifts can not be calculated for evident reasons, since the exact determination of a shift requires the unambiguous correspondence between nor-

Figure 4.6: Comparison between Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$.



Frequency differences are expressed in cm^{-1} and are defined as $\nu_{GS} - \nu_{BS}$. When * is present, the correspondence is dubious. When + is present, a clear one-to-one correspondence between normal modes could not be asserted and the shift was no calculated.

mal modes.

As one can realize from the inspection of Figure 4.6 for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ and excluding the lowest frequencies that are attributable to a combination of several BS modes having comparable weights (and whose values affected by considerable error), the shifts between the corresponding normal modes are not thoroughly negligible. As an example, a significant difference is observed for the EBS modes

6, 7 and 10 when using the B3LYP functional and for the EBS modes 8 and 10 when using the BP one. Such frequency shifts are often greater than 20 cm^{-1} and for the mode 8 calculated with the BP functional a remarkable change of 36% is observed with respect to the frequency obtained by employing the traditional Broken Symmetry approach. It is interesting to note that SCS-MP2 and B2PLYP methods show the least significant frequency changes when going from the BS to the EBS description among all the employed levels of theory. For the SCS-MP2 level of theory, the largest frequency shift just amounts to 5 cm^{-1} and that can be likely explained by the negligible structural difference found between the BS and EBS optimized structures, as shown in Table 4.1. These small structural differences are also reflected by the very constant (and close to zero) values that the exchange coupling constant keeps when varying the spin description.

As for the molecular system considered so far, the larger $\text{Fe}_4\text{S}_4\text{Cl}_4$ cluster shows some meaningful changes for some vibrational frequencies of some normal modes. The inspection of Table 6.7, Table 6.8 and Table 6.9 can be useful for this purpose.

For example, the modes 16 and 19 at the BP functional level have vibrational shifts equal to 32 cm^{-1} and 31 cm^{-1} , respectively, while the B3LYP functional provides 20 cm^{-1} for the mode 15 and for the mode 21 the shift -38 cm^{-1} is provided from the M06 functional. If the relative frequency changes are considered, in place of the absolute ones, the shifts are even more significant: for the former cases, referred to the BP functional, the change in fact amounts to 15%. No general trend seems to be present with respect to the sign of these frequency shifts.

It is fundamental to point out that the previous comparison was performed between the modes calculated on the same geometry; in particular, on the structure obtained from the optimization within EBS formalism. Doing this, the observed shifts in frequency (even though the use of the term "frequency" is a bit unrigorous in the case of BS description, being them not calculated in a minimum energy point for the given level of theory) are exclusively produced from the finer theoretical description achieved with Extended Broken Symmetry formalism, rather than being connected, even only partly, with a different geometrical arrangement.

Another hint inferred from the inspection of the normal modes showing a significant frequency shift is that the atomic displacements associated with them almost exclusively involve the iron-sulfur core, that is the portion of the cluster the description of which is most concerned with the employment of the EBS

method, without significantly affecting the chloride atoms.

It is also worth to observe that, for all the normal modes previously mentioned, the correspondence between the EBS and BS ones is remarkable. In the worst case, in fact, the superposition is around 79% whereas, on the opposite side, sometimes 100% can be reached. These noticeable percentages guarantee that the obtained shifts (reported in the last columns of the Tables in subsection 6.1.5) are fully a consequence of the theoretical approach adopted and are not due to any wrong assignment in the correspondence between modes.

At last, it cannot be omitted the observation that some inversions in the sequence of the computed frequencies may occur for some normal modes, as it is actually observed comparing the results shown in Tables from Table 6.5 to Table 6.9. This may for example occur for the modes 6 and 7 for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ cluster with the B3LYP functional and for the modes from 15 to 21 for $\text{Fe}_4\text{S}_4\text{Cl}_4$ cluster with the OPBE functional.

At this point of the discussion, it seems helpful to remind that the 3N-6 normal modes of the D_{2h} symmetry anion ($\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$) produce the following $\Gamma_{vib} = 4A_g + A_u + 2B_{1g} + 3B_{1u} + 2B_{2g} + 2B_{2u} + B_{3g} + 3B_{3u}$.

With the aim of decomposing the calculated normal modes into the more frequently adopted terms of bond stretchings, angle bendings or dihedral angle torsions, it was decided to employ the recently developed VMARD[153] (Vibrational Mode Automatic Relevance Determination) method, based on the Bayesian Ridge Regression technique[175]. The results accomplished from this procedure are shown in Table 4.3 and Table 4.4 for the levels of theory that allowed to reach the most accurate results from the structural and magnetic points of view, namely the M06/def2-TZVP and B2PLYP/def2-TZVP methods. The results concerning the other levels of theory are instead reported in Table 6.10, Table 6.11, Table 6.12 and Table 6.13.

Such a decomposition, along with the symmetries assigned to the normal modes of vibration, has provided an essential foothold for performing a comparison between the calculated frequencies and the experimental ones, reported in Smith *et al.* [167]. The results are listed in Table 4.5.

As it can be argued, the matching between the calculated and measured values may be regarded of satisfactory level, except for the normal modes at 337 and 332 cm^{-1} , respectively having B_{1g} and B_{2g} symmetry. Significant red-shifts, up to 60 cm^{-1} for the B2PLYP/def2-TZVP level of theory, can be appreciated for these modes. A deeper insight into these modes would indeed show that M06/def2-TZVP level of theory produces the highest discrepancy, underesti-

Table 4.3: Extended Broken Symmetry VMARD analysis on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - M06/def2-TZVP.

Mode	Point-group symmetry	ν	VMARD
1	B_{1u}	25	54% β_{FeFeCl}
2	A_u	40	35% ρ_{ClFeFeCl} + 20% β_{SFeCl} + 13% β_{FeSFe}
3	A_g	79	33% β_{FeFeCl} + 20% β_{FeSFe} + 21% β_{ClFeCl} + 21% ν_{FeCl}
4	B_{2u}	97	41% β_{SFeCl} + 19% ν_{FeS}
5	B_{2g}	115	30% β_{FeFeCl} + 29% ν_{FeCl} + 21% β_{FeSFe} + 11% β_{ClFeCl}
6	B_{3u}	126	16% β_{FeFeCl} + 13% β_{ClFeCl} + 11% ν_{FeCl}
7	B_{3g}	129	64% β_{SFeCl} + 10% β_{FeFeS}
8	B_{1g}	145	69% β_{SFeCl}
9	A_g	178	43% ν_{FeCl} + 31% β_{FeSFe} + 10% β_{ClFeCl}
10	B_{1u}	180	48% β_{ClFeS} + 34% ν_{FeCl}
11	B_{3u}	280	64% ν_{FeCl} + 33% β_{SFeS}
12	B_{1g}	298	67% β_{FeFeS} + 19% ν_{FeCl} + 11% ν_{FeS}
13	B_{2g}	311	87% ν_{FeCl}
14	B_{1u}	324	83% ν_{FeCl}
15	A_g	330	62% ν_{FeCl} + 24% ν_{FeS}
16	B_{3u}	341	47% ν_{FeCl} + 43% β_{SFeS}
17	A_g	395	74% ν_{FeS} + 22% ν_{FeCl}
18	B_{2u}	424	77% ν_{FeS}

Frequencies are expressed in cm^{-1} . ν indicates bond stretchings, β indicates angle bendings and ρ indicated dihedral angles torsions. The contributions smaller than 10% have been omitted. In order to compare the calculated frequencies with the experimental ones (which have been measured using the NRVS technique on the Fe^{57} isotope), the mass of the Fe atoms has been set to 56.93539 a. m. u..

mating these theoretical frequencies much more than B2PLYP. It is also interesting to point out that the theoretical spectra are generally shifted, except for a few modes, to lower wavenumbers.

Table 4.4: Extended Broken Symmetry VMARD analysis on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B2PLYP/def2-TZVP.

Mode	Point- group symmetry	ν	VMARD
1	B_{1u}	42	67% β_{FeFeCl}
2	A_u	66	92% ρ_{ClFeFeCl}
3	A_g	91	43% β_{ClFeCl} + 31% β_{FeSFe}
4	B_{2g}	97	51% β_{FeFeCl} + 45% ν_{FeCl}
5	B_{2u}	101	68% β_{SFeCl} + 31% β_{FeSFe}
6	B_{3u}	114	56% β_{ClFeCl} + 24% ν_{FeCl} + 15% β_{SFeS}
7	B_{3g}	119	100% ρ_{SFeFeCl}
8	B_{1g}	143	86% β_{SFeCl}
9	B_{1u}	156	58% β_{SFeCl} + 35% ν_{FeCl}
10	A_g	168	50% ν_{FeCl} + 36% β_{FeSFe} + 14% β_{ClFeCl}
11	B_{3u}	268	73% ν_{FeCl} + 26% β_{SFeS}
12	B_{1g}	275	90% β_{FeFeS}
13	B_{2g}	298	95% ν_{FeCl}
14	B_{1u}	310	85% ν_{FeCl}
15	A_g	311	68% ν_{FeCl} + 18% ν_{FeS} + 14% β_{FeSFe}
16	B_{3u}	338	65% β_{SFeS} + 32% ν_{FeCl}
17	A_g	383	76% ν_{FeS} + 15% ν_{FeCl}
18	B_{2u}	408	84% ν_{FeS}

Frequencies are expressed in cm^{-1} . ν indicates bond stretchings, β indicates angle bendings and ρ indicated dihedral angles torsions. The contributions smaller than 10% have been omitted. In order to compare the calculated frequencies with the experimental ones (which have been measured using the NRVS technique on the Fe^{57} isotope), the mass of the Fe atoms has been set to 56.93539 a. m. u..

The application of a scaling factor (that would generally be, for the considered spectral range, bigger than one[176]) should likely improve the matching between theory and experiment but this is out of the scope of the present study.

More than on the absolute match with experimental values, the attention is in fact focused on the difference in performance between BS and EBS approaches for Normal Mode Analysis tasks, and the application of the same scaling fac-

Table 4.5: Normal mode frequency comparison - $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ M06/def2-TZVP and B2PLYP/def2-TZVP.

Experimental[167]			M06		B2PLYP	
Point group symmetry	ν	Internal coordinate motion[167]	ν_{EBS}	Shift	ν_{EBS}	Shift
A_g	396	FeS stretch	395	1	383	13
A_g	313	$FeCl$ stretch	330	-17	311	2
A_g	177	Fe_2S_2 bend	178	-1	168	9
A_g	101	FeS ring bend	79	22	91	10
A_u	-	$FeCl_2$ twist	40	-	66	-
B_{1g}	337	FeS stretch	298	39	275	62
B_{1g}	145	$FeCl_2$ wag	145	0	143	2
B_{1u}	329	$FeCl$ stretch	324	5	310	19
B_{1u}	163	ring puckering and $FeCl_2$ rock	180	-17	156	7
B_{1u}	15-30	ring puckering	25	-10/5	42	-12/-27
B_{2g}	332	$FeCl$ stretch	311	21	298	35
B_{2g}	96	$FeCl_2$ rock	115	-19	97	-1
B_{2u}	415	FeS stretch	424	-9	408	7
B_{2u}	118	$FeCl_2$ wag	97	21	101	17
B_{3g}	-	cluster twist	129	-	119	-
B_{3u}	346	$FeCl$ stretch	341	5	338	8
B_{3u}	289	FeS stretch	280	9	268	21
B_{3u}	123	$ClFeCl$ bend	126	-3	114	9

Frequencies are expressed in cm^{-1} . The shifts are defined as $\nu_{Exp} - \nu_{EBS}$. In order to compare the calculated frequencies to the experimental ones (which have been measured using the NRVs technique on the Fe^{57} isotope), the mass of Fe atoms has been set to 56.93539.

tor for the two sets of frequencies would therefore produce any change in their relative accuracy. The quantification of the matching with experimental measures is not straightforward and I decided to roughly assume the sum of all the shifts ($\sum_{i=1}^{3N-6} |\nu_i^{exp} - \nu_i^{theo}|$) as a metric for expressing the error associated with a given level of theory. In other words, the sum of the absolute values of the "Shift" columns in Table 4.5 represents the mismatch of the EBS calculations. For the BS description, the shifts with respect to the experimental values are obtained by summing the quantities graphically depicted in Figure 4.6, and more exten-

sively reported in subsection 6.1.5, to the shifts reported in Table 4.5. Doing this, it can be definitely affirmed that the employ of the Extended Broken Symmetry method produces an overall improvement in the quality of the vibrational description with respect to the traditional Broken Symmetry approach. The overall frequency shift goes in fact from 201 to 192 cm^{-1} at the M06/def2-TZVP level and from 258 to 242 cm^{-1} at the B2PLYP/def2-TZVP level when going from the BS to the EBS description. In conclusion, it has again to be pointed out that, having both the BS and EBS calculations been performed on the same structure, such an improvement is entirely due to the more correct theoretical description delivered by the innovative EBS method.

4.3 F-RESP

The development of polarizable force fields is nowadays one of the most lively fields of classical Molecular Dynamics. The enormous increase in the available computer power, along with actual revolutions such as the appearance of GPU-computing, make now possible to describe many-body systems with unimaginable level of detail and accuracy. Due to the magnitude of the concerned interactions, the electrostatic side is obviously the natural choice for these improved descriptions.

The different formalisms available for including the effects of polarization in the classical Molecular Dynamics simulations have been listed in subsection 2.4.5. Among them, the use of virtual multipoles induced on the atomic positions is the most commonly adopted and the most resemblant to the actual polarization process, though it was recently pointed out that the currently available polarizable force fields generally underestimate the importance of the charge fluctuation process[177]. These formalisms require the use of a self-consistent or of an extended Lagrangian approach for taking into account the mutual interplay between the charge on the single atom and the electric field produced by the other atoms in the surroundings. As a natural consequence, Molecular Dynamics simulations using these methods are slower than the same calculation performed without a polarization treatment (*i.e.* with fixed atomic charges). The difference in the computational speed is generally around one order of magnitude.

It would therefore be very useful to develop faster algorithms, whose solution does not require an iterative procedure, for the description of polarization.

In the course of my studies in the group where I have performed this Thesis, I implemented the pre-existing F-RESP model into the well-known classical

Molecular Dynamics code LAMMPS[178]. Doing this, the method was made apt to the use in large-scale calculations and to be interfaced with all the useful features of a modern code as LAMMPS, at variance with the previous implementation, which had severe issues, such as the lack of parallelization. Moreover, I further extended the method, by implementing the use of Fennell method in place of Ewald summation for the fictitious electric field calculation. In this way, the virial of the forces is now easily available, making therefore possible to estimate the pressure of the system and to consequently perform simulations in the NPT statistical ensemble. Eventually, the method was employed for simulations on liquid water, obtaining important results.

4.3.1 The polarization model

In this model, the atomic charges are expressed as:

$$q_i = q_i^0 + \sum_{b \in M_i} \alpha_b^i \vec{E}^{fic}(\vec{r}_{h_b}) \cdot \hat{r}_b \quad (4.2)$$

where q_i^0 is a parameter typical of each atomic specie i , α_b^i is the polarization coefficient that depends on the atomic specie i and on the bond type b , $\vec{E}^{fic}(\vec{r}_{h_b})$ is the fictitious electric field exerted on the half-bond point of the bond b (\vec{r}_{h_b}), \hat{r}_b is the unit vector of bond b and M_i identifies the molecule to which atom i belongs. According to this equation, each atomic charge of a molecule can be influenced by the polarization of the electronic density of each covalent bond of the molecule. However, such a dependence is usually limited to the subset of bonds where the atom takes part to avoid to slow down the calculation heavily.

The electric field is "fictitious" because it is not produced by the current set of atomic charges but rather by a set of parametric charges, called "generating" charges, having the same positions of the atoms of the system at the current time-step. The slow self-consistent process that would be required by the use of the real charges can therefore be avoided.

The fictitious electric field can be calculated in two ways, that are completely independent from the technique employed in the main electrostatic calculation of the MD simulation. The first is through the use of the Ewald Summation technique[97] while in the second one a modification of Wolf's potential[179] is used, namely that proposed by Fennell and Gezelter [180].

While in the former way the classical method for the evaluation of the long range electrostatic interactions has been chosen, the latter and more recent one is reported in many articles to return results as good[181] as the smoothed ver-

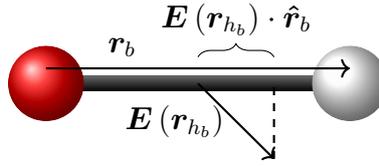


Figure 4.7: Projection of the fictitious electric field along the bond axis. The vectors representing the bond axis and the electric field exerted on the half-bond point are displayed. The projection of the electric field over the bond axis is also highlighted.

sions of the Ewald summation in shorter time[182] (lacking the part of the calculation performed in reciprocal space). Using this technique, moreover, the expression of the virial contribution coming from fluctuating charges has been obtained, permitting thus to correctly perform simulations under isobaric conditions.

A further improvement of the Fennell potential was introduced by Waibel[183] and it was implemented in the simulation code alongside the other two although no significant enhancement in the accuracy of the results could be observed.

Other contributions can be added to Equation 4.2, permitting to make charges depending on the molecular structure, in particular on the geometrical descriptors such as bond lengths, angle widths and improper dihedral angle widths, obtaining the final and general expression that is:

$$q_i = q_i^0 + \sum_{b \in M_i} \alpha_b^i \vec{E}^{fic}(\vec{r}_{h_b}) \cdot \hat{r}_b + \sum_{b \in M_i} \kappa_b^i (|\vec{r}_b| - |\vec{r}_b^0|) + \sum_{a \in M_i} \zeta_a^i (\theta_a - \theta_a^0) + \sum_{d \in M_i} \delta_d^i |\chi_d - \chi_d^0| \quad (4.3)$$

where the parameters κ_b^i , ζ_a^i and δ_d^i both depend on bond, angle and improper angle type and obviously on the atomic species. Eventually, $|\vec{r}_b^0|$, θ_a^0 and χ_d^0 are the natural values for the given bonds, angles or improper dihedrals, as defined in the underlying force field handling the bonded interactions. As written in Equation 4.2, even for the previous relation the influence of many geometrical descriptors can be ignored, because its magnitude, based on simple physical considerations, is negligible and in order not to slow down the calculation too much. This expression is moreover used for the parameterization, as described in detail in subsection 4.3.5 and therefore the inclusion of a quantity of terms as small as possible prevents the risk of over-fitting.

4.3.2 Fictitious electric field calculation

4.3.2.1 Ewald summation

Fictitious electric field calculated through the use of Ewald summation can be divided into three contributions: non-bonded, bonded and reciprocal space.

$$\vec{E}^{fic}(\vec{r}_{h_b}) = \vec{E}_{NB}^{fic}(\vec{r}_{h_b}) + \vec{E}_B^{fic}(\vec{r}_{h_b}) + \vec{E}_K^{fic}(\vec{r}_{h_b}) \quad (4.4)$$

While reciprocal space term is the same of the usual Ewald summation (namely the potential exerted by the shielding Gaussian distributions, evaluated in the Fourier reciprocal space), the non-bonded and bonded terms can be considered in correspondence with the real space and self-interaction terms. The non-bonded term is the electric field (evaluated in the real space) exerted by the atoms within the given F-RESP real-space cutoff that are not part of the molecule to which the considered bond (that giving the half-bond point \vec{r}_{h_b}) belongs. The bonded term, on the contrary, is needed to "cancel out" the electric field exerted by the atoms belonging to the same molecule of the bond considered. This undesired (in our model) contribution comes from the reciprocal space term, in which it is included for efficiency reasons (exactly as the self-interaction moiety in classical Ewald summation).

The non-bonded contribution is:

$$\vec{E}_{NB}^{fic}(\vec{r}_{h_b}) = \sum_{j \in V.L._b - M_b} \frac{q_j^{gen}}{|\vec{r}_{jh_b}|} \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} + \frac{erfc(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \hat{r}_{jh_b} \quad (4.5)$$

and the bonded one is:

$$\vec{E}_B^{fic}(\vec{r}_{h_b}) = \sum_{j \in M_b} \frac{q_j^{gen}}{|\vec{r}_{jh_b}|} \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{erf(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \hat{r}_{jh_b} \quad (4.6)$$

where $V.L._b$ is a Verlet list[93] of atom whose distance from the half-bond point of the bond b is smaller than the F-RESP cutoff in real space and M_b is the molecule to which the bond b belongs.

At last, the reciprocal space contribution is:

$$\vec{E}_K^{fic}(\vec{r}_{h_b}) = \sum_{\vec{k} \neq 0} \vec{k} \sum_{j=1}^{N_a} \frac{4\pi}{|\vec{k}|^2} q_j^{gen} \mathfrak{S} \left(e^{i\vec{k} \cdot (\vec{r}_{jh_b})} \right) e^{-\frac{|\vec{k}|^2}{4\alpha^2}} \quad (4.7)$$

It is important to point out that the generating charges (q_j^{gen}) are used instead of the real ones.

As in almost any polarizable force field, an additional contribution to atomic forces comes from the geometrical gradient of charges (equal to zero when dealing with fixed charges). Recalling Equation 4.2, it is possible to write:

$$\nabla_{\vec{r}_k} q_i = \sum_{b \in M_i} \alpha_b^i \nabla_{\vec{r}_k} \left(\vec{E}^{fic}(\vec{r}_{h_b}) \cdot \hat{r}_b \right) \quad (4.8)$$

Being the fictitious electric field expressed by three components (\vec{E}_{NB} , \vec{E}_B and \vec{E}_K) and considering the three possible distinct atomic positions \vec{r}_j , \vec{r}_{b_1} and \vec{r}_{b_2} (where j is the position of the atom exerting the electric field and $b_{1/2}$ are the extremes of the bond), nine expressions arise (they are reported in subsection 6.2.1). Along with them, it can be found a schema in pseudo-code showing how the computational order of the whole calculation can be maintained, thanks to some rearrangement in the flow of the calculations, $O\left(N^{\frac{3}{2}}\right)$, namely the same of the classical well-tuned Ewald summation.

4.3.2.2 Fennell potential

This approach is much simpler than the Ewald summation. In fact, the part in the reciprocal space is lacking and, consequently, the bonded correction is no more needed. The electric field on the half-bond point can be expressed as:

$$\vec{E}^{fic}(\vec{r}_{h_b}) = \sum_{j \in V.L.b-M_b} q_j^{gen} \left(\frac{erfc(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} - f_{shift} \right) \hat{r}_{jh_b} \quad (4.9)$$

where $f_{shift} = \frac{erfc(\alpha R_c)}{R_c^2} + \frac{2\alpha e^{-\alpha^2 R_c^2}}{\sqrt{\pi} R_c}$ and R_c is the cutoff radius. It is worth to note that f_{shift} is a constant term, depending on the chosen cutoff radius and on the α coefficient for the Fennell potential (that are, such as for the method based on the Ewald summation, completely independent on those employed by the main electrostatic solver).

Also in this instance, the $\nabla_{\vec{r}_k} \left(\vec{E}^{fic}(\vec{r}_{h_b}) \cdot \hat{r}_b \right)$ expression is changing according to k , in particular on its being one of the extremes of the bond or a third atom j that exerts the electric field. The expressions for the gradients can be found in subsection 6.2.2.

A crucial advantage of the employ of the Fennell potential for the calculation of the fictitious electric field with respect to the Ewald summation is that the contribution to the virial coming from the charge fluctuation can be easily derived. As shown in paragraph 4.3.2.2.1, I could in fact demonstrate that, due

to the lack of a long-range contribution to the fictitious electric field, the additional virial term reduces to the simple expression $\sum_{i=1}^N \vec{f}_i^{F-RESP} \cdot \vec{r}_i$.

A slight complication connected to the use of Fennell potential is lying in the need to determine the α and R_c parameters. Unluckily, the choice of these quantities is not straightforward as for the Ewald summation (or its smoothed versions), for which algorithms that correlate these parameters to the required accuracy have been derived[184, 185] and implemented in the most common simulation packages. A case by case evaluation is therefore needed.

Taking inspiration from what did by Hens and Vlugt [181], it was decided to determine the optimal set of parameters by using the subsequent formula for the calculation of the relative error:

$$err(\alpha, R_c) = \frac{100}{N_F} \cdot \sum_{\substack{f \in \\ Frames}} \sum_{b \in f} \left| \frac{\vec{E}_{\alpha, R_c}^{fic, Fennel}(\vec{r}_{h_b}) \cdot \hat{r}_b - \vec{E}^{fic, Ewald}(\vec{r}_{h_b}) \cdot \hat{r}_b}{\vec{E}^{fic, Ewald}(\vec{r}_{h_b}) \cdot \hat{r}_b} \right| \quad [\%] \quad (4.10)$$

where N_F is the number of frames (snapshots taken from a Molecular Dynamics trajectory) employed for the evaluation. In Equation 4.10, the set of fictitious electric field projections along the bonds ($\vec{E}^{fic}(\vec{r}_{h_b}) \cdot \hat{r}_b$) calculated using the Ewald summation with an high accuracy is chosen as reference.

For the simulations of liquid water reported in section 4.5, a scan was then performed by varying α and R_c over reasonable ranges of values (the range for the cutoff radius R_c was, for example, 8-18 Å).

The Figure 4.8 shows the error surface obtained from an average over 5 frames extracted from a NVE simulation having as starting point the final configuration of an equilibrated NPT simulation (298 K, 1 atm). The NVE ensemble was chosen in order for all the frames to have the same probability according to the Boltzmann distribution. As can be seen, the error roughly decreases when increasing R_c cutoff value (as could be easily imagined). It was decided to adopt an intermediate 13 Å value for the cutoff radius; having fixed it, the optimal α value, returning an error around 3%, is 0.095 \AA^{-1} .

4.3.2.2.1 F-RESP virial contribution with the Fennell potential employed for the fictitious electric field calculation This demonstration is obtained when the Ewald summation technique is employed in the main electrostatic calculation but the following analytical virial evaluation has shown to be numerically correct also when using PPPM[100] (as it is implemented in the LAMMPS package) method.

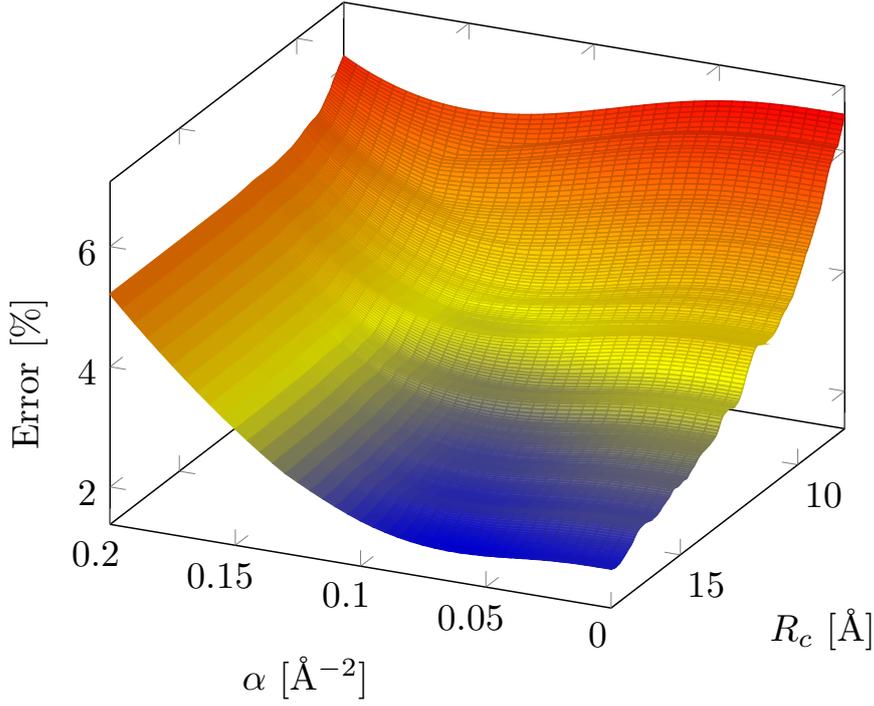


Figure 4.8: Relative error with respect to the α and R_c parameters for TIP3P water.

Knowing that the energy of the system is $E^{rec} = \frac{1}{2} \sum_{n=1}^N q_n \cdot \phi^{rec}(\vec{r}_n)$, where:

$$\phi^{rec}(\vec{r}_n) = \frac{8\pi}{V} \sum_{\vec{k} \neq 0} \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \left(\Re e^{i\vec{k} \cdot \vec{r}_n} \cdot \Re S(\vec{k}) + \Im e^{i\vec{k} \cdot \vec{r}_n} \cdot \Im S(\vec{k}) \right) \quad (4.11)$$

and that the virial is $\frac{\partial E}{\partial V}$, it follows that:

$$\frac{\partial E^{rec}}{\partial V} = \frac{1}{2} \sum_{n=1}^N \left(\frac{\partial q_n}{\partial V} \phi^{rec}(\vec{r}_n) + q_n \frac{\partial \phi^{rec}(\vec{r}_n)}{\partial V} \right) \quad (4.12)$$

While the first term between parentheses can be easily expressed using the classical virial form (namely $\frac{1}{3V} \sum_{n=1}^N \vec{f}_n^{rec} \cdot \vec{r}_n$, by using the derivative chain rule), the derivative of $\phi_{rec}(\vec{r}_n)$ with respect to the volume is:

$$\begin{aligned} \frac{\partial \phi^{rec}(\vec{r}_n)}{\partial V} = & 8\pi \left(\sum_{\vec{k} \neq 0} \left(\Re e^{i\vec{k} \cdot \vec{r}_n} \cdot \Re S(\vec{k}) + \Im e^{i\vec{k} \cdot \vec{r}_n} \cdot \Im S(\vec{k}) \right) \left(\frac{1}{V} \frac{\partial e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{\partial V} - \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \right) \right) + \\ & \sum_{\vec{k} \neq 0} \frac{1}{V} \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \left(\Re e^{i\vec{k} \cdot \vec{r}_n} \frac{\partial \Re S(\vec{k})}{\partial V} + \Im e^{i\vec{k} \cdot \vec{r}_n} \frac{\partial \Im S(\vec{k})}{\partial V} \right) \end{aligned} \quad (4.13)$$

by knowing that $e^{i\vec{k}\cdot\vec{r}_n}$ is constant with respect to the volume variation due to the compensation between the change in the vectors \vec{k} and \vec{r}_n .

While the first part of the summation is typical of a fixed-charge force field, the second is directly connected to the charge variation and permits to express the F-RESP contribution to the energy derivative with respect to the volume as:

$$\begin{aligned} & \sum_{n=1}^N q_n \frac{\partial \phi^{rec, F-RESP}(\vec{r}_n)}{\partial V} = \\ & \frac{8\pi}{V} \sum_{\vec{k} \neq 0} \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \left(\Re S(\vec{k}) \sum_n \frac{\partial q_n}{\partial V} \Re e^{i\vec{k}\cdot\vec{r}_n} + \Im S(\vec{k}) \sum_n \frac{\partial q_n}{\partial V} \Im e^{i\vec{k}\cdot\vec{r}_n} \right) \end{aligned} \quad (4.14)$$

By inverting the summation order:

$$\begin{aligned} & \sum_{n=1}^N q_n \frac{\partial \phi^{rec, F-RESP}(\vec{r}_n)}{\partial V} = \\ & \sum_n \frac{\partial q_n}{\partial V} \left(\frac{8\pi}{V} \sum_{\vec{k} \neq 0} \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \left(\Re S(\vec{k}) \Re e^{i\vec{k}\cdot\vec{r}_n} + \Im S(\vec{k}) \Im e^{i\vec{k}\cdot\vec{r}_n} \right) \right) \end{aligned} \quad (4.15)$$

As can be easily seen, this expression is equivalent to $\sum_{n=1}^N \frac{\partial q_n}{\partial V} \phi^{rec}(\vec{r}_n)$ and from this it follows that the contribution to the virial coming from the charge fluctuation is:

$$\frac{\partial E^{rec, F-RESP}}{\partial V} = \sum_{n=1}^N \frac{\partial q_n}{\partial V} \phi^{rec}(\vec{r}_n) \quad (4.16)$$

Knowing that q_n depends over all the atoms contained in the Verlet lists of all the bonds of the molecule whom atom n belongs and using the derivative chain rule, Equation 4.16 can be written as:

$$\begin{aligned} \frac{\partial E^{rec, F-RESP}}{\partial V} &= \sum_{n=1}^N \sum_{b \in M_n} \sum_{j \in V.L.b} \frac{\partial q_n}{\partial \vec{r}_j} \phi^{rec}(\vec{r}_n) \frac{\partial \vec{r}_j}{\partial V} = \\ & \sum_{n=1}^N \sum_{b \in M_n} \sum_{j \in V.L.b} \frac{\partial q_n}{\partial \vec{r}_j} \phi^{rec}(\vec{r}_n) \frac{\vec{r}_j}{3V} \end{aligned} \quad (4.17)$$

And eventually, recalling Equation 4.8:

$$\frac{\partial E^{rec, F-RESP}}{\partial V} = \sum_{n=1}^N \sum_{b \in M_n} \alpha_b^n \sum_{j \in V.L.b} \frac{\partial \left(\vec{E}(\vec{r}_{h_b}) \cdot \hat{r}_b \right)}{\partial \vec{r}_j} \phi^{rec}(\vec{r}_n) \cdot \frac{\vec{r}_j}{3V} \quad (4.18)$$

Concerning the self-interaction and real space energy terms, a similar but easier derivation can be made, eventually demonstrating that the total contribution of the F-RESP charge fluctuation to the virial is:

$$\frac{\partial E^{F-RESP}}{\partial V} = \sum_{b=1}^{N_b} \sum_{n \in M_b} \alpha_b^n \sum_{j \in V.L.b} \frac{\partial \left(\vec{E}(\vec{r}_{h_b}) \cdot \hat{r}_b \right)}{\partial \vec{r}_j} \phi(\vec{r}_n) \cdot \frac{\vec{r}_j}{3V} \quad (4.19)$$

4.3.2.3 Waibel potential

This further potential was very recently introduced[183] as an improvement of the Fennell one. Even in this case, only terms in the real space are present, at the price of an increase in the real space cutoff radius with respect to the Ewald and Ewald-based methods under the same accuracy. This feature permits, as in the previous instance, to easily express the contribution of the charge variation to the virial of forces. The expression of the electric field on the half-bond point is:

$$\vec{E}(\vec{r}_{h_b}) = \sum_{j \in V.L.b-M_b} q_j^{gen} \left(\frac{erfc(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} + \right. \\ \left. - f_{shift} \cdot e^{\beta(|\vec{r}_{jh_b}| - R_c)} \right) \hat{r}_{jh_b} \quad (4.20)$$

where the term f_{shift} has the expression shown in subsection 4.3.2.2.

As one can easily appreciate by comparing Equation 4.9 with Equation 4.20, the unique difference among the electric field expressions is represented by the term $e^{\beta(|\vec{r}_{jM}| - R_c)}$, that in the latter case contains the additional parameter β (independent from the analogous parameter α). The use of this exponential as a switching function (instead of 1) for the shifted force potential should, according to Waibel, Feinler, and Gross [183], reduce the artifacts of the DSF potential and improve the final results although we were not able to obtain, from our simulations, any notable improvement. It was therefore decided to rely on the Fennell potential for the simulations performed on liquid water and ionic liquids.

4.3.3 Damping for the electric field

The use of a damping function for the fictitious electric field has demonstrated to be fundamental to avoid the "polarization catastrophe"[114] connected to

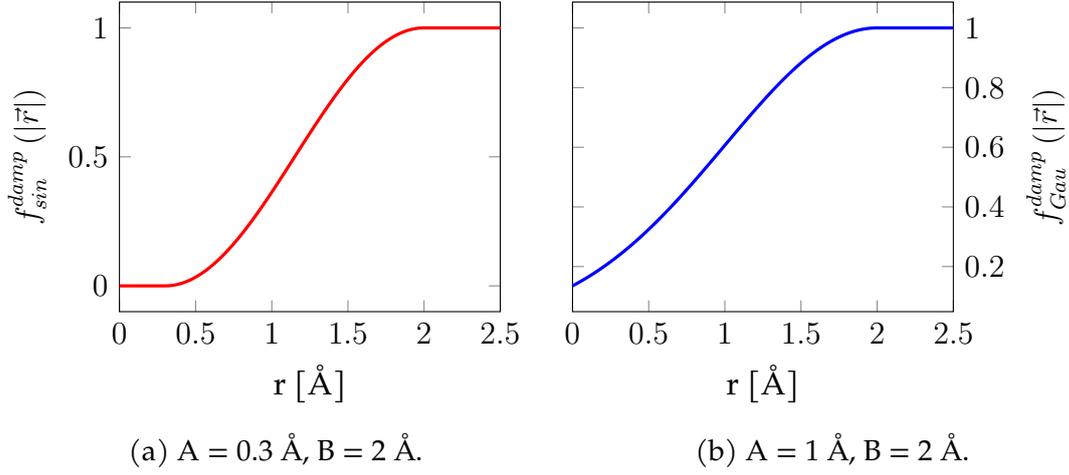


Figure 4.9: Example of the two damping function forms: sinusoidal (left) and Gaussian (right).

the unphysical use of point charges for the description of atomic charge densities. Even though the absence of a mutual dependence between induced and inductive charges prevents our method from the most dramatic consequences of this process, this erroneous description produces a significant alteration in the physics of the simulated system. The presence of functions approximating the smeared distribution of a real electron cloud is therefore very useful.

Two analytical expressions were tried for the damping function, a sinusoidal ($\sin^2\left(\frac{\pi}{2}\left(\frac{x-A}{B-A}\right)\right)$) and a Gaussian ($e^{-\frac{(x-B)^2}{2A^2}}$) one. Both of them gradually reduce the strength of the fictitious electric field when its source gets closer to the half-bond point. In order to assure a physically reasonable behaviour, both functions have to be piecewise-defined:

$$\begin{cases} 0 & 0 \leq x < A \\ \sin^2\left(\frac{\pi}{2}\left(\frac{x-A}{B-A}\right)\right) & A \leq x \leq B \\ 1 & x > B \end{cases} \quad \begin{cases} e^{-\frac{(x-B)^2}{2A^2}} & 0 \leq x \leq B \\ 1 & x > B \end{cases} \quad (4.21)$$

where x is the distance between the half-bond point and the atom exerting the fictitious electric field.

In these expressions, B can assume the meaning of an upper cutoff (meaning that, beyond that distance, a full interaction is acting) and A can assume, only for sinusoidal function, the meaning of a lower cutoff that neglects all interactions with a distance smaller than its value.

An example of the shape of these functions is reported in Figure 4.9.

The derivatives of the damping function (reported in subsection 6.2.4) have to be included in the expression of the electric field gradient, resulting in the following equation:

$$\begin{aligned} \nabla_{\vec{r}_k} \left(\left(\vec{E}^{fic}(\vec{r}_{jh_b}) \cdot f^{damp}(|\vec{r}_{jh_b}|) \right) \cdot \hat{r}_b \right) &= \nabla_{\vec{r}_k} \left(\left(\vec{E}^{fic}(\vec{r}_{jh_b}) \cdot \hat{r}_b \right) \cdot f^{damp}(|\vec{r}_{jh_b}|) \right) = \\ \nabla_{\vec{r}_k} \left(\vec{E}^{fic}(\vec{r}_{jh_b}) \cdot \hat{r}_b \right) \cdot f^{damp}(|\vec{r}_{jh_b}|) &+ \nabla_{\vec{r}_k} f^{damp}(|\vec{r}_{jh_b}|) \cdot \left(\vec{E}^{fic}(\vec{r}_{jh_b}) \cdot \hat{r}_b \right) \end{aligned} \quad (4.22)$$

4.3.4 Geometrical descriptors

The charge fluctuations connected to the variation of the geometrical descriptor values are independent from the technique employed for the fictitious electric field evaluation and the same way, their derivatives:

$$\nabla_{\vec{r}_k} q_i = \sum_{b \in M_i} \kappa_b^i \nabla_{\vec{r}_k} (|\vec{r}_b| - |\vec{r}_b^0|) + \sum_{a \in M_i} \zeta_b^i \nabla_{\vec{r}_k} (\theta_a - \theta_a^0) + \sum_{d \in M_i} \delta_d^i \nabla_{\vec{r}_k} |\chi_d - \chi_d^0| \quad (4.23)$$

where the right-hand gradients are non-zero only when k is one of the atoms involved in the definition of the geometrical descriptor. The effort required for the calculation of this gradient (and of the parent charge variation) is negligible, in analogy with the evaluation of the bonded interactions in any force field. Moreover, if the simulation box is composed by quite small molecules (such as it happens for the simulation of molecular liquids), the number of the geometrical descriptors grows linearly with the size of the system, keeping $O(N)$ the computational order of the gradient evaluation. The analytical expressions for the derivatives can be found in subsection 6.2.5.

4.3.5 Parameterization

The parameters of the F-RESP method are the generating charges $\{q^{gen}\}$, the charges $\{q^0\}$ and the polarization coefficients $\{\alpha\}$. As previously written, the dependence on the structural parameters of the studied system can be also taken into account. The corresponding parameters sets are, as shown in Equation 4.3, $\{\kappa\}$, $\{\gamma\}$ and $\{\delta\}$.

The parameterization is concerned with the aim of obtaining at first the set of the generating charges and, subsequently, the rest of the parameters and it is performed on one or more frames from an equilibrated (classical or, better, *ab*

initio) MD trajectory. This choice is done with the aim of sampling a physically significant configuration in the phase space and is to be preferred with respect to the use of random configurations, for example those obtained from the well-known Packmol[186] package.

After having assigned a starting value to the charge of each atomic species (for example, those obtained from an *in vacuo* RESP fitting), the QM/MM-RESP[104] charges are subsequently calculated for each molecule of the system. As previously outlined, calculating QM/MM-RESP charges means to describe the chosen molecule with an *ab initio* approach (the HF[9]/6-31G*[187] level of theory in the original article but we also adopted the B3LYP[188, 45, 42]/TZVP[189, 190] level, obtaining very similar results) and the rest of the system as composed by point charges (those assigned at the very beginning of the procedure). The presence of the surrounding point charges will, obviously, polarize the molecule and determine an electronic distribution different from that obtainable through an *in vacuo* calculation.

After having completed the cycle over all the molecules of the system, the calculated charges are averaged (for each atomic species) and compared to the charges used in the MM part. If they are equal (within a given threshold), the process is completed and the $\{q^{gen}\}$ values are obtained. Otherwise, the calculated averages are substituted to the charges assigned to the MM part and the process is repeated until the convergence is achieved. After having attained self-consistency, the electric fields over the bonds are calculated using the freshly determined $\{q^{gen}\}$.

The process is depicted in the first part of Figure 4.10.

At this point, two different approaches can be undertaken: charge matching (CM) or force matching[191] (FM).

In the first case, a new cycle over all the molecules of the system is performed in order to calculate their QM/MM-RESP charges (now using the final $\{q^{gen}\}$ values for the MM part). A linear regression is eventually performed using the QM/MM-RESP charges as scalar response, the Equation 4.2 as the functional expression and $\vec{E}(\vec{r}_{h_b}) \cdot \hat{r}_b$ as explanatory variable. The results are the sets of $\{q^0\}$ and $\{\alpha\}$ values which best permit to retrace the accurate *ab initio* QM/MM-RESP charges.

In the second case, instead, accurate electronic structure (in our case we adopted the Density Functional Theory, but other approaches are also fully valid) forces are calculated for the system on which the $\{q^{gen}\}$ parameterization has been performed and employed as target for an optimization having the $\{q^0\}$

and $\{\alpha\}$ values as unknowns. Cost function is defined as:

$$\chi^2 = \sum_{i=1}^{N_{atom}} \left| \vec{f}_i^{MD} - \vec{f}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{f}_i^{DFT} \right| \right) \quad (4.24)$$

following the prescriptions suggested by Masia, Guàrdia, and Nicolini [192] concerning the definition of the w weighting function (namely aiming at uniformly sampling all the possible forces, even those associated with uncommonly strong or weak interactions that would be underestimated without employing a weighting function).

The total atomic forces are composed by a sum of different terms, some of them being characterized by a considerable strength, such as those connected to the 1-2 bonded interactions. The strength of these contributions, that are left untouched by the changes of the polarization parameters, can "hinder" the electrostatic contribution to the force and negatively affect the cost function minimization process. In order to counterbalance this issue, the cost function reported in Equation 4.24 can be modified by the addition of two terms, completely analogous to the first but focused on molecular forces or torques instead of the atomic forces. In the two former quantities, in fact, the contribution coming from the intramolecular forces, such as the bonded interactions, is canceled by the application of Newton's third law, permitting to highlight the role of the intermolecular electrostatic interactions. The obtained cost function is shown in Equation 4.25.

$$\begin{aligned} \chi^2 = & \sum_{i=1}^{N_{atom}} \left| \vec{f}_i^{MD} - \vec{f}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{f}_i^{DFT} \right| \right) + \\ & \sum_{i=1}^{N_{mol}} \left| \vec{F}_i^{MD} - \vec{F}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{F}_i^{DFT} \right| \right) + \\ & \sum_{i=1}^{N_{mol}} \left| \vec{\tau}_i^{MD} - \vec{\tau}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{\tau}_i^{DFT} \right| \right) \end{aligned} \quad (4.25)$$

For sake of simplicity, the coefficients of the resulting linear combinations are considered to be equal but other choices are equally valid.

This last part of the parameterization process is shown in the second part of Figure 4.10.

The molecular charge conservation over the time is a basic requirement for a polarizable force field, as pointed out in subsection 2.4.5.1. In the F-RESP method, the conservation can be enforced by applying on the freshly optimized

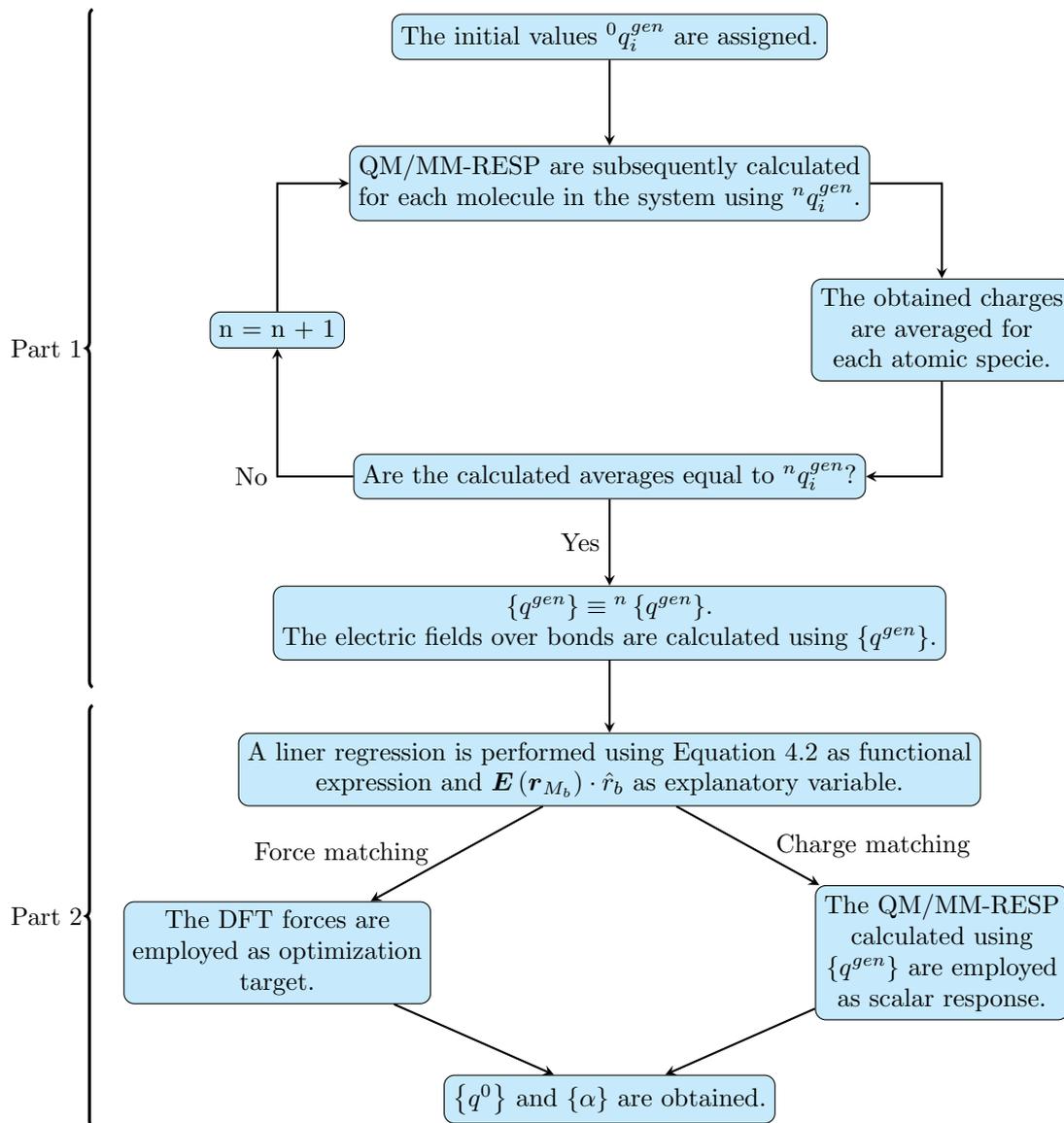


Figure 4.10: The parameterization process.

parameters the following relations:

$$\begin{aligned}
 \alpha_b^i &= \alpha_b^i - \frac{\sum_{j \in M_i} \alpha_b^j}{N_i} \\
 \kappa_b^i &= \kappa_b^i - \frac{\sum_{j \in M_i} \kappa_b^j}{N_i} \\
 \zeta_a^i &= \zeta_a^i - \frac{\sum_{j \in M_i} \zeta_a^j}{N_i} \\
 \delta_d^i &= \delta_d^i - \frac{\sum_{j \in M_i} \delta_d^j}{N_i}
 \end{aligned} \tag{4.26}$$

where N_i is the number of atoms contained in the molecule M_i . The proof for this procedure is trivial and it is worth to note that, if the parameters already ensure molecular charge conservation, its application will not alter them (being the right-hand summations equal to zero). In all the following of the Thesis, the parameters are assumed to be already corrected (thus, those with ' apex in Equation 4.26).

Another possible procedure for enforcing the molecular charge conservation is to "hardcode" it in the parameterization process. That can be easily done through the use of constraints in the force matching approach, where a numerical minimization of a cost function is performed.

Other approaches have been attempted for the parameterization task, even by slightly changing the F-RESP charge expression reported in Equation 4.2. In place of the polarization coefficient previously described, having $\left[\frac{q^2 \cdot t^2}{m \cdot l}\right]$ as dimensionality, it was tried to employ the real isotropic polarizability α (a "bond" polarizability in this case), having $\left[\frac{q^2 \cdot t^2}{m}\right]$ as dimensionality and connected to the induced dipole by the relationship:

$$\vec{\mu} = \alpha \cdot \vec{E} \tag{4.27}$$

As it can be appreciated, the two expressions for α only differs by the factor $\frac{1}{l}$, namely the inverse of a length; while the dot product in Equation 4.2 produces a charge, that in Equation 4.27 produces a dipole. It was chosen, in order to fill the gap between the two expressions, to employ $\frac{1}{|r_b|}$ as corrective factor, meaning that the induced dipole is imposed to have a length equal to the considered bond.

Two different approaches were undertaken with the aim of obtaining the bond polarizabilities: the first was to recover them from the decomposition of the experimentally measured polarizability tensor of the molecule, taking inspiration from the works of Long [193], Smith and Mortensen [194] and Guha *et al.* [195]. This kind of approach is however unfortunately strongly limiting: the number of independent values of the polarizability tensor (six, given its sym-

metric nature) that have to be inserted in a system of linear equations permits in fact to deal with compounds having a maximum of three different bond types (since, for each bond, two polarizabilities are obtained, namely the parallel and the perpendicular one).

The second method is based on the bond polarizability formula derived by Macchi and Krawczuk [196] in the field of Raman spectroscopy. They in fact express the bond polarizability as a function of the polarizabilities of the atoms at the edges of the bond:

$$\alpha_{A-A'} = \vec{r}_{A-A'}^T \cdot (\alpha_A + \alpha_{A'}) \cdot \vec{r}_{A-A'} \quad (4.28)$$

This kind of approach has, with respect to the former, the great advantage of being not limited by the size of the molecule. The atomic polarizabilities can be obtained by numerically deriving the atomic dipole moment with respect to the applied electric field, thus reverting Equation 4.27.

The calculation of the atomic dipole moments has been undertaken using both electrostatic potential-based population analyses like RESP[104] or HLY[197] (as implemented in Gaussian 16 package[14]) and electronic-density based ones, in particular the powerful Atoms In Molecules (AIM) technique by Bader [198].

Eventually and as already extensively done elsewhere, the use of Machine Learning approaches for the parameterization procedure has been attempted. In particular, their use has been figured in order to overcome the rigid linear relationship between the fictitious electric field projections over bonds and the atomic charges that is contained in Equation 4.2, thus permitting to model more complex functional forms.

However, all the previously described trials did neither produce any increase in the accuracy of the parameterization (quantified as in Equation 4.24) nor an higher accuracy in the simulation test results and both the trials were consequently discarded without performing productive simulations with them.

4.4 F-RESP implementation and performances

As previously stated, the F-RESP polarizable model was implemented during the course of my PhD into the C/C++[199] LAMMPS[178] classical Molecular Dynamics code. The package was chosen due to its large diffusion, its being open-source and, last but not least, the ease to customize it.

Following the LAMMPS philosophy, four new *fix* have been defined; these are *fix_fresp*, *fix_fresp/ewald*, *fix_fresp/dsf* and *fix_fresp/mdsf*. The former (*fix_fresp*)

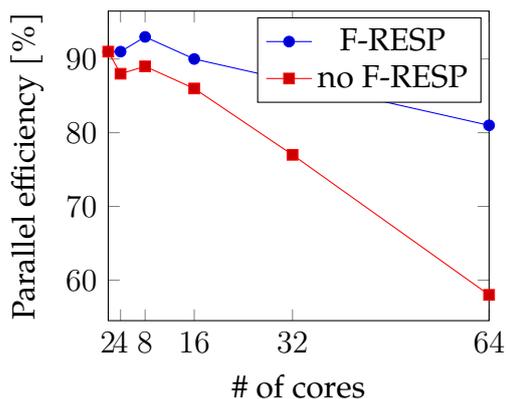
is the parent class and the remaining ones, quoted according to the method used for the calculation of the fictitious electric field (*ewald* is self-explanatory while the DSF and mDSF acronyms are employed, in the respective articles[180, 183], for referring to the Fennell and Waibel potentials), inherit from it the common methods. A guide for using this *fix* along with some example input files can be found in github.com/FrancescoCappelluti/lammps/tree/F-RESP. Additional fixes called *fix_fresp/aidsf* and *fix_fresp/aimdsf* have been also defined, in order to be employed with the polarization coefficients defined as bond polarizabilities, as described in subsection 4.3.5, but as previously written they have not been used for the productive simulations.

The MD simulations have been performed on the Tier-0 cluster Marconi of CINECA, using A1 and A2 partitions, and on the Tier-1 cluster Galileo. Marconi-A1 partition and Galileo are systems whose computing nodes are 2x18-cores Intel®Xeon E5-2697 v4 (Broadwell)@2.30 GHz and each node has 128 Gb RAM memory. Marconi-A2, instead, is composed by 1x68-cores Intel®Xeon Phi 7250 (Knight Landings)@1.4 GHz nodes, each of them having 112 Gb RAM memory. LAMMPS, in its 2 Feb 2019 version, has been compiled using Intel®MPI 2018 and MKL®[103] 2018 libraries and -O2 optimization level of Intel®C++ compiler. The MKL® libraries can be used by the *fix* for speeding up the computational intensive calculations (in particular the exponential, trigonometric and error functions) although their use is not mandatory.

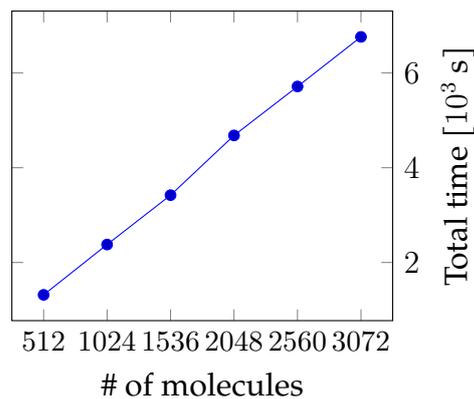
The speed ratio between the MD simulations with or without enabling the F-RESP charge variation has been calculated performing three 250 ps long simulations for each case (for a box containing 2880 TIP3P[200] water molecules). The obtained performances (expressed as $\frac{\text{simulated time}}{\text{computing time}}$) have been then averaged and divided, obtaining that the simulations with enabled F-RESP and DSF potential used for the fictitious electric field evaluation are roughly 5 times slower.

In Figure 4.11a, the parallel efficiencies (defined as the ratio between the observed and the maximum theoretically predictable speed-up) with or without enabled F-RESP are shown for 250 ps long simulations. As it can be seen, the F-RESP routines (which are the most computationally intensive ones and, thus, those imposing the overall rate) scale better than the pure LAMMPS ones, being the parallel efficiency of the calculations with enabled F-RESP always higher than the simulations with fixed charges.

In Figure 4.11b, the performances of different LAMMPS simulations (performed for 250 ps over 36 Broadwell cores) with enabled F-RESP are reported



(a) Parallel efficiencies of LAMMPS with and without enabled F-RESP.



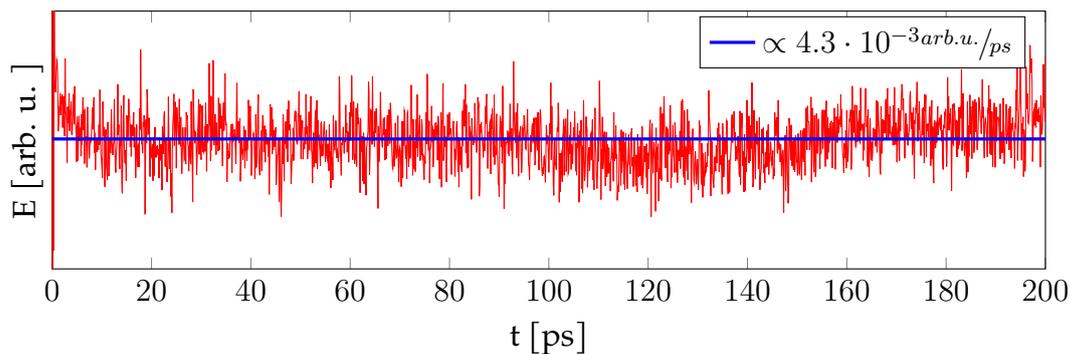
(b) Scaling of LAMMPS with F-RESP with respect to the number of molecules.

with respect to the number of water molecules contained in the MD box.

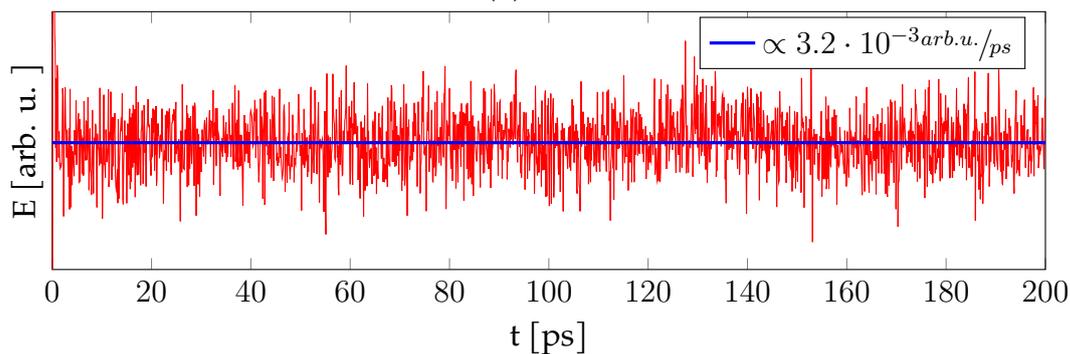
The scaling of the simulation time with respect to the number of molecules is roughly linear. Assuming, as done before, that the F-RESP routines impose the overall ratio (that is taken into account for the plot), an $O(N)$ scaling for the F-RESP/DSF can thus be stated. When Ewald summation is employed for the fictitious electric field evaluation, instead, an $O(N^{\frac{3}{2}})$ scaling is demonstrated to hold (see subsection 6.2.1).

The value of the total energy of the simulated box with respect to the time of a 200 ps long trajectory in the NVE ensemble is reported in Figure 4.12 for two simulations, performed on the same system but respectively with and without F-RESP charge fluctuation enabled. As can be seen by comparing the slope of the fits, the energy drift is small and, most importantly, comparable, meaning that the F-RESP algorithm does not deteriorate the charge conservation.

It has eventually to be noted that the choice (done for the sake of portability) of enclosing almost all the new code into the newly defined *fix* source files could result in a slight decrease in performance connected to the recalculation of some quantities that were already computed elsewhere in the code without being accessible.



(a) F-RESP.



(b) No F-RESP.

Figure 4.12: Energy vs time in a NVE simulation.

4.5 F-RESP simulation of liquid water

4.5.1 Parameterization

The parameterization of the force field has been performed on the top of the rigid three-site TIP3P[200] water model. It is one of the first model to be proposed and it has been chosen for its extreme simplicity, being bond stretching and angle bending, which could complicate the parameterization process, not allowed.

Both the Ewald summation and the Fennell potential approaches have been employed for the calculation of the fictitious electric field. First, the polarization parameters have been determined through the charge matching approach. The comparison between the sets of QM/MM-RESP charges employed as target for the least square minimization and the corresponding F-RESP charges (obtained employing the Ewald summation method for the fictitious electric field calculation) are show in Figure 4.13.

The coefficient of determination R^2 [201] has been chosen as a measure of quality and its values are shown in the plots. The remarkable values attained

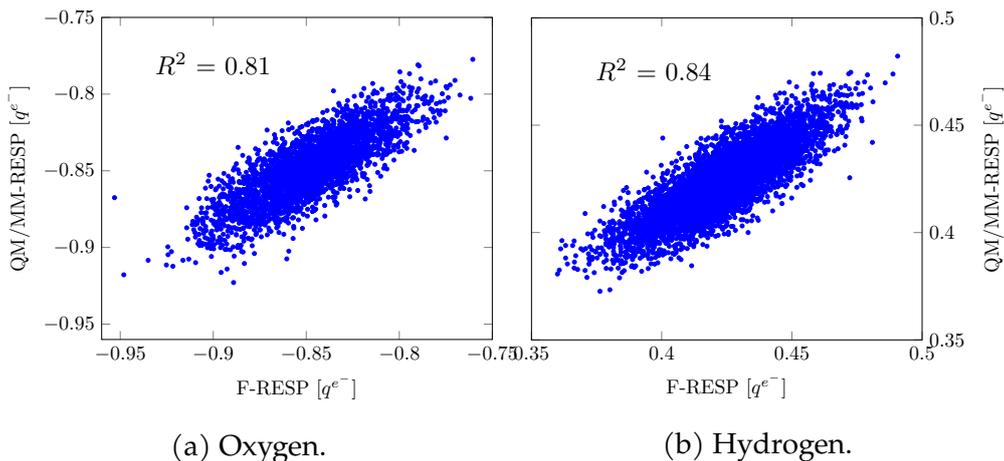


Figure 4.13: Water HF/6-31G* QM/MM-RESP vs F-RESP atomic charges (with parameters obtained from charge matching) for oxygen (left) and hydrogen (right).

by the coefficient testify that a significant correlation exists between the atomic charges and the fictitious electric field projections over the covalent bonds, therefore certifying the goodness of the assumption underlying the force field, summarized in the charge variation function reported in Equation 4.2.

After the success of the charge matching based parameterization, the force matching approach was also applied. The correlation plots between the atomic forces calculated at the DFT (using PBE[137]-D3[202] functional) and the F-RESP levels are reported in Figure 4.14 and, as in the previous case, the coefficient R^2 is employed for checking the accuracy of the results.

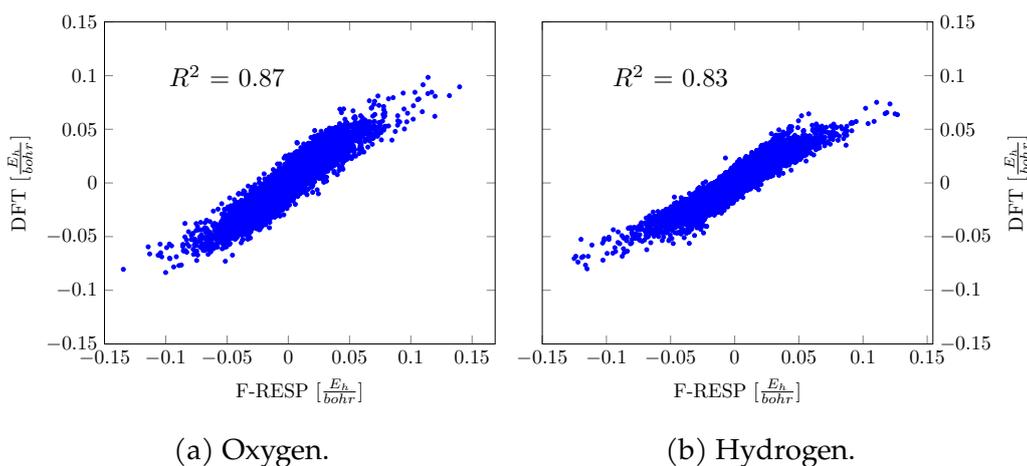


Figure 4.14: Water PBE-D3 vs F-RESP atomic forces (with parameters obtained from force matching) for oxygen (left) and hydrogen (right).

In this latter case, the correlation with the *ab initio* target quantities is even

better than when charge matching is employed for the parameterization. For the atomic forces of the oxygen atom, a correlation coefficient of 0.9 was almost touched.

Each atomic charge has been let free to depend on the fictitious electric field projection over the two bonds of the molecule. With both the approaches, the two hydrogen atoms have been constrained, due to the evident symmetry reasons (having the molecule C_{2v} point-group symmetry), to have specular parameters. The obtained parameters are reported in Table 4.6 and Table 4.7, respectively using the charge and the force matching procedure.

For sake of completeness, in Table 4.8 shows the parameters of the underlying TIP3P water model that are left untouched (namely those not concerning the electrostatic side).

Table 4.6: F-RESP parameters for TIP3P water - charge matching.

	O	H_1	H_2
q^{gen}	-0.834	0.417	0.417
q^0	-0.664	0.332	0.332
α_{O-H_1}	2.783	-2.783	0.0
α_{O-H_2}	2.783	0.0	-2.783

The parameters are reported in [a.u.].

The parameters obtained from the charge matching approach are, concerning q^{gen} and q^0 sets of charges, not far from those of the non-polarizable force field and quite easy to rationalize when referred to the α coefficients. Knowing that the fictitious electric field projections along the O-H bonds virtually always have negative sign (due to water structural organization through hydrogen bonds and to how the bond vector is defined in the simulation code), the positive sign of the coefficient with respect to the oxygen atom and the negative sign with respect to hydrogen ones produce a charge variation that makes the former atom more negative and the latter more positive, namely what could be intuitively expected from the polarization process.

When the force matching approach is used for the parameterization, on the contrary, the situation is less straightforward than in the previous case. The values of the set of generating charges (calculated at the B3LYP/TZVP level of theory) are very similar to the former calculated using the classical HF/6-31G* approach. At variance, the $\{q^0\}$ parameters significantly increase their magnitude, assuming values that are out of the range of the commonly employed fixed-charge force fields. Concerning the α polarization coefficients, it is in-

Table 4.7: F-RESP parameters for TIP3P water - force matching.

	<i>O</i>	<i>H</i> ₁	<i>H</i> ₂
q^{gen}	-0.848	0.424	0.424
q^0	-1.478	0.739	0.739
α_{O-H_1}	-2.473	0.951	1.522
α_{O-H_2}	-2.473	1.522	0.951

The parameters are reported in [a.u.].

Table 4.8: TIP3P water parameters.

m_O	15.9994 a.m.u.
m_H	1.008 a.m.u.
ϵ_{O-O}	0.1521 kcal/mol
σ_{O-O}	3.1507 Å
r_{O-H}^0	0.9572 Å
θ_{H-O-H}^0	104.52°

The Lennard-Jones parameters for the O-H and H-H interactions are zero.

interesting to observe that they change their sign with respect to those obtained from the charge matching. Taking into account the preferred sign possessed by the fictitious electric field projection, cited a few lines above, the change in sign comes out in a charge variation that compensates the high $\{q^0\}$ values. Moreover, the charges of the hydrogen atoms are strongly influenced by the electric field exerted along the bond whom they are not part, even more than from the bond they are concerned in, while in the previous set of parameters this influence was found to be null. Such findings are counter-intuitive and quite difficult to explain. Concerning the "unexpected" sign of the polarization coefficients, a possible interpretation is that they would improperly represent the charge transfer process between the oxygen and hydrogen atoms connected by an hydrogen bond. The high $\{q^0\}$ value would therefore implicitly include the effects of increase of the charge magnitude by molecular polarization and the F-RESP charge variation function would act by reducing the charge magnitudes as much as the electric field magnitude increases (and thus the distance decreases), taking into account the inter-penetration of the electron clouds of the two atoms (the one exerting the electric field and the one whose charge is varied).

Supposing that these effects are an artifact of the force matching approach (which is well known for producing parameters far from physically reasonable values) and of the system got catch in a local minimum of the cost function

hyper-surface, many other optimizations were performed. In particular, it was tried to use the global optimizer SLSQP-Basin-hopping and to start the minimization from the values obtained by the charge matching approach, adopting a very small optimization step (in order not to overtake the global minimum in the case of its being located in a very sharp basin). Nonetheless, the same set of parameters was obtained as the optimal one and this finding eventually pushed us to retain the unrestrained parameters, even in the lack of an intuitive physical comprehension.

4.5.2 Simulations

The charge fluctuation of a randomly chosen oxygen atom is represented in Figure 4.15 for a F-RESP classical Molecular Dynamics simulation and an *Ab Initio* Molecular Dynamics simulation (performed at the PBE/DZVP-MOLOPT-GTH[203] level of theory). The Hirshfeld[204] population analysis is employed for estimating the charge of the atom in the AIMD simulation. As can be seen, the magnitude of the fluctuations is fully comparable even though the response of the F-RESP model to the external perturbations is less prompt than in the rigorous *ab initio* description.

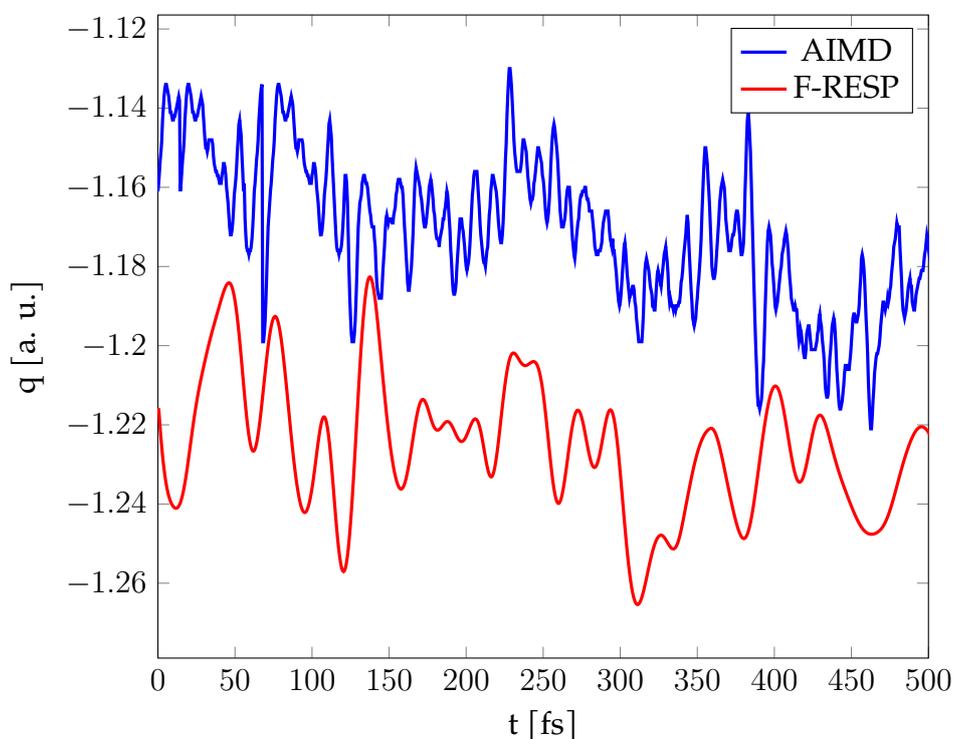


Figure 4.15: Comparison of the charge fluctuation for an oxygen atom for a F-RESP and an AIMD trajectory.

A comparison between some important experimental observables and the corresponding results from DFT and classical Molecular Dynamics simulations (using the traditional TIP3P and the new F-RESP-TIP3P water models) are reported in Table 4.9 and Figure 4.16. In particular, in the Table the average bulk density and diffusion coefficient are shown while in the Figure the comparison is performed for the oxygen-oxygen Radial Distribution Function, one of the most important and used structural observables for water.

For the DFT level, the PBE functional with D3 formalism for the accounting of the dispersive interactions was chosen. This is, in fact, the level of theory employed for the calculation of the atomic forces that have been used during the force matching parameterization procedure.

As can be seen in Table 4.9 and Figure 4.16, the F-RESP technique permits to obtain, with both the parameterization procedures presented in the current Thesis, a diffusion coefficient and a Radial Distribution Function that are much more similar to the experimental ones with respect to the results obtained from the non-polarizable TIP3P water model.

Table 4.9: Density and diffusion coefficient comparison.

	Experimental [205, 206]	PBE-D3 [207]	TIP3P [200, 208]	F-RESP- TIP3P CM	F-RESP- TIP3P FM
Density [g/cm^3]	0.997	-	0.982	-	1.018
Diffusion coefficient [$10^{-5}cm^2/s$]	2.30	0.73	5.06	2.22	2.33

Data referred to normal conditions (T = 298 K, P = 1 atm).

Concerning the diffusion coefficient, whose calculation has been performed by means of the Einstein's relation [209, 133, 134] over the mean squared displacement, the value obtained from the force matching-derived parameter set is nearly equal to the measured one, within the $\pm 1\%$ experimental uncertainty reported by Harris and Woolf [206]. The density is, instead, slightly worse than the unpolarized TIP3P one (with a 2.1% error in the former case with respect to a 1.5% one in the latter).

Eventually switching to a richly significant observable, the oxygen-oxygen RDF, the most striking difference can be appreciated by observing the appearance of the second and third peaks, that are almost absent in the Radial Distribu-

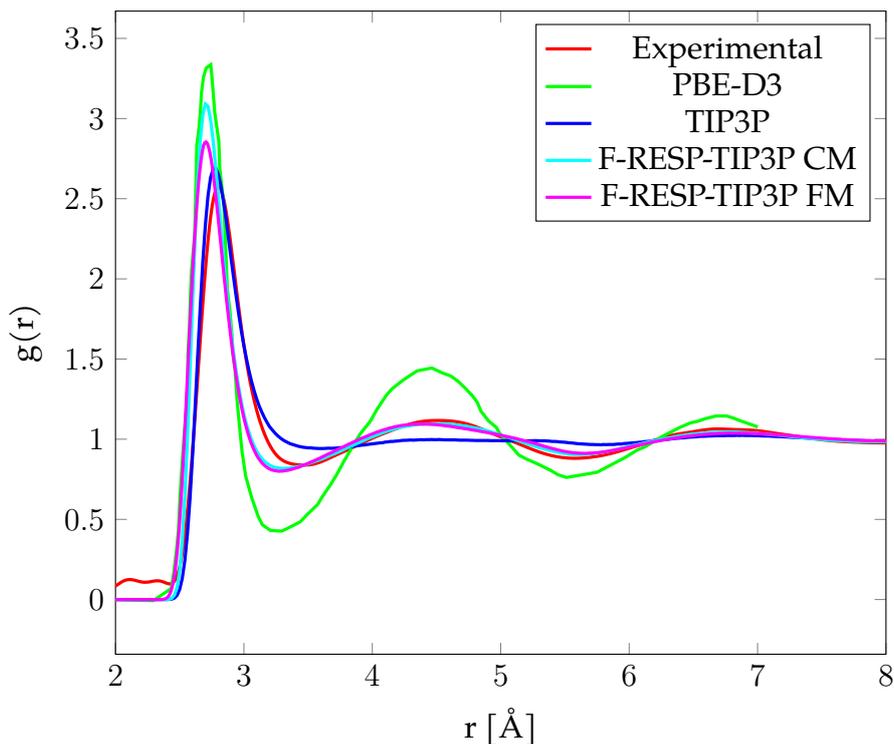


Figure 4.16: Oxygen-oxygen Radial Distribution Function comparison.
Data referred to normal conditions ($T = 298$ K, $P = 1$ atm).

tion Function of the non-polarizable TIP3P model underlying ours. The matching of these peaks with the experimental data is remarkable, both in position and height. The location of the first peak, on the contrary, is slightly more compressed than in experiment (and when using the simple, non polarizable TIP3P model); with both the sets of parameters, a shift to the left slightly smaller than 0.1 Å is observed. An increase of the peak height, quite significant when the set of charge matching-derived parameters is employed, can also be observed.

As shown in Figure 4.17 and, in a more detailed way, in section 6.3 and section 6.4, these effects are closely related to the hyper-polarization phenomenon already quoted and the position and shape of the RDF peaks are strongly influenced by the interplay between the A and B parameters of the damping function and the α coefficient (which expresses the strength of the polarization). As a general observation, we found that the increase of the steepness of the damping function produces the shift of the first RDF peak to the right and the augment of its height.

It is very pleasant to observe that the results produced by the F-RESP-TIP3P model generally supersede those obtained both by classical non-polarizable and *Ab Initio* Molecular Dynamics. In particular, the diffusion coefficient and the height of the second and third peak of oxygen-oxygen RDF are correctly caught

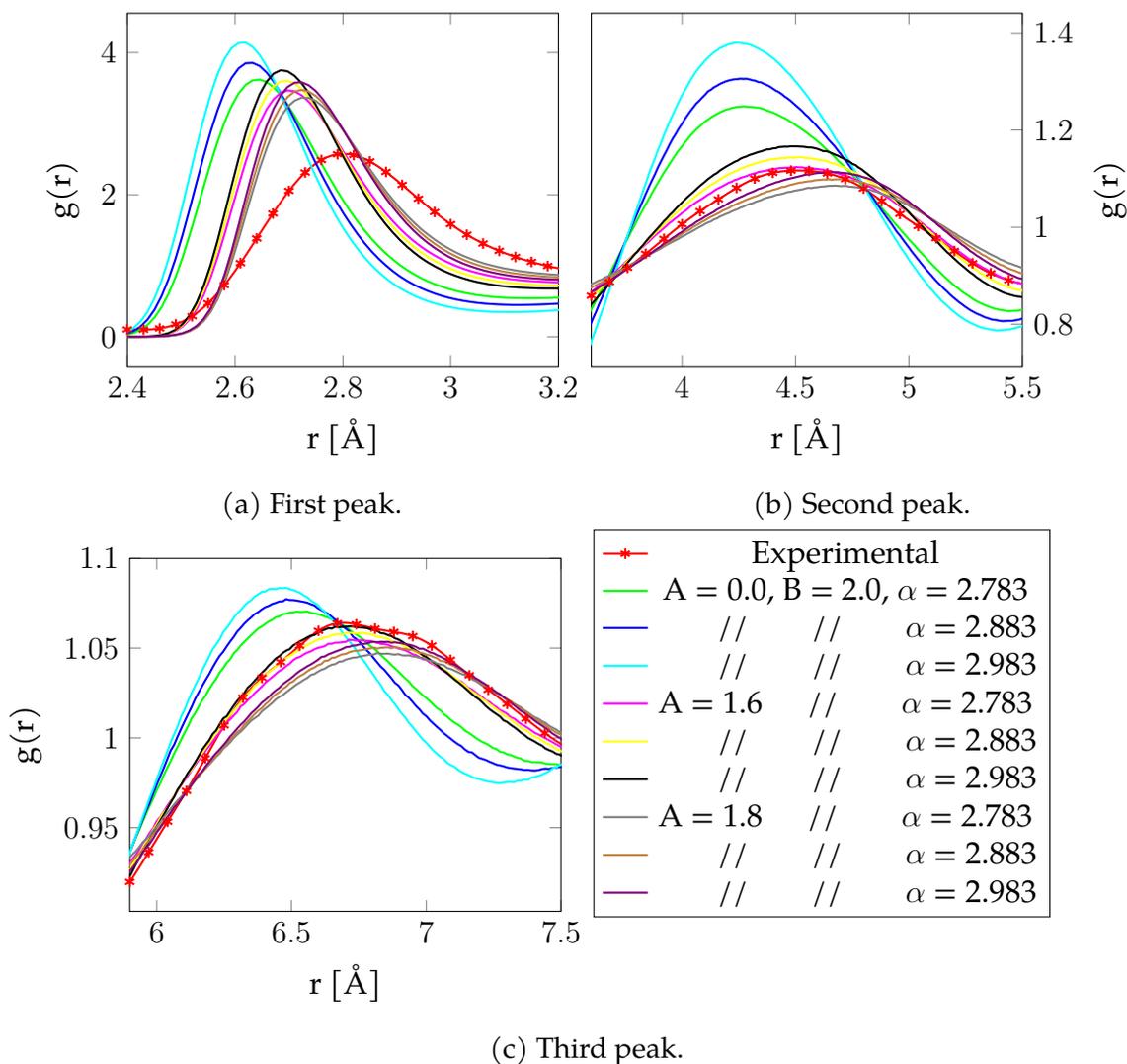


Figure 4.17: Changes in O-O RDF with respect to the damping (sinusoidal) function parameters - Fennell potential TIP3P charge matching.

by our approach, whereas the traditional TIP3P model and the *ab initio* approach with PBE-D3 functional alternatively strongly over- or underestimate them. We could therefore find a good balance between the two approaches, correcting their respective weaknesses.

As a final point, the data reported here clearly show that the force matching parameterization approach supersedes charge matching for both the structural and dynamical point of view, as can be realized by inspecting the height of the first peak in Figure 4.16 or the diffusion coefficient in Table 4.9. The better description is for all likelihood produced by the targeting of well defined physical quantities as the atomic forces with respect to the non rigorously defined atomic charges.

4.6 F-RESP simulation of ionic liquids

The importance of the polarization phenomenon for the ionic liquids class of compounds cannot be underestimated, as largely testified in the scientific literature[210, 211]. For this reason, it was chosen to apply the F-RESP formalism to perform polarizable Molecular Dynamics simulations on the two ionic liquids methylammonium nitrate and 2-methoxyethylammonium nitrate represented in Figure 4.18.

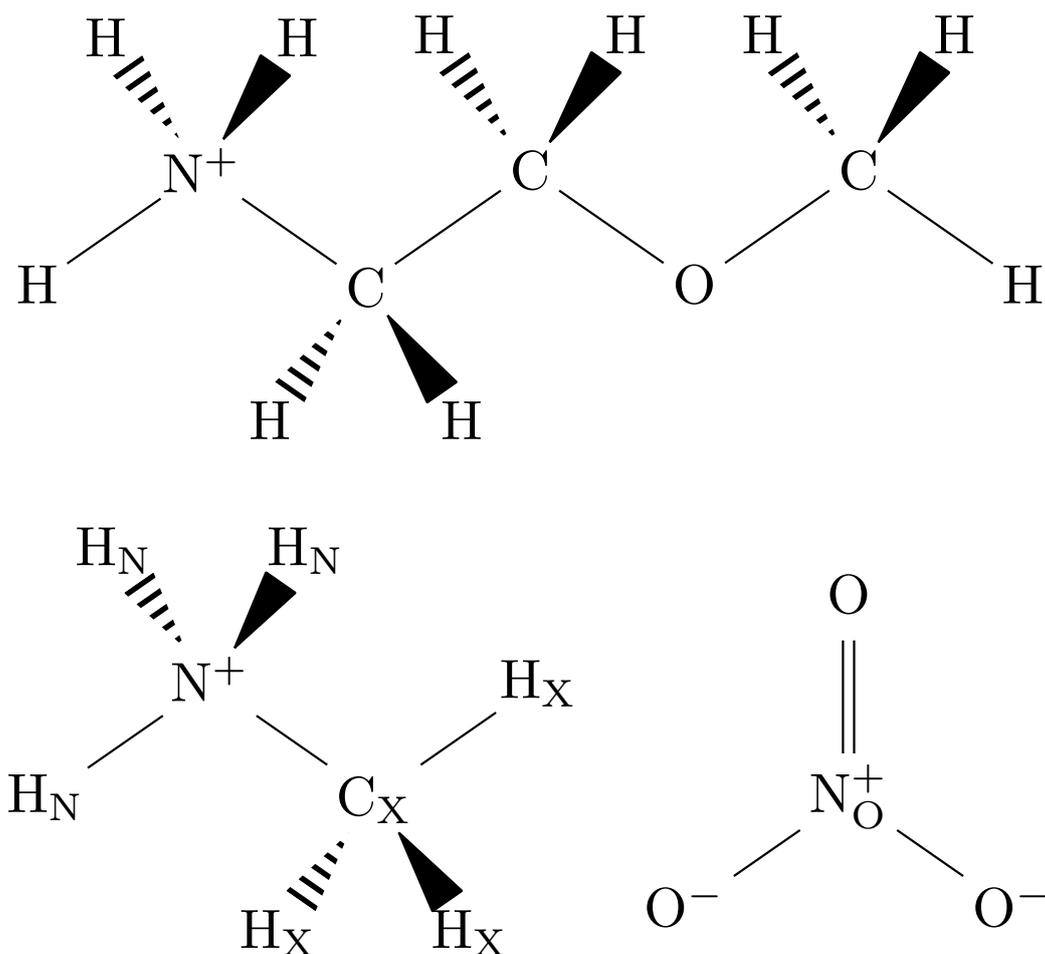


Figure 4.18: 2-methoxyethylammonium (top), methylammonium (bottom left) and nitrate (bottom right) ions. For methylammonium and nitrate, the GAFF atom types are depicted.

These ionic liquids were chosen for their structural simplicity and for the availability of both experimental and computational data[212, 213] (produced by accurate *Ab Initio* Molecular Dynamics simulations). The availability of such data would in fact permit to check the reliability of our approach for the description of the polarization process.

4.6.1 Parameterization

The two ionic liquids have been parameterized using both the charge and the force matching approaches described in subsection 4.3.5.

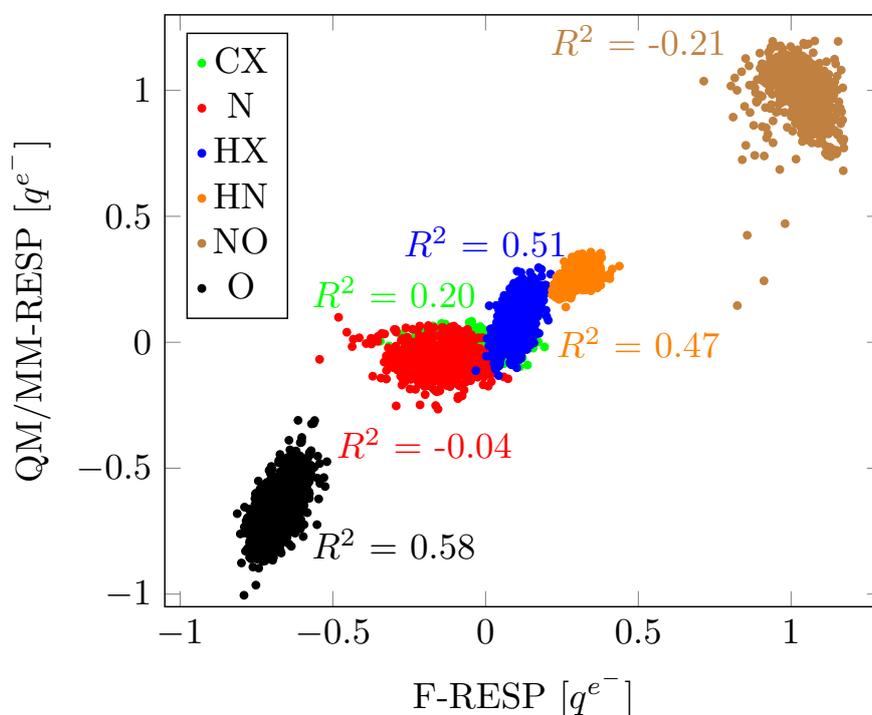


Figure 4.19: Methylammonium nitrate HF/6-31G* QM/MM-RESP vs F-RESP atomic charges (with parameters obtained from charge matching) for the different GAFF atomic types.

At variance with water, described in the previous chapter, the matching between QM/MM-RESP charges and the F-RESP ones is quite scarce, as shown in Figure 4.19 for methylammonium nitrate.

Many trials were done, making the F-RESP charges depending either on a more or less vast set of fictitious electric field projections. In particular, it was tried to make charges depending only on the electric field projections over the bonds whom they belong to or, on the contrary, on all the covalent bonds of the moiety they are contained in. The first kind of approach has, moreover, the advantage of impeding the unphysical long-range charge transfer, a common drawback of the polarizable force fields based on the fluctuating charges formalism. In all these cases, the results were not satisfying, as one can see by looking at the low R^2 determination coefficients reported in Figure 4.19. Such coefficients testify that, for this kind of molecule, the high correlation existing between the atomic charges and the electric field projections over the covalent bonds does not completely explain the polarization processes taking place in the

system. More specifically, the behaviour of some atomic types can be described in conjunction with the electric field projections with a sufficient degree of accuracy (for example concerning the oxygen atoms of the nitrate ions) but, in other cases, the correlation is completely absent. Consequently, the subsequent Molecular Dynamics simulations performed using the obtained polarization parameters on top of the GAFF force field[214] led to severe issues preventing to produce a complete trajectory.

After having ascertained that the charge matching approach is not able to produce a reliable set of polarization parameters, it was decided to employ the force matching method, which demonstrated its superiority in performance with respect to the former in the parameterization of the TIP3P water model.

At variance with the TIP3P model, which is a rigid system, the two considered ionic liquids have a high conformational flexibility, even when the SHAKE[215] algorithm is employed for constraining the stretching motions of the hydrogen-containing bonds or even the bending motions of the angles they are involved in. For this reason, the dependency of the fluctuating atomic charges on the variations of the geometrical descriptors outlined in Equation 4.3 was enabled, and the cost function reported in Equation 4.25 was made depending on $\{\kappa\}$, $\{\zeta\}$ and $\{\delta\}$ variables too:

$$\chi^2(\{q^0\}, \{\alpha\}, \{\kappa\}, \{\gamma\}, \{\delta\}) = \sum_{i=1}^{N_{atom}} \left| \vec{f}_i^{MD} - \vec{f}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{f}_i^{DFT} \right| \right) + \sum_{i=1}^{N_{mol}} \left| \vec{F}_i^{MD} - \vec{F}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{F}_i^{DFT} \right| \right) + \sum_{i=1}^{N_{mol}} \left| \vec{\tau}_i^{MD} - \vec{\tau}_i^{DFT} \right|^2 \cdot w \left(\left| \vec{\tau}_i^{DFT} \right| \right) \quad (4.29)$$

In particular, the dependence on the value of the improper angle of nitrate ion, whose strength is expressed by the coefficient δ , was demonstrated in a previous unpublished study accomplished in our research group to have a non-negligible importance for correctly catching the polarization process underwent by that moiety.

Also in this case, a huge number of trials were carried out, by making the cost function depend on a considerable variety of sets of parameters. In particular concerning 2-methoxyethylammonium nitrate, it is unphysical and overfitting-prone to make a charge depending on a descriptor which is 4 or 5 Å far from it. It was therefore generally chosen to prevent that by allowing a charge to only depend on the geometrical descriptors it is contained in.

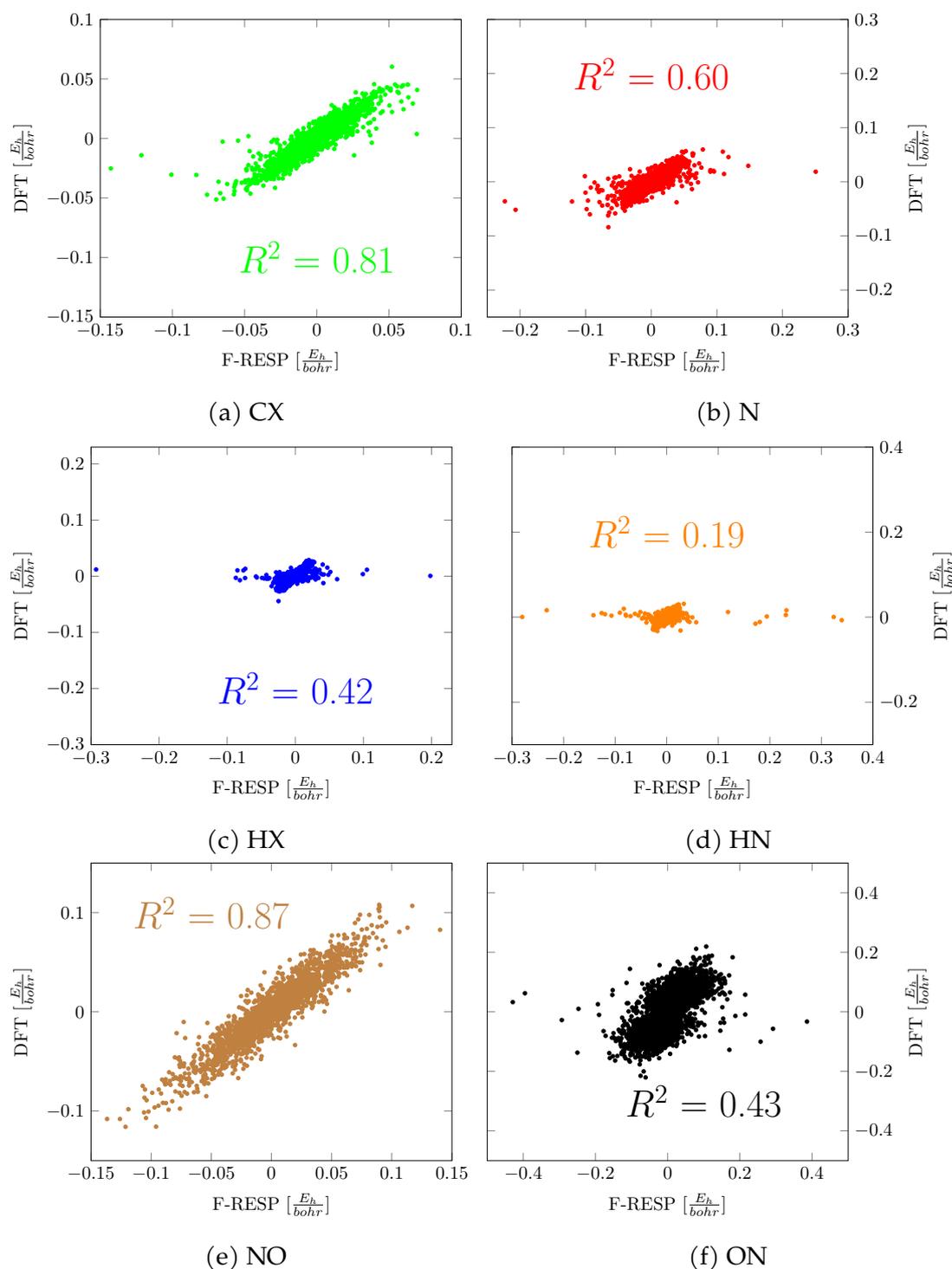


Figure 4.20: Methylammonium nitrate PBE-D3 vs F-RESP atomic forces (with parameters obtained from force matching) for the different GAFF atomic types.

It was also tried to make the cost function depending on the parameters that express Van der Waals interactions. Some of the forces (*i.e.* the Debye one) contained in this family of interactions are in fact connected to polarization and it is therefore theoretically correct to modify them when an explicit term (the

F-RESP charge variation, in this case) is added for describing it. These interactions are accounted by the underlying GAFF force field using the Lennard-Jones' expression reported in subsection 2.4.1.2 (with the neglect of 1-2 and 1-3 interactions and the scaling of 1-4 interactions by a factor smaller than one) and, in order for σ and ϵ to maintain reasonable values (bigger than zero) throughout the minimization procedure, the use of appropriate constraints has been required. The correlation plot between the DFT and the F-RESP atomic forces for an optimization of this type is reported in Figure 4.20. As can be seen, while the atomic forces of some atomic types can be caught very well (with R^2 determination coefficients of almost 0.9), the result of the parameterization procedure is for other cases very poor.

Unfortunately, despite of the high number of trials, it was not possible to obtain a valuable set of parameters, in spite of the significant decrease of the cost function during the parameterization processes. In most of the cases, in fact, our Molecular Dynamics simulations led to unstable trajectories that irreversibly stopped after few picoseconds due to severe issues, even in the absence of any externally imposed thermal excitation (by means of the thermostat algorithms outlined in section 2.4). In rare instances, some sets of parameters permitted, when employed, to perform a correct equilibration and eventually to produce a stable productive trajectory. The physical observables obtained from these trajectories (such as, for example, the structure factor $S(Q)$) were however worse than the corresponding ones obtained from a conventional non-polarizable MD simulation.

In conclusion, the improvement over the fixed-charge description was not achieved for both the considered ionic liquids in my study. Some explanations can be proposed to explain the failure, indeed unexpected, especially if thinking about the good performance of the method for the simulations on liquid water. The main could probably be associated with the lacking account of intermolecular charge transfer[216, 217] by the F-RESP force field. This process, that produces an apparent reduction of the magnitude of the atomic point charges, has a considerable importance for this class of compounds. As a demonstration for this, the most simple way of accounting for mean-field polarization in these systems (and obtaining a more realistic description of the dynamical properties) is represented by the counter-intuitive scaling of the atomic charges by a factor smaller than one[218]. This kind of phenomenon can unfortunately not yet be accounted by the current F-RESP charge expression (Equation 4.3). The inclusion of a term able to describe this process would therefore represent a much

interesting further development topic.

Chapter 5

Conclusions

In this Thesis, the two main research topics carried on in the course of my PhD are presented. They are the development of the Extended Broken Symmetry (EBS) approach for electronic structure calculations on low spin open-shell compounds and the development of the Fluctuating-RESP (F-RESP) force field for classical polarizable Molecular Dynamics simulations. In both these cases, important results have been obtained and robust approaches, ready for being made available for the scientific community, have been established. In particular for the latter case, much effort has been devoted to the implementation of the algorithm, its parallelization and its inclusion into a community Molecular Dynamics code. Nevertheless, as always happening in Science, the attained achievements provide just the basis on which new questions and further developments can be settled up, opening new roads and interests for the research. All this is now going to be more extensively reviewed in the following two sections, referred to the EBS and F-RESP projects.

5.1 EBS

The theoretical description of the spin frustrated, low spin open-shell systems has always represented a challenge for chemists and physicists, due to the spin contamination problem that arises when, as traditionally done, a single Slater determinant is employed for expressing the wavefunction of the species. This phenomenon produces a qualitatively wrong description of the studied systems and an artificial raise in the estimated energy reflecting also on the geometrical gradient, leading the structural optimizations to land more or less far from the true energy minima. Many techniques have been developed with the aim of facing this issue without requiring the employ of the qualitatively correct, however

computationally expensive, multi-reference techniques but they are restricted to systems with only two spin centres or do not permit to perform geometrical optimizations.

For such a reason, few years ago the development of the generalized Extended Broken Symmetry (EBS) technique was undertaken in the group of Prof. Guidoni. In 2017, the first article about the method was published. Within the EBS framework, the gradients on the nuclei are calculated using a linear combination of several Broken Symmetry gradients, with coefficients that permit to recover the correct spin-uncontaminated value. The resulting relaxed geometry is fully consistent with the magnetic coupling constants, for the inclusion of the latter in the HDvV Hamiltonian. The technique is fully generalized, therefore permitting to study systems that contain an arbitrary number of spin centres. The approach therefore permits to alleviate the effects of symmetry breaking on the geometrical properties of the studied systems.

In the course of my PhD, I further extended the method to the calculation of the Hessian matrix of the uncontaminated low-spin ground state. This was achieved by using a linear combination of the Broken Symmetry Hessian matrices, analogously to what did for the calculation of the gradient.

Also in this case, it was chosen to test the newly developed method on molecules belonging to the iron-sulfur clusters class. This category of chemical species is in fact largely characterized by the high spin nature of the iron atoms coming along with the global low-spin organization produced by the sulfur atoms that act as bridges between the spin centres. Moreover, these species possess remarkable importance in biochemistry and the interest is growing also for the promising role as catalytic centres for hydrogen production[219] from the water splitting reaction[220]. The two clusters $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ and $\text{Fe}_4\text{S}_4\text{Cl}_4$ were chosen for the large availability of experimental data and to reliably represent two of the main structural motifs of both natural and synthetic iron-sulfur clusters (ferredoxin-like and cubane-like).

Calculations were performed employing the OPBE, BP, B3LYP, M06 and B2PLYP DFT functionals, in order to test the two main families of DFT functionals and the newly developed double-hybrid class of functionals. SCS-MP2 calculations have also been performed with the aim of testing a fully *ab initio* approach.

Concerning the magnetic point of view, the double-hybrid functional B2PLYP has permitted to obtain a J-coupling constant value that is perfectly coinciding with the experimentally measured one. Even the extensively used hybrid functional B3LYP could return satisfying results. On the geometrical side, on

the contrary, the best match we were able to obtain with the structure determined by X-ray diffraction measures was provided by the recent hybrid M06 functional. We could therefore in general conclude that, in agreement with the first article[84] published on the EBS approach, the hybrid and double-hybrid DFT functionals give the best description of this kind of systems although we still could not find a functional equally well describing both the magnetic and geometrical features.

Concerning the most significant feature developed during my PhD, namely the Normal Mode Analysis, it was possible to observe non-negligible differences in the frequencies obtained for the same normal modes within the BS and EBS approaches for all the chosen functionals, therefore demonstrating the importance of a correct treatment of the static correlation phenomenon for the vibrational study. The visual inspection of the involved normal modes provided a further evidence of the deep connection existing between the shifts in wavenumbers and the enforced spin-symmetrization. A final comparison with the measured vibrational frequencies definitively testified that the EBS description permits to improve the agreement with experiments.

As previously written, the satisfying results obtained until now are not ultimate and further improvements can be outlined. Among the possible evolutionary paths that can be imagined, the most important is probably the further generalization of the Extended Broken Symmetry technique with the aim of being able to describe compounds containing spin centres having mixed-valence. These systems in fact require to include the contributions coming from the double exchange process[171, 221] in the effective Hamiltonian and a major revision of the code would be required. Full inclusion of the EBS approach into the ORCA code employed for the single-point electronic structure calculations (which would result in an overall speedup of the calculations and in a larger diffusion of the method) can also be envisaged.

5.2 F-RESP

The explicit description of the polarization of both atoms and of the entire molecules within a classical Molecular Dynamics simulation has in the last decades become a vibrant research topic. The enormous increase of the available computer power has permitted to deal with the problem, common for many systems ranging from the ubiquitous water to important classes of compounds such as ionic liquids, of a very poor description unveiled by the classical fixed-charge approx-

imation, used by the most widespread force fields.

The force fields that are able to take into account the polarization processes are quoted polarizable force fields and are divided into three main categories that use fluctuating atomic charges, additional Drude particles or virtual induced dipoles. These approaches, however, have in common the requirement of performing an iterative cycle for taking into account the mutual polarization between sites and this procedure slows them down significantly with respect to the traditional force fields. For such a reason, a few years ago the development of a new method for performing polarizable Molecular Dynamics simulations, the Fluctuating-RESP (F-RESP) force field, was undertaken by the group of Prof. Guidoni.

This new force field is able to include the effects of molecular polarization (due to the use of the fluctuating charges approach) without increasing the computational order of the overall simulation. The point charge of each atom varies according to the charge variation function reported in Equation 4.2 and depends on the effect of the surrounding fictitious electric field. In particular, we found a remarkable correlation between the atomic QM/MM-RESP charges and the projections of the fictitious electric field on the covalent bonds of the molecule whom the atom belongs.

The fictitious electric field used by the charge variation function is calculated using a set of parametric charges, said "generating charges", having the same positions of the atoms of the system in the current time-step but a fixed value. They therefore represent an "average value" of the atomic charges in the system that takes into account the polarization, and this peculiarity makes the algorithm much faster and stable than other polarizable force fields, avoiding the need of a self-consistent cycle for including the mutual interaction between the atomic charges.

Two approaches are available for the calculation of the fictitious electric field, that is the Ewald summation method and the Fennell method. The latter method permits to obtain an accuracy comparable with the former with a significant saving in time and the equations underlying it were entirely developed during the course of my PhD. An additional advantage of this approach is that it permits to very easily express the contribution to the virial coming from the variation of the charges, as demonstrated in paragraph 4.3.2.2.1. The exact knowledge of the virial is of crucial importance for the estimation of the pressure inside the simulation box, which in turn permits to perform simulations in the NPT statistical ensemble.

The F-RESP algorithm has been implemented into the open-source software LAMMPS in a fully productive version. It is in fact fully parallelized and much attention has been devoted to the optimization of the code for the use on HPC systems. A complete documentation, including working examples, has been also produced. The inclusion of the new feature in the main LAMMPS code will be proposed in a short time, as soon as the article about the method that is currently being written will be accepted by a peer-reviewed scientific journal.

The method was applied on liquid water and very satisfying results have been obtained. The new TIP3P-FRESP water model, produced by the application of the F-RESP charge variation function on the pre-existing TIP3P model, is able to simulate a box of water molecules in their liquid phase much more accurately than its non-polarizable counterpart (the TIP3P model), with remarkable improvements in both structural (RDF) and dynamical (self-diffusion coefficient) description, at the price of a slight increase in the computational times of the simulation.

Even though the results on liquid water were significant, many problems arose when trying to obtain the polarization parameters for other compounds, such as ionic liquids, thus testifying the complexity of the polarization phenomenon. Much work can still therefore be done for extending the method to cover a wider variety of systems.

Chapter 6

Supplementary material

6.1 EBS

6.1.1 Clebsch-Gordan coefficients calculation

The scheme[222] for the calculation of the CG coefficients required for the expression of the Hamiltonian matrix is:

$$\begin{aligned}
 \langle s_i s_j; m_i m_j | SM \rangle &= \delta_{M, m_i + m_j} \sqrt{\frac{(2S + 1)(S + s_i - s_j)!(S - s_i + s_j)!(s_i + s_j - S)!}{(s_i + s_j + S + 1)!}} \\
 &\quad \times \sqrt{(S + M)!(S - M)!(s_i - m_i)!(s_i + m_i)!(s_j - m_j)!(s_j + m_j)!} \\
 &\quad \times \sum_k \frac{(-1)^k}{k!(s_i + s_j - S - k)!(s_i - m_i - k)!(s_j + m_j - k)!(S - s_j + m_i + k)!(S - s_i - m_j + k)!}
 \end{aligned} \tag{6.1}$$

where the summation on the third line is extended over all the integers k for which the argument of every factorial is non-negative[223]. Under the following conditions:

- $k \geq 0$
- $k \leq s_i + s_j - S$
- $k \leq s_i - m_i$
- $k \leq s_j + m_j$
- $k \geq -S + s_j - m_i$
- $k \geq -S + s_i + m_j$

Equation 6.1 reduces to:

$$\begin{aligned}
 \langle s_i s_j; m_i m_j | SM \rangle &= (-1)^{S - s_i - s_j} \langle s_i s_j; -m_i, -m_j | S, -M \rangle \\
 \langle s_i s_j; m_i m_j | SM \rangle &= (-1)^{S - s_i - s_j} \langle s_j s_i; m_j m_i | SM \rangle
 \end{aligned} \tag{6.2}$$

6.1.2 Finite difference gradient test

The test on finite difference gradients, necessary for checking if the approximation done in Equation 2.38 is meaningful, is reported in Table 6.1.

Table 6.1: Test of finite difference gradient.

Deviation [\AA] ^a	ΔE [cm^{-1}] ^b	E_{HS} [E_h] ^c	E_{GS} [E_h] ^d	GS_{FND}^e	GS_{engrad}^f
-0.01		-8389.1049633	-8389.1083292		
0.00	-737.959	-8389.1049676	-8389.1083300	0.000381	0.000378
0.005	-737.583	-8389.1049637	-8389.1083244	0.000802	0.000802
0.01	-737.208	-8389.1049558	-8389.1083148	0.001231	0.001227
0.015	-736.830	-8389.1049439	-8389.1083011	0.001661	0.001659
0.02	-736.466	-8389.1049278	-8389.1082834		

^a Deviations with respect to the optimized structure under the HS state. 0.00 means that this structure is the original optimized structure. 0.005 means a movement of 0.005 \AA of the first Fe atom along the x axis. The length unit is in \AA .

^b Energy difference between the HS and the GS state.

^c HS state energy calculated from ORCA.

^d GS state energy calculated with EBS method.

^e Gradient on the GS state numerically calculated by the finite difference method with the central difference scheme.

^f Gradient on GS state analytically calculated with EBS method.

6.1.3 EBS modes decomposition

An example of EBS mode decomposition is reported in Table 6.2 for the $\text{Fe}_4\text{S}_4\text{Cl}_4$ cluster with the OPBE functional (only the BS modes that have a contribution larger than 10% one are shown). The percentages are calculated as normalized squares of the dot product results.

Table 6.2: Normal modes decomposition - Fe₄S₄Cl₄ OPBE.

EBS mode	BS modes	EBS mode	BS modes
1	62% 1 + 15% 2 + 10% 3 + 10% 4	16	82% 15 + 16% 17
2	63% 4 + 18% 1 + 12% 5	17	91% 18
3	57% 2 + 17% 1 + 11% 3 + 8% 4	18	76% 19 + 18% 17
4	72% 3 + 15% 5 + 11% 2	19	50% 17 + 17% 20 + 16% 19 + 10% 15
5	54% 5 + 16% 4 + 12% 2 + 6% 6	20	95% 21
6	90% 6	21	76% 20 + 15% 17
7	69% 8 + 26% 7	22	99% 22
8	68% 7 + 22% 8	23	100% 23
9	96% 9	24	93% 25
10	60% 11 + 38% 10	25	93% 24
11	62% 10 + 37% 11	26	93% 26
12	95% 12	27	92% 27
13	98% 13	28	99% 28
14	95% 14	29	99% 29
15	97% 16	30	100% 30

6.1.4 Exchange coupling constants for Fe₂S₂Cl₄²⁻ and Fe₄S₄Cl₄

The exchange coupling constants evaluated at different levels of theory and with different spin descriptions are reported in Table 6.3 and Table 6.4.

Table 6.3: J coupling constants - $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$.

		J [cm^{-1}]
Exp[168]		-158
B3LYP	HS	-105.8
	BS	-159.8
	EBS	-174.1
BP	HS	-201.5
	BS	-356.0
	EBS	-377.3
M06	HS	-141.1
	BS	-193.5
	EBS	-205.8
OPBE	HS	-171.6
	BS	-297.8
	EBS	-328.1
MP2	HS	-47.5
	BS	-60.9
	EBS	-64.0
B2PLYP	HS	-91.2
	BS	-141.9
	EBS	-156.0

Table 6.4: J coupling constants - $\text{Fe}_4\text{S}_4\text{Cl}_4$.

		J_1 [cm^{-1}]	J_2 [cm^{-1}]	J_3 [cm^{-1}]	J_4 [cm^{-1}]	J_5 [cm^{-1}]	J_6 [cm^{-1}]
B3LYP	HS	-74.4	-71.5	-73.5	-73.4	-74.0	-73.0
	BS	-117.8	-70.2	-115.5	-114.3	-73.7	-115.7
	EBS	-123.9	-70.4	-121.1	-120.1	-73.8	-121.7
BP	HS	-114.3	-114.3	-114.2	-114.4	-114.4	-113.7
	BS	-306.1	-306.4	-274.7	-274.8	-307.1	-306.7
	EBS	-323.9	-324.1	-298.7	-298.6	-324.2	-324.2
M06	HS	-86.9	-82.6	-85.7	-84.7	-87.8	-84.9
	BS	-128.0	-77.5	-126.4	-123.5	-82.2	-125.2
	EBS	-133.0	-77.4	-131.2	-128.3	-82.1	-130.4
OPBE	HS	-126.4	-126.1	-125.7	-126.0	-126.2	-125.8
	BS	-252.0	-152.7	-252.0	-252.2	-152.6	-251.9
	EBS	-268.3	-177.7	-267.9	-268.0	-177.6	-268.2

6.1.5 EBS vs BS normal modes frequencies comparison

The comparisons between the EBS and BS vibrational frequencies are reported in a series of Tables (from Table 6.5 to Table 6.9) for the $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ and $\text{Fe}_4\text{S}_4\text{Cl}_4$ clusters.

Table 6.5: Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B3LYP and BP.

B3LYP					BP				
EBS		BS		Shift	EBS		BS		Shift
mode	ν	mode	ν		mode	ν	mode	ν	
1	37	1	27	10	1	57	1	46	11
2	72	2	63	9	2	64	2	65	-1
3	88	3	84	4	3	88	3	87	1
4	99	4	96	3	4	94	4	91	3
5	108	5	118	-10	5	99	5	100	-1
6	120	7*	101	19	6	106	6	105	1
7	124	6	105	19	7	151	8	137	14
8	143	8	135	8	8	152	7	112	40
9	166	10*	164	2	9	173	10	172	1
10	173	9	149	24	10	183	9	152	31
11	263	11	263	0	11	240	11	237	3
12	286	12, 13	-	-	12	254	12	254	0
13	289	12, 13	-	-	13	277	13	271	6
14	301	14	297	4	14	291	14	290	1
15	306	15	305	1	15	320	15, 16	-	-
16	336	16	331	5	16	322	15, 16	-	-
17	378	17	383	-5	17	368	17	376	-8
18	401	18	401	0	18	402	18	406	-4

The frequencies and their differences are in cm^{-1} . The shifts are defined as $\nu_{EBS} - \nu_{BS}$. When * is present, the correspondence is dubious. When more than one normal mode is present, the EBS mode is a linear combination of several BS modes and the shift has consequently not been calculated.

Table 6.6: Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - M06 and OPBE.

M06					OPBE				
EBS		BS		Shift	EBS		BS		Shift
mode	ν	mode	ν		mode	ν	mode	ν	
1	25	1, 2		-	1	41	1	37	4
2	40	1, 2		-	2	62	2	61	1
3	79	3	71	8	3	90	3	89	1
4	97	4	94	3	4	94	4	91	3
5	115	5	111	4	5	96	5	96	0
6	126	7	123	3	6	108	6	107	1
7	129	6	121	8	7	116	7	112	4
8	145	8	140	5	8	142	8	138	4
9	178	9	175	3	9	158	9	153	5
10	180	10	171	9	10	174	10	175	-1
11	280	11	280	0	11	248	11	250	-2
12	298	12	285	13	12	276	12	276	0
13	311	13	307	4	13	293	13	291	2
14	324	14	319	5	14	305	14	304	1
15	330	15	326	4	15	322	15	321	1
16	341	16	339	2	16	328	16	332	-4
17	395	17	400	-5	17	379	17	389	-10
18	424	18	423	1	18	415	18	421	-6

The frequencies and their differences are in cm^{-1} . The shifts are defined as $\nu_{EBS} - \nu_{BS}$. When * is present, the correspondence is dubious. When more than one normal mode is present, the EBS mode is a linear combination of several BS modes and the shift has consequently not been calculated.

Table 6.7: Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B2PLYP and SCS-MP2.

B2PLYP					SCS-MP2				
EBS		BS		Shift	EBS		BS		Shift
mode	ν	mode	ν		mode	ν	mode	ν	
1	42	1	39	3	1	29	1	28	1
2	66	2	65	1	2	60	2	60	0
3	91	3	91	0	3	93	3	94	-1
4	97	4	89	8	4	100	4	100	0
5	101	5	100	1	5	101	5	100	1
6	114	6	114	0	6	108	6	107	1
7	119	7	114	5	7	120	7	120	0
8	144	8	140	4	8	138	8	138	0
9	156	9	151	5	9	156	9	154	2
10	169	10	170	-1	10	174	10	176	-2
11	268	11	269	-1	11	276	11	277	-1
12	276	12	272	4	12	287	12	287	0
13	299	13	299	0	13	315	13	316	-1
14	312	14	311	1	14	326	14	325	1
15	312	15	312	0	15	328	15	328	0
16	340	16	342	-2	16	365	16	365	0
17	384	17	391	-7	17	401	17	406	-5
18	409	18	413	-4	18	422	18	425	-3

The frequencies and their differences are in cm^{-1} . The shifts are defined as $\nu_{EBS} - \nu_{BS}$. When * is present, the correspondence is dubious. When more than one normal mode is present, the EBS mode is a linear combination of several BS modes and the shift has consequently not been calculated.

Table 6.8: Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_4\text{S}_4\text{Cl}_4$ - B3LYP and BP.

B3LYP					BP				
EBS		BS		Shift	EBS		BS		Shift
mode	ν	mode	ν		mode	ν	mode	ν	
1	82	2, 3*	-	-	1	56	1	49	7
2	89	4*	90	-1	2	61	3, 4, 5*	-	-
3	89	1, 2, 4*	-	-	3	64	3	60	+4
4	103	5*	96	+7	4	71	2	56	+15
5	112	1, 2*	-	-	5	81	4*	68	+13
6	122	7*	107	+15	6	102	6, 8	-	-
7	126	6*	104	+22	7	114	7	106	+8
8	130	8*	117	+13	8	119	8	110	+9
9	137	9, 10	-	-	9	129	9	125	+4
10	139	9, 10	-	-	10	150	11	134	+16
11	145	11	137	+8	11	151	10	133	+18
12	152	12*	142	+10	12	166	13	156	+10
13	158	13*	147	+11	13	168	14	157	+11
14	164	14*	154	+10	14	172	12	155	+17
15	247	15	227	+20	15	235	15, 16	-	-
16	250	16*	236	+14	16	249	15*	217	+32
17	259	17*	239	+20	17	256	19	256	0
18	263	16, 18*	-	-	18	265	17	239	+26
19	280	19*	263	+17	19	270	18	239	+31
20	283	20*	275	+8	20	277	20, 21	-	-
21	290	21	277	+13	21	285	20, 21	-	-
22	295	22	279	+16	22	303	22	292	+11
23	353	23	348	+5	23	364	23	352	+12
24	367	24	359	+8	24	385	25	375	+10
25	372	25	367	+5	25	391	24	374	+17
26	388	26	383	+5	26	415	26	403	+12
27	424	27	422	+2	27	426	27	421	+5
28	426	28	424	+2	28	430	29	425	+5
29	427	29	425	+2	29	430	28	425	+5
30	443	30	442	+1	30	447	30	443	+4

The frequencies and their differences are in cm^{-1} . The shifts are defined as $\nu_{EBS} - \nu_{BS}$. When * is present, the correspondence is dubious. When more than one normal mode is present, the EBS mode is a linear combination of several BS modes and the shift has consequently not been calculated.

Table 6.9: Comparison between the Extended Broken Symmetry and Broken Symmetry frequencies calculated on $\text{Fe}_4\text{S}_4\text{Cl}_4$ - M06 and OPBE.

M06					OPBE				
EBS		BS		Shift	EBS		BS		Shift
mode	ν	mode	ν		mode	ν	mode	ν	
1	45	1, 3*			1	54	1*	51	+3
2	76	3	98	-22	2	60	4*	64	-4
3	86	2*	86	0	3	62	2*	58	+4
4	98	4, 5*	-	-	4	67	3*	61	+6
5	103	4, 5, 6*	-	-	5	75	5*	78	-3
6	112	6*	120	-8	6	95	6	94	+1
7	118	8*	130	-12	7	109	8*	114	-5
8	128	7*	-	-	8	110	7*	105	+5
9	143	9	145	-2	9	130	9	128	+2
10	148	10*	147	-1	10	151	10, 11	-	-
11	151	11*	150	+1	11	154	10, 11	-	-
12	156	12*	157	-1	12	162	12	158	+4
13	165	13	165	0	13	165	13	160	+5
14	169	14	168	+1	14	170	14	164	+6
15	256	16*	267	-11	15	240	16	242	-2
16	261	17, 18	-	-	16	243	15	235	+8
17	270	17*	274	-4	17	257	18	262	-5
18	274	20*	292	-18	18	260	19	268	-8
19	289	19*	290	-1	19	265	17*	253	+12
20	293	15, 19, 21*	-	-	20	283	21	280	+3
21	298	22	336	-38	21	284	20	277	+7
22	300	21*	304	-4	22	309	22	300	+9
23	368	23	371	-3	23	362	23	358	+4
24	394	25	395	-1	24	393	25	388	+5
25	397	24	395	+2	25	394	24	387	+7
26	416	26	415	+1	26	421	26	415	+6
27	444	27	443	+1	27	428	27	425	+3
28	448	28, 29	-	-	28	433	28	430	+3
29	449	28, 29	-	-	29	434	29	431	+3
30	468	30	468	0	30	450	30	447	+3

The frequencies and their differences are in cm^{-1} . The shifts are defined as $\nu_{EBS} - \nu_{BS}$. When * is present, the correspondence is dubious. When more than one normal mode is present, the EBS mode is a linear combination of several BS modes and the shift has consequently not been calculated.

6.1.6 VMARD analysis for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$

The VMARD (Vibrational Mode Automatic Relevance Determination) analysis for $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ is reported in Table 6.10, Table 6.11, Table 6.12 and Table 6.13 for B3LYP, BP, M06 and OPBE levels of theory.

Table 6.10: Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - B3LYP.

B3LYP		
Mode	ν	VMARD decomposition
1	37	37% ρ_{ClFeFeCl} + 21% β_{FeFeCl}
2	72	59% β_{FeFeCl}
3	88	58% β_{ClFeS} + 24% β_{FeSFe}
4	99	20% β_{FeSFe} + 18% ρ_{FeSFeCl} + 21% β_{ClFeCl}
5	108	43% β_{FeFeCl} + 41% ν_{FeCl}
6	120	31% β_{ClFeCl} + 13% β_{SFeCl} + 18% ν_{FeCl}
7	124	45% ρ_{ClFeFeS} + 22% β_{FeFeCl}
8	143	73% β_{SFeCl}
9	166	18% β_{FeSFe} + 47% ν_{FeCl} + 15% β_{ClFeCl} + 12% β_{FeSFe}
10	173	37% ν_{FeCl} + 36% β_{SFeCl}
11	263	72% ν_{FeCl} + 22% β_{SFeS}
12	286	56% ν_{FeCl} + 36% β_{FeFeS}
13	289	50% β_{FeFeS} + 38% ν_{FeCl}
14	301	76% ν_{FeCl}
15	306	67% ν_{FeCl} + 12% β_{FeSFe} + 15% ν_{FeS}
16	336	51% β_{SFeS} + 28% ν_{FeCl}
17	378	76% ν_{FeS}
18	401	82% ν_{FeS}

The frequencies are in cm^{-1} . ν indicates the bond stretching motions, β indicates the angle bending motions and ρ indicated the dihedral angles torsions. The contributions smaller than 10% have been ignored.

Table 6.11: Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - BP.

BP		
Mode	ν	VMARD decomposition
1	57	59% β_{FeFeCl} + 12% β_{ClFeCl}
2	64	67% ρ_{ClFeFeCl} + 16% β_{SFeCl}
3	88	27% β_{FeSFe} + 45% β_{ClFeCl}
4	94	45% β_{FeFeCl} + 11% β_{ClFeCl} + 41% ν_{FeCl}
5	99	26% β_{FeSFe} + 66% β_{SFeCl}
6	106	44% β_{ClFeCl} + 22% ν_{FeCl} + 20% β_{SFeCl}
7	151	84% β_{SFeCl}
8	152	62% ρ_{SFeFeCl} + 25% β_{SFeCl}
9	173	33% β_{FeSFe} + 54% ν_{FeCl}
10	183	42% ν_{FeCl} + 40% β_{SFeCl}
11	240	74% ν_{FeCl} + 23% β_{SFeS}
12	254	99% ν_{FeCl}
13	277	69% ν_{FeCl}
14	291	64% ν_{FeCl} + 27% β_{FeSFe}
15	320	63% β_{FeFeS} + 21% ν_{FeCl}
16	322	78% β_{FeFeS}
17	368	82% ν_{FeS}
18	402	81% ν_{FeS}

The frequencies are in cm^{-1} . ν indicates the bond stretching motions, β indicates the angle bending motions and ρ indicated the dihedral angles torsions. The contributions smaller than 10% have been ignored.

Table 6.12: Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - OPBE.

OPBE		
Mode	ν	VMARD decomposition
1	41	69% β_{FeFeCl}
2	62	100% ρ_{ClFeFeCl}
3	90	35% β_{FeSFe} + 54% β_{ClFeCl}
4	94	99% β_{ClFeCl}
5	96	32% β_{FeSFe} + 68% β_{SFeCl}
6	108	56% β_{ClFeCl} + 17% β_{SFeS} + 24% ν_{FeCl}
7	116	96% ρ_{ClFeFeS}
8	142	84% β_{SFeCl}
9	158	37% ν_{FeCl} + 54% β_{SFeCl}
10	174	36% β_{FeSFe} + 55% ν_{FeCl} + 10% β_{ClFeCl}
11	248	72% ν_{FeCl} + 28% β_{SFeS}
12	276	96% ν_{FeCl}
13	293	79% ν_{FeCl}
14	305	66% ν_{FeCl} + 21% β_{FeSFe} + 13% ν_{FeS}
15	322	98% β_{FeFeS}
16	328	59% β_{SFeS} + 37% ν_{FeCl}
17	379	77% ν_{FeS} + 13% ν_{FeCl}
18	415	83% ν_{FeS}

The frequencies are in cm^{-1} . ν indicates the bond stretching motions, β indicates the angle bending motions and ρ indicated the dihedral angles torsions. The contributions smaller than 10% have been ignored.

Table 6.13: Extended Broken Symmetry Vibrational Mode Automatic Relevance Determination on $\text{Fe}_2\text{S}_2\text{Cl}_4^{2-}$ - SCS-MP2.

SCS-MP2		
Mode	ν	VMARD decomposition
1	29	66% β_{FeFeCl} + 10% ρ_{SFeFeS}
2	60	100% ρ_{ClFeFeCl}
3	93	34% β_{FeSFe} + 49% β_{ClFeCl}
4	100	32% β_{FeSFe} + 68% β_{SFeCl}
5	101	51% β_{FeFeCl} + 47% ν_{FeCl}
6	108	93% ρ_{ClFeFeS}
7	120	59% β_{ClFeCl} + 18% β_{SFeS} + 24% ν_{FeCl}
8	138	90% β_{SFeCl}
9	156	35% ν_{FeCl} + 61% β_{SFeCl}
10	174	35% β_{FeSFe} + 50% ν_{FeCl} + 15% β_{ClFeCl}
11	276	84% β_{FeFeS}
12	287	75% ν_{FeCl} + 25% β_{SFeS}
13	315	88% ν_{FeCl}
14	326	84% ν_{FeCl}
15	328	60% ν_{FeCl} + 16% β_{FeSFe} + 20% ν_{FeS}
16	365	71% β_{SFeS} + 25% ν_{FeCl}
17	401	75% ν_{FeS} + 21% ν_{FeCl}
18	422	83% ν_{FeS}

The frequencies are in cm^{-1} . ν indicates the bond stretching motions, β indicates the angle bending motions and ρ indicated the dihedral angles torsions. The contributions smaller than 10% have been ignored.

6.2 F-RESP

6.2.1 Electric field gradients using the Ewald summation

Defining $\vec{r}_b = \vec{r}_{b_2} - \vec{r}_{b_1}$ and $\vec{r}_{jh_b} = \vec{r}_{h_b} - \vec{r}_j$, it can be written:

$$\begin{aligned} \nabla_{r_j} \left(\vec{E}_{NB}(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= \frac{q_j^{gen} \hat{r}_{jh_b}}{|\vec{r}_b| |\vec{r}_{jh_b}|} \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \left(\frac{3}{|\vec{r}_{jh_b}|^2} + 2\alpha^2 \right) + \frac{3 \operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^3} \right) \\ &(\vec{r}_b \cdot \vec{r}_{jh_b}) - \frac{q_j^{gen} \hat{r}_b}{|\vec{r}_{jh_b}|^2 |\vec{r}_b|} \left(\frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \right) \end{aligned} \quad (6.3)$$

$$\begin{aligned} \nabla_{r_{b_1}} \left(\vec{E}_{NB}(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= \frac{q_j^{gen} \hat{r}_b}{|\vec{r}_b|^2 |\vec{r}_{jh_b}|} \left(\frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \right) + \\ &- \frac{q_j^{gen} \hat{r}_{jh_b}}{2 |\vec{r}_{jh_b}| |\vec{r}_b|} \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \left(\frac{3}{|\vec{r}_{jh_b}|^2} + 2\alpha^2 \right) + \frac{3 \operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^3} \right) (\vec{r}_{jh_b} \cdot \vec{r}_b) + \\ &\frac{q_j^{gen} \vec{r}_{j,b_1}}{|\vec{r}_{jh_b}|^2 |\vec{r}_b|} \left(\frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \right) \end{aligned} \quad (6.4)$$

$$\begin{aligned} \nabla_{r_{b_2}} \left(\vec{E}_{NB}(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= - \frac{q_j^{gen} \hat{r}_b}{|\vec{r}_b|^2 |\vec{r}_{jh_b}|} \left(\frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \right) + \\ &- \frac{q_j^{gen} \hat{r}_{jh_b}}{2 |\vec{r}_{jh_b}| |\vec{r}_b|} \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \left(\frac{3}{|\vec{r}_{jh_b}|^2} + 2\alpha^2 \right) + \frac{3 \operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^3} \right) (\vec{r}_{jh_b} \cdot \vec{r}_b) + \\ &\frac{q_j^{gen} \vec{r}_{j,b_2}}{|\vec{r}_{jh_b}|^2 |\vec{r}_b|} \left(\frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \right) \end{aligned} \quad (6.5)$$

$$\begin{aligned} \nabla_{r_j} \left(\vec{E}_B(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= \frac{q_j^{gen} (\vec{r}_{jh_b} \cdot \hat{r}_b)}{|\vec{r}_{jh_b}|^2} \hat{r}_{jh_b} \left(\frac{2}{|\vec{r}_{jh_b}|} \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) + \right. \\ &\left. \frac{2\alpha}{\sqrt{\pi}} |\vec{r}_{jh_b}| e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \left(2\alpha^2 + \frac{1}{|\vec{r}_{jh_b}|^2} \right) + \frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} - \frac{1}{|\vec{r}_{jh_b}|} \right) + \\ &\frac{q_j^{gen}}{|\vec{r}_{jh_b}|^2} \hat{r}_b \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{\operatorname{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \end{aligned} \quad (6.6)$$

$$\begin{aligned}
\nabla_{r_{b_1}} \left(\vec{E}_B(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= \frac{q_j^{gen}}{|\vec{r}_{jh_b}|} \hat{r}_{jh_b} \left(\left(\frac{2\alpha}{\sqrt{\pi}} |\vec{r}_{jh_b}| e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \left(2\alpha^2 + \frac{1}{|\vec{r}_{jh_b}|^2} \right) + \right. \right. \\
&\frac{erfc(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{1}{|\vec{r}_{jh_b}|^2} \left. \right) (\vec{r}_{jh_b} \cdot \hat{r}_b) - \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{erf(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \cdot \\
&\left. \left(\frac{1}{|\vec{r}_b|} + \frac{\vec{r}_{jh_b} \cdot \hat{r}_b}{|\vec{r}_{jh_b}|^2} \right) \right) + \frac{q_j^{gen}}{|\vec{r}_{jh_b}|^2} \hat{r}_b \cdot \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{erf(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \left(\frac{1}{2} + \frac{\vec{r}_{jh_b} \cdot \hat{r}_b}{|\vec{r}_b|} \right)
\end{aligned} \tag{6.7}$$

$$\begin{aligned}
\nabla_{r_{b_2}} \left(\vec{E}_B(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= \frac{q_j^{gen}}{|\vec{r}_{jh_b}|} \hat{r}_{jh_b} \left(\left(\frac{2\alpha}{\sqrt{\pi}} |\vec{r}_{jh_b}| e^{-\alpha^2 |\vec{r}_{jh_b}|^2} \left(2\alpha^2 + \frac{1}{|\vec{r}_{jh_b}|^2} \right) + \right. \right. \\
&\frac{erfc(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} + \frac{1}{|\vec{r}_{jh_b}|^2} \left. \right) (\vec{r}_{jh_b} \cdot \hat{r}_b) + \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{erf(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \cdot \\
&\left. \left(\frac{1}{|\vec{r}_b|} - \frac{\vec{r}_{jh_b} \cdot \hat{r}_b}{|\vec{r}_{jh_b}|^2} \right) \right) + \frac{q_j^{gen}}{|\vec{r}_{jh_b}|^2} \hat{r}_b \cdot \left(\frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2 |\vec{r}_{jh_b}|^2} - \frac{erf(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|} \right) \left(\frac{1}{2} - \frac{\vec{r}_{jh_b} \cdot \hat{r}_b}{|\vec{r}_b|} \right)
\end{aligned} \tag{6.8}$$

$$\nabla_{r_j} \left(\vec{E}_K(\vec{r}_{h_b}) \cdot \hat{r}_b \right) = - \sum_{\vec{k} \neq 0} \frac{4\pi}{|\vec{k}|^2} e^{-\frac{|\vec{k}|^2}{4\alpha^2}} (\vec{k} \cdot \hat{r}_b) q_j^{gen} \Re \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) \vec{k} \tag{6.9}$$

$$\begin{aligned}
\nabla_{r_{b_1}} \left(\vec{E}_K(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= 4\pi \sum_{\vec{k} \neq 0} \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \left(\frac{(\vec{k} \cdot \hat{r}_b)}{|\vec{r}_b|} \left(\sum_{j=1}^N q_j^{gen} \Im \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) \right) \hat{r}_b + \right. \\
&- \left. \left(\sum_{j=1}^N q_j^{gen} \left(\frac{1}{|\vec{r}_b|} \Im \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) + \frac{(\vec{k} \cdot \hat{r}_b)}{2} \Re \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) \right) \right) \vec{k} \right)
\end{aligned} \tag{6.10}$$

$$\begin{aligned}
\nabla_{r_{b_2}} \left(\vec{E}_K(\vec{r}_{h_b}) \cdot \hat{r}_b \right) &= 4\pi \sum_{\vec{k} \neq 0} \frac{e^{-\frac{|\vec{k}|^2}{4\alpha^2}}}{|\vec{k}|^2} \left(\left(\sum_{j=1}^N q_j^{gen} \left(\frac{1}{|\vec{r}_b|} \Im \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) + \right. \right. \right. \\
&\left. \left. \frac{(\vec{k} \cdot \hat{r}_b)}{2} \Re \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) \right) \right) \vec{k} - \frac{(\vec{k} \cdot \hat{r}_b)}{|\vec{r}_b|} \left(\sum_{j=1}^N q_j^{gen} \Im \left(e^{i\vec{k} \cdot \vec{r}_{jh_b}} \right) \right) \hat{r}_b
\end{aligned} \tag{6.11}$$

Attention must be paid to the different meaning of j , which can belong to the $\{V.L.b - M_b\}$, $\{M_b\}$ or $\{System\}$ sets according to the considered fictitious electric field contribution (non-bonded, bonded and reciprocal space, respectively).

All the previous expressions are calculated along with the charge variations except for Equation 6.9, whose evaluation would require an order $O(N^2)$ calculation (considering the bonds to have a N cardinality such as the atoms), thus resulting in a computational order overcoming that of the classical, well-tuned Ewald summation ($O(N^{\frac{3}{2}})$). Its evaluation is performed subsequently, in a more clever way that permits the entire algorithm to remain $O(N^{\frac{3}{2}})$.

Being the force connected to Equation 6.9 (being careful not to confuse the α coefficient from the Ewald summation and the α_b^i parameters from the F-RESP model):

$$\vec{f}_K^{F-RESP}(\vec{r}_z) = -q_z \Re \left(\sum_{\vec{k} \neq 0} e^{-i\vec{k} \cdot \vec{r}_z} \sum_{i=1}^N \phi(i) \sum_{b \in M_i} \alpha_b^i (\vec{k} \cdot \hat{r}_b) \vec{k} \frac{4\pi}{|\vec{k}|^2} e^{-\frac{|\vec{k}|^2}{4\alpha^2}} e^{i\vec{k} \cdot \vec{r}_{h_b}} \right) \quad (6.12)$$

where $\phi(i)$ is the electrostatic potential for the atom i , the following algorithm can be written:

```

FOR bond in System:
  FOR k in K-space:
    appo2[bond][k] = c(b, k) * exp(i * k * r_h_b)
  END
END
FOR bond in System:
  FOR atom in Molecules[bond]:
    FOR k in K-space:
      appo3[k] += phi[atom] * appo2[bond][k]
    END
  END
END
FOR atom in System:
  FOR k in K-space:
    partial += appo3[k] * exp(-i * k * r_atom)
  END
  f[atom] += qgen[atom] * partial
END

```

where $c(b, k)$ is equal to $(\vec{k} \cdot \hat{r}_b) \vec{k} \frac{4\pi}{|\vec{k}|^2} e^{-\frac{|\vec{k}|^2}{4\alpha^2}}$.

Assuming a (reasonable) \sqrt{N} cardinality for the Fourier space, the overall computational order of the previous algorithm can be easily seen to be $O\left(N^{\frac{3}{2}}\right)$.

6.2.2 Electric field gradients using the Fennell potential

The bond is defined as $\vec{r}_{b_1} - \vec{r}_{b_2}$ and it should be pointed out that j always belongs to the $\{V.L.b - M_b\}$ set.

$$\begin{aligned} \nabla_{\vec{r}_j} \left(\vec{E}(\vec{r}_h) \cdot \hat{r}_b \right) &= \sum_{\substack{j \in V.L.b, \\ j \notin M_b}} \frac{q_j^{gen}}{|\vec{r}_{jh_b}| |\vec{r}_b|} \left(\left(\frac{\text{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} - f_{shift} \right) \right. \\ &\quad \left. \left(3 \frac{(\vec{r}_{jh_b} \cdot \vec{r}_b)}{|\vec{r}_{jh_b}|^2} \vec{r}_{jh_b} - \vec{r}_b \right) + 2 (\vec{r}_{jh_b} \cdot \vec{r}_b) \left(\frac{f_{shift}}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha^3 e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} \right) \vec{r}_{jh_b} \right) \end{aligned} \quad (6.13)$$

$$\begin{aligned} \nabla_{\vec{r}_{b_1}} \left(\vec{E}(\vec{r}_h) \cdot \hat{r}_b \right) &= \sum_{\substack{j \in V.L.b, \\ j \notin M_b}} \frac{q_j^{gen}}{|\vec{r}_{jh_b}| |\vec{r}_b|} \left(\left(\frac{\text{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} - f_{shift} \right) \cdot \right. \\ &\quad \left(\left(\frac{|\vec{r}_b|}{2} - \frac{(\vec{r}_{jh_b} \cdot \vec{r}_b)}{|\vec{r}_b|} \right) \hat{r}_b + \left(1 - \frac{3 (\vec{r}_{jh_b} \cdot \vec{r}_b)}{2 |\vec{r}_{jh_b}|^2} \right) \vec{r}_{jh_b} \right) - (\vec{r}_{jh_b} \cdot \vec{r}_b) \cdot \\ &\quad \left. \left(\frac{f_{shift}}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha^3 e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} \right) \vec{r}_{jh_b} \right) \end{aligned} \quad (6.14)$$

$$\begin{aligned} \nabla_{\vec{r}_{b_2}} \left(\vec{E}(\vec{r}_h) \cdot \hat{r}_b \right) &= \sum_{\substack{j \in V.L.b, \\ j \notin M_b}} \frac{q_j^{gen}}{|\vec{r}_{jh_b}| |\vec{r}_b|} \left(\left(\frac{\text{erfc}(\alpha |\vec{r}_{jh_b}|)}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} - f_{shift} \right) \cdot \right. \\ &\quad \left(\left(\frac{|\vec{r}_b|}{2} + \frac{(\vec{r}_{jh_b} \cdot \vec{r}_b)}{|\vec{r}_b|} \right) \hat{r}_b - \left(1 + \frac{3 (\vec{r}_{jh_b} \cdot \vec{r}_b)}{2 |\vec{r}_{jh_b}|^2} \right) \vec{r}_{jh_b} \right) - (\vec{r}_{jh_b} \cdot \vec{r}_b) \cdot \\ &\quad \left. \left(\frac{f_{shift}}{|\vec{r}_{jh_b}|^2} + \frac{2\alpha^3 e^{-\alpha^2 |\vec{r}_{jh_b}|^2}}{\sqrt{\pi} |\vec{r}_{jh_b}|} \right) \vec{r}_{jh_b} \right) \end{aligned} \quad (6.15)$$

The use of the Fennell potential for the calculation of the fictitious electric field permits to easily derive the additional (with respect to a first-generation, fixed-charge force field) contribution to the virial that comes from the charge fluctuation. The great advantage is, in fact, that $\{V.L.b\}$ is a finite set (differently from what happens with the Ewald summation, that makes each atom of the infinite system to exert its influence over the bond) and that Fennell potential is

evaluated only in the real space. This permits to express the virial contribution coming from the polarization using the $\frac{1}{3V} \sum_{i=1}^N \vec{f}_i^{F-RESP} \cdot \vec{r}_i$ formula, independently from the use of long-range solvers for the main electrostatic calculation.

6.2.3 Electric field gradients using the Waibel potential

Also in this case, the bond is defined as $\vec{r}_{b_1} - \vec{r}_{b_2}$ and j always belongs to the $\{V.L.b - M_b\}$ set.

$$\begin{aligned} \nabla_{\vec{r}_j} \left(\vec{E}(\vec{r}_M) \cdot \hat{r}_b \right) &= \frac{q_j^{gen}}{|\vec{r}_{jM}| |\vec{r}_b|} \left(\left(\frac{\text{erfc}(\alpha |\vec{r}_{jM}|)}{|\vec{r}_{jM}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jM}|^2}}{\sqrt{\pi} |\vec{r}_{jM}|} - f_{shift} \cdot e^{\beta(|r_{jM}| - R_c)} \right) \right. \\ &\left. \left(\frac{3(\vec{r}_{jM} \cdot \vec{r}_b)}{|\vec{r}_{jM}|^2} \vec{r}_{jM} - \vec{r}_b \right) + 2(\vec{r}_{jM} \cdot \vec{r}_b) \left(\frac{2\alpha^3 e^{-\alpha^2 |\vec{r}_{jM}|^2}}{\sqrt{\pi} |\vec{r}_{jM}|} + f_{shift} \left(\frac{2 + \beta |\vec{r}_{jM}|}{2 |\vec{r}_{jM}|^2} \right) \right) \right) \vec{r}_{jM} \end{aligned} \quad (6.16)$$

$$\begin{aligned} \nabla_{\vec{r}_{b_1}} \left(\vec{E}(\vec{r}_M) \cdot \hat{r}_b \right) &= \sum_{(b,j)} \frac{q_j^{gen}}{|\vec{r}_{jM}| |\vec{r}_b|} \left(\left(\frac{\text{erfc}(\alpha |\vec{r}_{jM}|)}{|\vec{r}_{jM}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jM}|^2}}{\sqrt{\pi} |\vec{r}_{jM}|} - f_{shift} \right) \cdot \right. \\ &\left(\left(\frac{|\vec{r}_b|}{2} - \frac{(\vec{r}_{jM} \cdot \vec{r}_b)}{|\vec{r}_b|} \right) \hat{r}_b + \left(1 - \frac{3(\vec{r}_{jM} \cdot \vec{r}_b)}{2 |\vec{r}_{jM}|^2} \right) \vec{r}_{jM} \right) - (\vec{r}_{jM} \cdot \vec{r}_b) \cdot \\ &\left(f_{shift} \left(\frac{2 + \beta |\vec{r}_{jM}|}{2 |\vec{r}_{jM}|^2} \right) + \frac{2\alpha^3 e^{-\alpha^2 |\vec{r}_{jM}|^2}}{\sqrt{\pi} |\vec{r}_{jM}|} \right) \vec{r}_{jM} \end{aligned} \quad (6.17)$$

$$\begin{aligned} \nabla_{\vec{r}_{b_2}} \left(\vec{E}(\vec{r}_M) \cdot \hat{r}_b \right) &= \sum_{(b,j)} \frac{q_j^{gen}}{|\vec{r}_{jM}| |\vec{r}_b|} \left(\left(\frac{\text{erfc}(\alpha |\vec{r}_{jM}|)}{|\vec{r}_{jM}|^2} + \frac{2\alpha e^{-\alpha^2 |\vec{r}_{jM}|^2}}{\sqrt{\pi} |\vec{r}_{jM}|} - f_{shift} \right) \cdot \right. \\ &\left(\left(\frac{|\vec{r}_b|}{2} + \frac{(\vec{r}_{jM} \cdot \vec{r}_b)}{|\vec{r}_b|} \right) \hat{r}_b - \left(1 + \frac{3(\vec{r}_{jM} \cdot \vec{r}_b)}{2 |\vec{r}_{jM}|^2} \right) \vec{r}_{jM} \right) - (\vec{r}_{jM} \cdot \vec{r}_b) \cdot \\ &\left(f_{shift} \left(\frac{2 + \beta |\vec{r}_{jM}|}{2 |\vec{r}_{jM}|^2} \right) + \frac{2\alpha^3 e^{-\alpha^2 |\vec{r}_{jM}|^2}}{\sqrt{\pi} |\vec{r}_{jM}|} \right) \vec{r}_{jM} \end{aligned} \quad (6.18)$$

The same demonstration that is shown few lines above about the virial contribution is valid for Waibel potential too.

6.2.4 Damping function gradients

The derivatives of the damping functions with respect to \vec{r}_j and $\vec{r}_{b_1 \vee b_2}$ are (defining $\vec{r}_{jh_b} = \vec{r}_{h_b} - \vec{r}_j$):

$$\nabla_{\vec{r}_j} f_{damp}^{sin}(|\vec{r}_{jh_b}|) = -\frac{\pi}{B-A} \sin\left(\frac{\pi}{2} \left(\frac{|\vec{r}_{jh_b}| - A}{B-A}\right)\right) \cos\left(\frac{\pi}{2} \left(\frac{|\vec{r}_{jh_b}| - A}{B-A}\right)\right) \hat{r}_{jh_b} \quad (6.19)$$

$$\nabla_{\vec{r}_{b_1 \vee b_2}} f_{damp}^{sin}(|\vec{r}_{jh_b}|) = \frac{\pi}{2(B-A)} \sin\left(\frac{\pi}{2} \left(\frac{|\vec{r}_{jh_b}| - A}{B-A}\right)\right) \cos\left(\frac{\pi}{2} \left(\frac{|\vec{r}_{jh_b}| - A}{B-A}\right)\right) \hat{r}_{jh_b} \quad (6.20)$$

$$\nabla_{\vec{r}_j} f_{damp}^{gauss}(|\vec{r}_{jh_b}|) = \frac{(|\vec{r}_{jh_b}| - B)}{A^2} e^{-\frac{(|\vec{r}_{jh_b}| - A)^2}{2A^2}} \hat{r}_{jh_b} \quad (6.21)$$

$$\nabla_{\vec{r}_{b_1 \vee b_2}} f_{damp}^{gauss}(|\vec{r}_{jh_b}|) = -\frac{(|\vec{r}_{jh_b}| - B)}{2A^2} e^{-\frac{(|\vec{r}_{jh_b}| - A)^2}{2A^2}} \hat{r}_{jh_b} \quad (6.22)$$

6.2.5 Geometrical descriptors gradients

If the bond \vec{r}_b is defined as $\vec{r}_{b_1} - \vec{r}_{b_2}$, the angle θ_a as $\arccos\left(\frac{(\vec{r}_{a_1} - \vec{r}_{a_2}) \cdot (\vec{r}_{a_3} - \vec{r}_{a_2})}{|\vec{r}_{a_1} - \vec{r}_{a_2}| |\vec{r}_{a_3} - \vec{r}_{a_2}|}\right)$ and the improper dihedral χ_{im} as $\text{sign}(\vec{r}_{32} \cdot (\vec{r}_{12} \times \vec{r}_{32}) \times (\vec{r}_{32} \times \vec{r}_{34})) \arccos\left(\frac{(\vec{r}_{12} \times \vec{r}_{32}) \cdot (\vec{r}_{32} \times \vec{r}_{34})}{|\vec{r}_{12} \times \vec{r}_{32}| |\vec{r}_{32} \times \vec{r}_{34}|}\right)$, the following expressions can be written:

$$\nabla_{\vec{r}_{b_1}} (|\vec{r}_b| - |\vec{r}_b^0|) = \hat{r}_b \quad (6.23)$$

$$\nabla_{\vec{r}_{b_2}} (|\vec{r}_b| - |\vec{r}_b^0|) = -\hat{r}_b \quad (6.24)$$

$$\nabla_{\vec{r}_{a_1}} (\theta_a - \theta_a^0) = \frac{1}{\sqrt{1 - \cos^2(\theta_a)}} \frac{1}{|\vec{r}_{a_1} - \vec{r}_{a_2}|} \left(\frac{\vec{r}_{a_1} - \vec{r}_{a_2}}{|\vec{r}_{a_1} - \vec{r}_{a_2}|} \cos(\theta_a) - \frac{\vec{r}_{a_3} - \vec{r}_{a_2}}{|\vec{r}_{a_3} - \vec{r}_{a_2}|} \right) \quad (6.25)$$

$$\nabla_{\vec{r}_{a_2}} (\theta_a - \theta_a^0) = \frac{1}{\sqrt{1 - \cos^2(\theta_a)}} \left(\frac{\vec{r}_{a_1} + \vec{r}_{a_3} - 2\vec{r}_{a_2}}{|\vec{r}_{a_1} - \vec{r}_{a_2}| |\vec{r}_{a_3} - \vec{r}_{a_2}|} + \cos(\theta_a) \left(\frac{\vec{r}_{a_1} - \vec{r}_{a_2}}{|\vec{r}_{a_1} - \vec{r}_{a_2}|} + \frac{\vec{r}_{a_3} - \vec{r}_{a_2}}{|\vec{r}_{a_3} - \vec{r}_{a_2}|} \right) \right) \quad (6.26)$$

$$\nabla_{\vec{r}_{a_3}} (\theta_a - \theta_a^0) = \frac{1}{\sqrt{1 - \cos^2(\theta_a)}} \frac{1}{|\vec{r}_{a_3} - \vec{r}_{a_2}|} \left(\frac{\vec{r}_{a_3} - \vec{r}_{a_2}}{|\vec{r}_{a_3} - \vec{r}_{a_2}|} \cos(\theta_a) - \frac{\vec{r}_{a_1} - \vec{r}_{a_2}}{|\vec{r}_{a_1} - \vec{r}_{a_2}|} \right) \quad (6.27)$$

$$\frac{\partial |\chi - \chi^0|}{\partial \vec{r}_1} = -\frac{1}{\sin(\chi)} \cdot \frac{\chi - \chi^0}{|\chi - \chi^0|} \cdot \vec{r}_{32} \times \frac{1}{|\vec{r}_{12} \times \vec{r}_{32}|} \left(\frac{\vec{r}_{32} \times \vec{r}_{34}}{|\vec{r}_{32} \times \vec{r}_{34}|} - \cos(\chi) \frac{\vec{r}_{12} \times \vec{r}_{32}}{|\vec{r}_{12} \times \vec{r}_{32}|} \right) \quad (6.28)$$

$$\begin{aligned} \frac{\partial |\chi - \chi^0|}{\partial \vec{r}_2} = & -\frac{1}{\sin(\chi)} \cdot \frac{\chi - \chi^0}{|\chi - \chi^0|} \cdot \left(\vec{r}_{13} \times \frac{1}{|\vec{r}_{12} \times \vec{r}_{32}|} \left(\frac{\vec{r}_{32} \times \vec{r}_{34}}{|\vec{r}_{32} \times \vec{r}_{34}|} - \cos(\chi) \frac{\vec{r}_{12} \times \vec{r}_{32}}{|\vec{r}_{12} \times \vec{r}_{32}|} \right) + \right. \\ & \left. - \vec{r}_{34} \times \frac{1}{|\vec{r}_{32} \times \vec{r}_{34}|} \left(\frac{\vec{r}_{12} \times \vec{r}_{32}}{|\vec{r}_{12} \times \vec{r}_{32}|} - \cos(\chi) \frac{\vec{r}_{32} \times \vec{r}_{34}}{|\vec{r}_{32} \times \vec{r}_{34}|} \right) \right) \end{aligned} \quad (6.29)$$

$$\begin{aligned} \frac{\partial |\chi - \chi^0|}{\partial \vec{r}_3} = & -\frac{1}{\sin(\chi)} \cdot \frac{\chi - \chi^0}{|\chi - \chi^0|} \cdot \left(\vec{r}_{24} \times \frac{1}{|\vec{r}_{32} \times \vec{r}_{34}|} \left(\frac{\vec{r}_{12} \times \vec{r}_{32}}{|\vec{r}_{12} \times \vec{r}_{32}|} - \cos(\chi) \frac{\vec{r}_{32} \times \vec{r}_{34}}{|\vec{r}_{32} \times \vec{r}_{34}|} \right) + \right. \\ & \left. - \vec{r}_{12} \times \frac{1}{|\vec{r}_{12} \times \vec{r}_{32}|} \left(\frac{\vec{r}_{32} \times \vec{r}_{34}}{|\vec{r}_{32} \times \vec{r}_{34}|} - \cos(\chi) \frac{\vec{r}_{12} \times \vec{r}_{32}}{|\vec{r}_{12} \times \vec{r}_{32}|} \right) \right) \end{aligned} \quad (6.30)$$

$$\frac{\partial |\chi - \chi^0|}{\partial \vec{r}_4} = -\frac{1}{\sin(\chi)} \cdot \frac{\chi - \chi^0}{|\chi - \chi^0|} \cdot \vec{r}_{32} \times \frac{1}{|\vec{r}_{32} \times \vec{r}_{34}|} \left(\frac{\vec{r}_{12} \times \vec{r}_{32}}{|\vec{r}_{12} \times \vec{r}_{32}|} - \cos(\chi) \frac{\vec{r}_{32} \times \vec{r}_{34}}{|\vec{r}_{32} \times \vec{r}_{34}|} \right) \quad (6.31)$$

6.3 Effects of the changes in the parameters of the damping function - Ewald summation

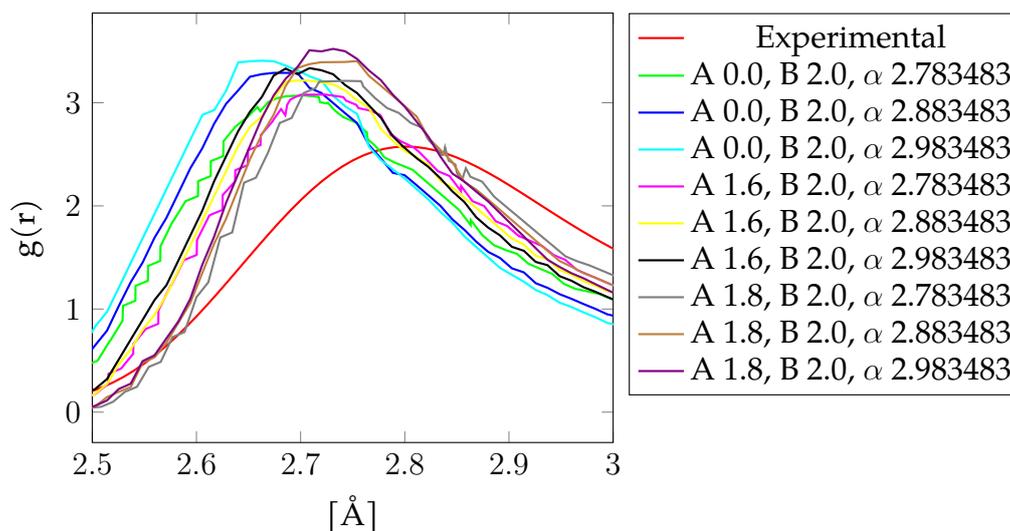


Figure 6.1: Changes in the O-O RDF with respect to the parameters of the damping (sinusoidal) function - Ewald summation TIP3P charge matching.

Table 6.14: Changes in the O-O RDF maxima and minima position and in the diffusion coefficient with respect to the parameters of the damping (sinusoidal) function - Ewald summation TIP3P charge matching.

A [Å]	α	First max [Å]	First min [Å]	Second max [Å]	Second min [Å]	Third max [Å]	D [$10^{-5}cm^2/s$]
Exp[224, 205]		2.80	3.47	4.51	5.58	6.69	2.30
0.0	2.783483	2.71	3.30	4.42	5.64	6.56	2.96
	2.883483	2.68	3.24	4.42	5.62	6.82	2.22
	2.983483	2.66	3.28	4.44	5.54	6.60	1.68
1.6	2.783483	2.72	3.28	4.56	5.56	6.81	2.98
	2.883483	2.70	3.32	4.56	5.60	6.69	3.08
	2.983483	2.70	3.32	4.54	5.60	6.82	2.48
1.8	2.783483	2.76	3.48	4.70	5.73	6.88	3.32
	2.883483	2.77	3.32	4.60	5.56	6.49	2.48
	2.983483	2.72	3.28	4.58	5.66	6.86	2.24

B is always 2.0 Å.

6.4 Effects of the changes in the parameters of the damping function - Fennell potential

Table 6.15: Changes in the density and the diffusion coefficient with respect to the parameters of the damping (sinusoidal) function - Fennell potential TIP3P charge matching.

A [Å]	α	ρ [g/cm ³]	D [10 ⁻⁵ cm ² /s]
Exp[224, 205]		0.997	2.30
	TIP3P	1.024	5.37
	2.783483	1.078	0.71
0.0	2.883483	1.089	0.36
	2.983483	1.099	0.14
	2.783483	1.037	2.00
1.6	2.883483	1.043	1.59
	2.983483	1.049	1.2
	2.783483	1.021	2.82
1.8	2.883483	1.027	2.34
	2.983483	1.033	2.04

B is always 2.0 Å.

Chapter 7

Publication list

- Francesco Cappelluti *et al.* “F-RESP polarizable force field and its application on liquid water”. In: (in preparation)
- M Campetella *et al.* “Physical-chemical studies on putrescine (butane-1, 4-diamine) and its solutions: Experimental and computational investigations”. In: *Journal of Molecular Liquids* 322 (2021), p. 114568
- O Palumbo *et al.* “New insights into chloromethyl-oxirane and chloromethylthiirane in liquid and solid phase from low-temperature infrared spectroscopy and ab initio modeling”. In: *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 247 (2021), p. 119061
- Nicolas Nieuwjaer *et al.* “IRMPD spectroscopy and quantum chemistry calculations on mono- and bi-metallic complexes of acetylacetonate ligands with aluminum, iron, and ruthenium ions”. In: *The Journal of Chemical Physics* 153.23 (2020), p. 234303
- Francesco Cappelluti, Luigi Bencivenni, and Leonardo Guidoni. “Spin-symmetrised structures and vibrational frequencies of iron–sulfur clusters”. In: *Physical Chemistry Chemical Physics* 22.29 (2020), pp. 16655–16664
- Marco Campetella, Francesco Cappelluti, and Lorenzo Gontrani. “Medium range interactions evidences in compounds with aliphatic lateral chain: 1-pentanoic acid, 1-pentanol and pentylammonium nitrate as test cases”. In: *Chemical Physics Letters* 734 (2019), p. 136738
- Shibing Chu *et al.* “Effects of Static Correlation between Spin Centers in Multicenter Transition Metal Complexes”. In: *Journal of chemical theory and computation* 13.10 (2017), pp. 4675–4683

Acknowledgments

I would like to thank Prof. L. Guidoni and all the CBBC research group for their support during my PhD course. I would also like to acknowledge Dr. L. Bencivenni for the many fruitful discussions and his kind help.

References

- [1] Clifford Dykstra *et al.* *Theory and applications of computational chemistry: the first forty years*. Elsevier, 2011.
- [2] Stefan Grimme and Peter R Schreiner. “Computational chemistry: the fate of current methods and future challenges”. In: *Angewandte Chemie International Edition* 57.16 (2018), pp. 4170–4176.
- [3] Douglas Rayne Hartree. “The wave mechanics of an atom with a non-coulomb central field. Part II. Some results and discussion”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 24. 1. Cambridge University Press. 1928, pp. 111–132.
- [4] John Clarke Slater. “The self consistent field and the structure of atoms”. In: *Physical Review* 32.3 (1928), p. 339.
- [5] JA Gaunt. “A theory of Hartree’s atomic fields”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 24. 2. Cambridge University Press. 1928, pp. 328–342.
- [6] John C Slater. “Note on Hartree’s method”. In: *Physical Review* 35.2 (1930), p. 210.
- [7] Vladimir Fock. “Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems”. In: *Zeitschrift für Physik* 61.1-2 (1930), pp. 126–148.
- [8] Douglas Rayner Hartree and W Hartree. “Self-consistent field, with exchange, for beryllium”. In: *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences* 150.869 (1935), pp. 9–33.
- [9] Clemens Carel Johannes Roothaan. “New developments in molecular orbital theory”. In: *Reviews of modern physics* 23.2 (1951), p. 69.
- [10] GG Hall. “The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 205.1083 (1951), pp. 541–552.
- [11] S Francis Boys. “Electronic wave functions-I. A general method of calculation for the stationary states of any molecular system”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 200.1063 (1950), pp. 542–554.

- [12] SF Boys *et al.* "Automatic fundamental calculations of molecular structure". In: *Nature* 178.4544 (1956), pp. 1207–1209.
- [13] Peter Pulay. "Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules: I. Theory". In: *Molecular Physics* 17.2 (1969), pp. 197–204.
- [14] MJ Frisch *et al.* *Gaussian 16*. 2016.
- [15] Pierre Hohenberg and Walter Kohn. "Inhomogeneous electron gas". In: *Physical review* 136.3B (1964), B864.
- [16] Walter Kohn and Lu Jeu Sham. "Self-consistent equations including exchange and correlation effects". In: *Physical review* 140.4A (1965), A1133.
- [17] Llewellyn H Thomas. "The calculation of atomic fields". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 23. 5. Cambridge University Press. 1927, pp. 542–548.
- [18] Enrico Fermi. "Un metodo statistico per la determinazione di alcune proprietà dell'atomo". In: *Rend. Accad. Naz. Lincei* 6.602-607 (1927), p. 32.
- [19] Chr. Møller and Milton S. Plesset. "Note on an approximation treatment for many-electron systems". In: *Physical Review* 46.7 (1934), p. 618.
- [20] Fritz Coester and Hermann Kümmel. "Short-range correlations in nuclear wave functions". In: *Nuclear Physics* 17 (1960), pp. 477–485.
- [21] J Paldus, J Čížek, and I Shavitt. "Correlation Problems in Atomic and Molecular Systems. IV. Extended Coupled-Pair Many-Electron Theory and Its Application to the B H 3 Molecule". In: *Physical Review A* 5.1 (1972), p. 50.
- [22] Nicholas Metropolis *et al.* "Equation of state calculations by fast computing machines". In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.
- [23] Nicholas Metropolis and Stanislaw Ulam. "The monte carlo method". In: *Journal of the American statistical association* 44.247 (1949), pp. 335–341.
- [24] Berni J Alder and Thomas Everett Wainwright. "Studies in molecular dynamics. I. General method". In: *The Journal of Chemical Physics* 31.2 (1959), pp. 459–466.
- [25] JB Gibson *et al.* "Dynamics of radiation damage". In: *Physical Review* 120.4 (1960), p. 1229.
- [26] Aneesur Rahman. "Correlations in the motion of atoms in liquid argon". In: *Physical review* 136.2A (1964), A405.
- [27] J Andrew McCammon, Bruce R Gelin, and Martin Karplus. "Dynamics of folded proteins". In: *Nature* 267.5612 (1977), pp. 585–590.
- [28] Michael Levitt and Arieh Warshel. "Computer simulation of protein folding". In: *Nature* 253.5494 (1975), pp. 694–698.

- [29] Richard Car and Mark Parrinello. “Unified approach for molecular dynamics and density-functional theory”. In: *Physical review letters* 55.22 (1985), p. 2471.
- [30] Max Born and Robert Oppenheimer. “Zur quantentheorie der molekeln”. In: *Annalen der physik* 389.20 (1927), pp. 457–484.
- [31] John E Stone *et al.* “GPU-accelerated molecular modeling coming of age”. In: *Journal of Molecular Graphics and Modelling* 29.2 (2010), pp. 116–125.
- [32] Stefan Chmiela *et al.* “Towards exact molecular dynamics simulations with machine-learned force fields”. In: *Nature communications* 9.1 (2018), pp. 1–10.
- [33] W Patrick Walters, Matthew T Stahl, and Mark A Murcko. “Virtual screening—an overview”. In: *Drug discovery today* 3.4 (1998), pp. 160–178.
- [34] Don Howard. “Who invented the “Copenhagen Interpretation”? A study in mythology”. In: *Philosophy of Science* 71.5 (2004), pp. 669–682.
- [35] Andreas K Theophilou. “The energy density functional formalism for excited states”. In: *Journal of Physics C: Solid State Physics* 12.24 (1979), p. 5419.
- [36] R Gaudoin and K Burke. “Lack of Hohenberg-Kohn theorem for excited states”. In: *Physical review letters* 93.17 (2004), p. 173001.
- [37] Per-Olov Löwdin. “Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects”. In: *Physical review* 97.6 (1955), p. 1509.
- [38] David C Langreth and MJ Mehl. “Beyond the local-density approximation in calculations of ground-state electronic properties”. In: *Physical Review B* 28.4 (1983), p. 1809.
- [39] Paul AM Dirac. “Note on exchange phenomena in the Thomas atom”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 26. 3. Cambridge University Press. 1930, pp. 376–385.
- [40] WJ Carr Jr. “Energy, specific heat, and magnetic properties of the low-density electron gas”. In: *Physical Review* 122.5 (1961), p. 1437.
- [41] WJ Carr Jr and AA Maradudin. “Ground-state energy of a high-density electron gas”. In: *Physical Review* 133.2A (1964), A371.
- [42] Seymour H. Vosko, Leslie Wilk, and Marwan Nusair. “Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis”. In: *Canadian Journal of physics* 58.8 (1980), pp. 1200–1211.
- [43] John P Perdew and Yue Wang. “Pair-distribution function and its coupling-constant average for the spin-polarized electron gas”. In: *Physical Review B* 46.20 (1992), p. 12947.

- [44] Axel D. Becke. "Density-functional exchange-energy approximation with correct asymptotic behavior". In: *Phys. Rev. A* 38.6 (1988), p. 3098.
- [45] Chengteh Lee, Weitao Yang, and Robert G. Parr. "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density". In: *Physical review B* 37.2 (1988), p. 785.
- [46] J Harris. "Adiabatic-connection approach to Kohn-Sham theory". In: *Physical Review A* 29.4 (1984), p. 1648.
- [47] Yan Zhao, Benjamin J. Lynch, and Donald G. Truhlar. "Doubly hybrid meta DFT: New multi-coefficient correlation and density functional methods for thermochemistry and thermochemical kinetics". In: *The Journal of Physical Chemistry A* 108.21 (2004), pp. 4786–4791.
- [48] Stefan Grimme. "Semiempirical hybrid density functional with perturbative second-order correlation". In: *The Journal of chemical physics* 124.3 (2006), p. 034108.
- [49] Tobias Schwabe and Stefan Grimme. "Towards chemical accuracy for the thermodynamics of large molecules: new hybrid density functionals including non-local correlation effects". In: *Physical Chemistry Chemical Physics* 8.38 (2006), pp. 4398–4401.
- [50] Stefan Kurth, John P Perdew, and Peter Blaha. "Molecular and solid-state tests of density functional approximations: LSD, GGAs, and meta-GGAs". In: *International journal of quantum chemistry* 75.4-5 (1999), pp. 889–909.
- [51] John C Slater. "The theory of complex spectra". In: *Physical Review* 34.10 (1929), p. 1293.
- [52] H.A. Kramers. "L'interaction entre les atomes magnétogènes dans un cristal paramagnétique". In: *Physica* 1.1-6 (1934), pp. 182–192.
- [53] P.W. Anderson. "Antiferromagnetism. Theory of superexchange interaction". In: *Physical Review* 79.2 (1950), p. 350.
- [54] JH Van Vleck. "Scientific Books: The Theory of Electric and Magnetic Susceptibilities". In: *Science* 76 (1932), pp. 326–328.
- [55] Werner Heisenberg. "Zur theorie des ferromagnetismus". In: *Original Scientific Papers Wissenschaftliche Originalarbeiten*. Springer, 1985, pp. 580–597.
- [56] Paul Adrien Maurice Dirac. "Quantum mechanics of many-electron systems". In: *Proc. R. Soc. Lond. A* 123.792 (1929), pp. 714–733.
- [57] Louis Noodleman. "Valence bond description of antiferromagnetic coupling in transition metal dimers". In: *The Journal of Chemical Physics* 74.10 (1981), pp. 5737–5743.

- [58] Frank Neese. "Definition of Corresponding Orbitals and the Diradical Character in Broken Symmetry DFT Calculations on Spin Coupled Systems". In: *J. Phys. Chem. Solids*. Design, Characterization and Modelling of Molecule-Based Magnetic Materials Proceedings of Symposium K, EMRS Spring Meeting, June 2003, Strasbourg, France 65.4 (Apr. 2004), pp. 781–785. ISSN: 0022-3697. DOI: 10.1016/j.jpcs.2003.11.015.
- [59] Dermot Hegarty and Michael A. Robb. "Application of unitary group methods to configuration interaction calculations". In: *Molecular Physics* 38.6 (1979), pp. 1795–1812.
- [60] Hans-Joachim Werner and Peter J. Knowles. "An efficient internally contracted multiconfiguration–reference configuration interaction method". In: *The Journal of chemical physics* 89.9 (1988), pp. 5803–5814.
- [61] C. Angeli *et al.* "Introduction of n-electron valence states for multireference perturbation theory". In: *The Journal of Chemical Physics* 114.23 (2001), pp. 10252–10264.
- [62] K Yamaguchi, Y Takahara, and T Fueno. "Applied quantum chemistry". In: *Reidel, Dordrecht* 155 (1986).
- [63] Toru Saito and Walter Thiel. "Analytical gradients for density functional calculations with approximate spin projection". In: *The Journal of Physical Chemistry A* 116.44 (2012), pp. 10864–10869.
- [64] Hrant P Hratchian. *Communication: An efficient analytic gradient theory for approximate spin projection methods*. 2013.
- [65] M Shoji *et al.* "A general algorithm for calculation of Heisenberg exchange integrals J in multispin systems". In: *Chemical physics letters* 432.1-3 (2006), pp. 343–347.
- [66] Yihan Shao, Martin Head-Gordon, and Anna I Krylov. "The spin–flip approach within time-dependent density functional theory: Theory and applications to diradicals". In: *The Journal of chemical physics* 118.11 (2003), pp. 4807–4818.
- [67] Nicholas J Mayhall and Martin Head-Gordon. "Computational Quantum Chemistry for Multiple-Site Heisenberg Spin Couplings Made Simple: Still Only One Spin–Flip Required". In: *The journal of physical chemistry letters* 6.10 (2015), pp. 1982–1988.
- [68] Jean-Paul Malrieu and Georges Trinquier. "A recipe for geometry optimization of diradicalar singlet states from broken-symmetry calculations". In: *The Journal of Physical Chemistry A* 116.31 (2012), pp. 8226–8237.

- [69] Ibério de PR Moreira and Francesc Illas. "A unified view of the theoretical description of magnetic coupling in molecular chemistry and solid state physics". In: *Physical Chemistry Chemical Physics* 8.14 (2006), pp. 1645–1659.
- [70] Pragya Verma, Rémi Maurice, and Donald G Truhlar. "Adsorbate-Induced Changes in Magnetic Interactions in Fe₂ (dobdc) with Adsorbed Hydrocarbon Molecules". In: *The Journal of Physical Chemistry C* 120.18 (2016), pp. 9933–9948.
- [71] Mirosław Antoni Czarnecki *et al.* "Advances in molecular structure and interaction studies using near-infrared spectroscopy". In: *Chemical reviews* 115.18 (2015), pp. 9707–9744.
- [72] Ove Christiansen. "Vibrational structure theory: new vibrational wave function methods for calculation of anharmonic vibrational energies and vibrational contributions to molecular properties". In: *Physical Chemistry Chemical Physics* 9.23 (2007), pp. 2942–2953.
- [73] Attila G Csaszar. "Anharmonic molecular force fields". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2.2 (2012), pp. 273–289.
- [74] Vincenzo Barone, Malgorzata Biczysko, and Julien Bloino. "Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation". In: *Physical Chemistry Chemical Physics* 16.5 (2014), pp. 1759–1787.
- [75] Eduard Schreiner *et al.* "Dynamical magnetostructural properties of Anabaena ferredoxin". In: *Proceedings of the National Academy of Sciences* 104.52 (2007), pp. 20725–20730.
- [76] Nisanth N. Nair *et al.* "Magnetostructural dynamics with the extended broken symmetry formalism: Antiferromagnetic [2Fe-2S] complexes". In: *Journal of chemical theory and computation* 4.8 (2008), pp. 1174–1188.
- [77] D. Bovi and L. Guidoni. "Magnetic coupling constants and vibrational frequencies by extended broken symmetry approach with hybrid functionals". In: *The Journal of Chemical Physics* 137.11 (2012), p. 114107.
- [78] Md Ehesan Ali, Volker Staemmler, and Dominik Marx. "Magnetostructural dynamics of Rieske versus ferredoxin iron–sulfur cofactors". In: *Physical Chemistry Chemical Physics* 17.9 (2015), pp. 6289–6296.
- [79] Matteo Barborini and Leonardo Guidoni. "Geometries of low spin states of multi-centre transition metal complexes through extended broken symmetry variational Monte Carlo". In: *The Journal of chemical physics* 145.12 (2016), p. 124107.
- [80] MJ Carney *et al.* "Ground spin state variability in [Fe₄S₄ (SR)₄]³⁻. Synthetic analogs of the reduced clusters in ferredoxins and other iron-sulfur proteins: cases of extreme sensitivity of electronic state and structure to extrinsic factors". In: *Journal of the American Chemical Society* 110.18 (1988), pp. 6084–6095.

- [81] Helmut Beinert, Richard H Holm, and Eckard Münck. "Iron-sulfur clusters: nature's modular, multipurpose structures". In: *Science* 277.5326 (1997), pp. 653–659.
- [82] Deborah C. Johnson *et al.* "Structure, function, and formation of biological iron-sulfur clusters". In: *Annu. Rev. Biochem.* 74 (2005), pp. 247–281.
- [83] Rhonda A Torres *et al.* "Density functional and reduction potential calculations of Fe₄S₄ clusters". In: *Journal of the American Chemical Society* 125.7 (2003), pp. 1923–1936.
- [84] Shibing Chu *et al.* "Effects of Static Correlation between Spin Centers in Multi-center Transition Metal Complexes". In: *Journal of chemical theory and computation* 13.10 (2017), pp. 4675–4683.
- [85] Francesco Cappelluti, Luigi Bencivenni, and Leonardo Guidoni. "Spin-symmetrised structures and vibrational frequencies of iron-sulfur clusters". In: *Physical Chemistry Chemical Physics* 22.29 (2020), pp. 16655–16664.
- [86] Frank Neese. "The ORCA Program System". en. In: *WIREs. Comput. Mol. Sci.* 2.1 (Jan. 2012), pp. 73–78. ISSN: 1759-0884. DOI: 10.1002/wcms.81.
- [87] Gene Golub and William Kahan. "Calculating the singular values and pseudo-inverse of a matrix". In: *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis* 2.2 (1965), pp. 205–224.
- [88] Robert E Rudd and Jeremy Q Broughton. "Coarse-grained molecular dynamics and the atomic limit of finite elements". In: *Physical review B* 58.10 (1998), R5893.
- [89] Paul AM Dirac. "The Lagrangian in quantum mechanics". In: *Feynman's Thesis—A New Approach To Quantum Theory*. World Scientific, 2005, pp. 111–119.
- [90] Richard Phillips Feynman. "Space-time approach to non-relativistic quantum mechanics". In: *Feynman's Thesis—A New Approach To Quantum Theory*. World Scientific, 2005, pp. 71–109.
- [91] Mohsen Razavy. *Quantum theory of tunneling*. World Scientific, 2003.
- [92] Philippe H Hünenberger. "Thermostat algorithms for molecular dynamics simulations". In: *Advanced computer simulation*. Springer, 2005, pp. 105–149.
- [93] Loup Verlet. "Computer" experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules". In: *Physical review* 159.1 (1967), p. 98.
- [94] Judith A Harrison *et al.* "Review of force fields and intermolecular potentials used in atomistic computational materials research". In: *Applied Physics Reviews* 5.3 (2018), p. 031104.
- [95] David A Case *et al.* "The Amber biomolecular simulation programs". In: *Journal of computational chemistry* 26.16 (2005), pp. 1668–1688.

- [96] Richard A Buckingham. "The classical equation of state of gaseous helium, neon and argon". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 168.933 (1938), pp. 264–283.
- [97] Paul P. Ewald. "Die Berechnung optischer und elektrostatischer Gitterpotentiale". In: *Annalen der physik* 369.3 (1921), pp. 253–287.
- [98] Tom Darden, Darrin York, and Lee Pedersen. "Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems". In: *The Journal of chemical physics* 98.12 (1993), pp. 10089–10092.
- [99] Ulrich Essmann *et al.* "A smooth particle mesh Ewald method". In: *The Journal of chemical physics* 103.19 (1995), pp. 8577–8593.
- [100] Roger W. Hockney and James W. Eastwood. *Computer simulation using particles*. crc Press, 1988.
- [101] William H Press *et al.* *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.
- [102] Matteo Frigo and Steven G Johnson. "FFTW: An adaptive software architecture for the FFT". In: *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*. Vol. 3. IEEE. 1998, pp. 1381–1384.
- [103] Endong Wang *et al.* "Intel math kernel library". In: *High-Performance Computing on the Intel® Xeon Phi™*. Springer, 2014, pp. 167–188.
- [104] Christopher I. Bayly *et al.* "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model". In: *The Journal of Physical Chemistry* 97.40 (1993), pp. 10269–10280.
- [105] Brent H Besler, Kenneth M Merz Jr, and Peter A Kollman. "Atomic charges derived from semiempirical methods". In: *Journal of computational chemistry* 11.4 (1990), pp. 431–439.
- [106] A van Bondi. "van der Waals volumes and radii". In: *The Journal of physical chemistry* 68.3 (1964), pp. 441–451.
- [107] Alessandro Laio *et al.* "A variational definition of electrostatic potential derived charges". In: *The Journal of Physical Chemistry B* 108.23 (2004), pp. 7963–7968.
- [108] Arieh Warshel and Michael Levitt. "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme". In: *Journal of molecular biology* 103.2 (1976), pp. 227–249.
- [109] Teodoro Laino *et al.* "An efficient real space multigrid QM/MM electrostatic coupling". In: *Journal of Chemical Theory and Computation* 1.6 (2005), pp. 1176–1184.

- [110] Hai Lin and Donald G Truhlar. "QM/MM: what have we learned, where are we, and where do we go from here?" In: *Theoretical Chemistry Accounts* 117.2 (2007), pp. 185–199.
- [111] Piotr Cieplak *et al.* "Polarization effects in molecular mechanical force fields". In: *Journal of Physics: Condensed Matter* 21.33 (2009), p. 333102.
- [112] Dmitry V Matyushov. "Solvent reorganization energy of electron-transfer reactions in polar solvents". In: *The Journal of chemical physics* 120.16 (2004), pp. 7532–7556.
- [113] Haibo Yu and Wilfred F Van Gunsteren. "Accounting for polarization in molecular simulation". In: *Computer Physics Communications* 172.2 (2005), pp. 69–85.
- [114] Thomas A Halgren and Wolfgang Damm. "Polarizable force fields". In: *Current opinion in structural biology* 11.2 (2001), pp. 236–242.
- [115] Arieh Warshel, Mitsunori Kato, and Andrei V Pisliakov. "Polarizable force fields: history, test cases, and prospects". In: *Journal of Chemical Theory and Computation* 3.6 (2007), pp. 2034–2045.
- [116] Christopher M Baker. "Polarizable force fields for molecular dynamics simulations of biomolecules". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 5.2 (2015), pp. 241–254.
- [117] Robert T Sanderson. "Electronegativity and bond energy". In: *Journal of the American Chemical Society* 105.8 (1983), pp. 2259–2261.
- [118] Filippo Lipparini and Vincenzo Barone. "Polarizable force fields and polarizable continuum model: a fluctuating charges/PCM approach. 1. theory and implementation". In: *Journal of chemical theory and computation* 7.11 (2011), pp. 3711–3724.
- [119] Sandeep Patel, Alexander D Mackerell Jr, and Charles L Brooks III. "CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model". In: *Journal of computational chemistry* 25.12 (2004), pp. 1504–1514.
- [120] AJ Stone. "Distributed multipole analysis, or how to describe a molecular charge distribution". In: *Chemical Physics Letters* 83.2 (1981), pp. 233–239.
- [121] B Th Thole. "Molecular polarizabilities calculated with a modified dipole interaction". In: *Chemical Physics* 59.3 (1981), pp. 341–350.
- [122] Jay W. Ponder *et al.* "Current status of the AMOEBA polarizable force field". In: *The journal of physical chemistry B* 114.8 (2010), pp. 2549–2564.
- [123] Paul Drude. *The theory of optics*. Longmans, Green and Company, 1901.

- [124] Guillaume Lamoureux and Benoit Roux. "Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm". In: *The Journal of chemical physics* 119.6 (2003), pp. 3025–3039.
- [125] Justin A Lemkul *et al.* "An empirical polarizable force field based on the classical drude oscillator model: development history and recent applications". In: *Chemical reviews* 116.9 (2016), pp. 4983–5013.
- [126] Robert S. Mulliken. "Electronic population analysis on LCAO–MO molecular wave functions. I". In: *The Journal of Chemical Physics* 23.10 (1955), pp. 1833–1840.
- [127] Benedict Leimkuhler and Sebastian Reich. *Simulating hamiltonian dynamics*. Vol. 14. Cambridge university press, 2004.
- [128] Shuichi Nosé. "A unified formulation of the constant temperature molecular dynamics methods". In: *The Journal of chemical physics* 81.1 (1984), pp. 511–519.
- [129] William G. Hoover. "Canonical dynamics: equilibrium phase-space distributions". In: *Physical review A* 31.3 (1985), p. 1695.
- [130] D McQuarrie. "Statistical mechanics. 2000". In: *Sausalito, Calif.: University Science Books* 12 (2004), p. 641.
- [131] Enrico Matteoli and G Ali Mansoori. "A simple expression for radial distribution functions of pure fluids and mixtures". In: *The Journal of chemical physics* 103.11 (1995), pp. 4672–4677.
- [132] Adolf Fick. "Ueber diffusion". In: *Annalen der Physik* 170.1 (1855), pp. 59–86.
- [133] Albert Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der physik* 322.8 (1905), pp. 549–560.
- [134] Marian Von Smoluchowski. "Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen". In: *Annalen der physik* 326.14 (1906), pp. 756–780.
- [135] Edward J Maginn. "Atomistic simulation of the thermodynamic and transport properties of ionic liquids". In: *Accounts of chemical research* 40.11 (2007), pp. 1200–1207.
- [136] Florian Weigend and Reinhart Ahlrichs. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy". In: *Phys. Chem. Chem. Phys.* 7.18 (2005), pp. 3297–3305.
- [137] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized gradient approximation made simple". In: *Phys. Rev. Lett.* 77.18 (1996), p. 3865.

- [138] Nicholas C. Handy and Aron J. Cohen. "Left-right correlation energy". In: *Mol. Phys.* 99.5 (2001), pp. 403–412.
- [139] John P. Perdew. "Density-functional approximation for the correlation energy of the inhomogeneous electron gas". In: *Phys. Rev. B* 33.12 (1986), p. 8822.
- [140] Axel D. Becke. "Becke's three parameter hybrid method using the LYP correlation functional". In: *J. Chem. Phys.* 98 (1993), pp. 5648–5652.
- [141] Yan Zhao and Donald G. Truhlar. "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals". In: *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* 120.1 (2008), pp. 215–241.
- [142] Stefan Grimme. "Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies". In: *The Journal of chemical physics* 118.20 (2003), pp. 9095–9102.
- [143] Lars Goerigk and Stefan Grimme. "Double-hybrid density functionals". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 4.6 (2014), pp. 576–600.
- [144] EJ Baerends, DE Ellis, and P Ros. "Self-consistent molecular Hartree–Fock–Slater calculations I. The computational procedure". In: *Chemical Physics* 2.1 (1973), pp. 41–51.
- [145] Jerry L Whitten. "Coulombic potential energy integrals and approximations". In: *The Journal of Chemical Physics* 58.10 (1973), pp. 4496–4501.
- [146] Florian Weigend. "Hartree–Fock exchange fitting basis sets for H to Rn". In: *Journal of computational chemistry* 29.2 (2008), pp. 167–175.
- [147] Florian Weigend. "Accurate Coulomb-fitting basis sets for H to Rn". In: *Physical chemistry chemical physics* 8.9 (2006), pp. 1057–1065.
- [148] Charles H Koelbel *et al.* *The high performance Fortran handbook*. MIT press, 1994.
- [149] Guido Rossum. "Python reference manual". In: (1995).
- [150] W. Robert Scheidt, Stephen M. Durbin, and J. Timothy Sage. "Nuclear resonance vibrational spectroscopy–NRVS". In: *Journal of inorganic biochemistry* 99.1 (2005), pp. 60–71.
- [151] John R. de Laeter *et al.* "Atomic weights of the elements. Review 2000 (IUPAC Technical Report)". In: *Pure and applied chemistry* 75.6 (2003), pp. 683–800.

- [152] Filipe Teixeira. *VibAnalysis - Tools for performing vibrational analysis on molecular systems*. July 4, 2017. URL: <https://github.com/teixeirafilipe/vibAnalysis> (visited on 07/04/2017).
- [153] Filipe Teixeira and M Natália DS Cordeiro. "Improving Vibrational Mode Interpretation Using Bayesian Regression". In: *Journal of chemical theory and computation* 15.1 (2018), pp. 456–470.
- [154] Michał H Jamróz. "Vibrational energy distribution analysis (VEDA): scopes and limitations". In: *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 114 (2013), pp. 220–230.
- [155] Pauli Virtanen *et al.* "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* (2020). DOI: <https://doi.org/10.1038/s41592-019-0686-2>.
- [156] Dieter Kraft. "A software package for sequential quadratic programming". In: *Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt* (1988).
- [157] Marucha Lalee, Jorge Nocedal, and Todd Plantenga. "On the implementation of an algorithm for large-scale equality constrained optimization". In: *SIAM Journal on Optimization* 8.3 (1998), pp. 682–706.
- [158] David J. Wales and Jonathan P.K. Doye. "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms". In: *The Journal of Physical Chemistry A* 101.28 (1997), pp. 5111–5116.
- [159] Jürg Hutter *et al.* "CP2K: atomistic simulations of condensed matter systems". In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 4.1 (2014), pp. 15–25.
- [160] RHWJ Ditchfield, W J_ Hehre, and John A Pople. "Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules". In: *The Journal of Chemical Physics* 54.2 (1971), pp. 724–728.
- [161] Michael J Frisch, John A Pople, and J Stephen Binkley. "Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets". In: *The Journal of chemical physics* 80.7 (1984), pp. 3265–3269.
- [162] Stefan Goedecker, Michael Teter, and Jürg Hutter. "Separable dual-space Gaussian pseudopotentials". In: *Physical Review B* 54.3 (1996), p. 1703.
- [163] Martín Abadi *et al.* "Tensorflow: A system for large-scale machine learning". In: *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*. 2016, pp. 265–283.

- [164] Fabian Pedregosa *et al.* "Scikit-learn: Machine learning in Python". In: *the Journal of machine Learning research* 12 (2011), pp. 2825–2830.
- [165] Ryutaro Tsuchida. "Absorption spectra of co-ordination compounds. I". In: *Bulletin of the Chemical Society of Japan* 13.5 (1938), pp. 388–400.
- [166] M.A. Bobrik, Keith O. Hodgson, and R.H. Holm. "Inorganic derivatives of iron-sulfide-thiolate dimers and tetramers. Structures of tetrachloro- μ -disulfido-diferrate (III) and tetrakis (chloro- μ -3-sulfido-iron) dianions". In: *Inorganic Chemistry* 16.8 (1977), pp. 1851–1858.
- [167] Matt C. Smith *et al.* "Normal-Mode Analysis of $FeCl_4^-$ and $Fe_2S_2Cl_4^{2-}$ via Vibrational Mössbauer, Resonance Raman, and FT-IR Spectroscopies". In: *Inorganic chemistry* 44.16 (2005), pp. 5562–5570.
- [168] Geoffrey B. Wong, M.A. Bobrik, and R.H. Holm. "Inorganic derivatives of iron sulfide thiolate dimers and tetramers: synthesis and properties of the halide series $[Fe_2S_2X_4]^{2-}$ and $[Fe_4S_4X_4]^{2-}$ (X= chlorine, bromine, iodine)". In: *Inorganic Chemistry* 17.3 (1978), pp. 578–584.
- [169] Christopher R. Sharp, Jeremiah S. Duncan, and Sonny C. Lee. " $[Fe_4S_4]^q$ cubane clusters (q = 4+, 3+, 2+) with terminal amide ligands". In: *Inorganic chemistry* 49.14 (2010), pp. 6697–6705.
- [170] Clarence Zener. "Interaction between the d-shells in the transition metals. II. Ferromagnetic compounds of manganese with perovskite structure". In: *Physical Review* 82.3 (1951), p. 403.
- [171] Patric W. Anderson and H. Hasegawa. "Considerations on double exchange". In: *Physical Review* 100.2 (1955), p. 675.
- [172] J-J Girerd. "Electron transfer between magnetic ions in mixed valence binuclear systems". In: *The Journal of Chemical Physics* 79.4 (1983), pp. 1766–1775.
- [173] Louis Noodleman and Evert Jan Baerends. "Electronic structure, magnetic properties, ESR, and optical spectra for 2-iron ferredoxin models by LCAO-X. alpha. valence bond theory". In: *Journal of the American Chemical Society* 106.8 (1984), pp. 2316–2327.
- [174] Golam Moula *et al.* "Synthesis of an All-Ferric Cuboidal Iron–Sulfur Cluster $[Fe^{III}_4S_4(SAr)_4]$ ". In: *Angewandte Chemie International Edition* 57.36 (2018), pp. 11594–11597.
- [175] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.
- [176] SG Andrade, Luisa CS Gonçalves, and Francisco Elias Jorge. "Scaling factors for fundamental vibrational frequencies and zero-point energies obtained from HF, MP2, and DFT/DZP and TZP harmonic frequencies". In: *W J. Stalin Journal of Molecular Structure: THEOCHEM* 864.1-3 (2008), pp. 20–25.

- [177] Ye Mei *et al.* “Numerical study on the partitioning of the molecular polarizability into fluctuating charge and induced atomic dipole contributions”. In: *The Journal of Physical Chemistry A* 119.22 (2015), pp. 5865–5882.
- [178] Steve Plimpton. “Fast parallel algorithms for short-range molecular dynamics”. In: *Journal of computational physics* 117.1 (1995), pp. 1–19.
- [179] D. Wolf *et al.* “Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r^{-1} summation”. In: *The Journal of chemical physics* 110.17 (1999), pp. 8254–8282.
- [180] Christopher J. Fennell and J. Daniel Gezelter. “Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics”. In: *The Journal of chemical physics* 124.23 (2006), p. 234104.
- [181] Remco Hens and Thijs J.H. Vlugt. “Molecular Simulation of Vapor–Liquid Equilibria Using the Wolf Method for Electrostatic Interactions”. In: *Journal of Chemical & Engineering Data* 63.4 (2017), pp. 1096–1102.
- [182] John A. Baker and Jonathan D. Hirst. “Accelerating electrostatic pair methods on graphical processing units to study molecules in supercritical carbon dioxide”. In: *Faraday discussions* 169 (2014), pp. 343–357.
- [183] Christian Waibel, Mathias Simon Feinler, and Joachim Gross. “A Modified Shifted Force Approach to the Wolf Summation”. In: *Journal of chemical theory and computation* (2018).
- [184] Henrik G. Petersen. “Accuracy and efficiency of the particle mesh Ewald method”. In: *The Journal of chemical physics* 103.9 (1995), pp. 3668–3679.
- [185] Markus Deserno and Christian Holm. “How to mesh up Ewald sums. II. An accurate error estimate for the particle–particle–particle-mesh algorithm”. In: *The Journal of chemical physics* 109.18 (1998), pp. 7694–7701.
- [186] Leandro Martínez *et al.* “PACKMOL: a package for building initial configurations for molecular dynamics simulations”. In: *Journal of computational chemistry* 30.13 (2009), pp. 2157–2164.
- [187] Warren J. Hehre, Robert Ditchfield, and John A. Pople. “Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules”. In: *The Journal of Chemical Physics* 56.5 (1972), pp. 2257–2261.
- [188] Axel D. Becke. “Density-functional thermochemistry. III. The role of exact exchange”. In: *The Journal of chemical physics* 98.7 (1993), pp. 5648–5652.
- [189] Ansgar Schäfer, Hans Horn, and Reinhart Ahlrichs. “Fully optimized contracted Gaussian basis sets for atoms Li to Kr”. In: *The Journal of Chemical Physics* 97.4 (1992), pp. 2571–2577.

- [190] Ansgar Schäfer, Christian Huber, and Reinhart Ahlrichs. “Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr”. In: *The Journal of Chemical Physics* 100.8 (1994), pp. 5829–5835.
- [191] Furio Ercolessi and James B. Adams. “Interatomic potentials from first-principles calculations: the force-matching method”. In: *EPL (Europhysics Letters)* 26.8 (1994), p. 583.
- [192] Marco Masia, Elvira Guàrdia, and Paolo Nicolini. “The force matching approach to multiscale simulations: Merits, shortcomings, and future perspectives”. In: *International journal of quantum chemistry* 114.16 (2014), pp. 1036–1040.
- [193] DA Long. “Intensities in Raman spectra I. A bond polarizability theory”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 217.1129 (1953), pp. 203–221.
- [194] Richard P Smith and Earl M Mortensen. “Bond and molecular polarizability tensors. I. Mathematical treatment of bond tensor additivity”. In: *The Journal of Chemical Physics* 32.2 (1960), pp. 502–507.
- [195] S Guha *et al.* “Empirical bond polarizability model for fullerenes”. In: *Physical Review B* 53.19 (1996), p. 13106.
- [196] Piero Macchi and Anna Krawczuk. “The polarizability of organometallic bonds”. In: *Computational and theoretical chemistry* 1053 (2015), pp. 165–172.
- [197] Hao Hu, Zhenyu Lu, and Weitao Yang. “Fitting molecular electrostatic potentials from quantum mechanical calculations”. In: *Journal of chemical theory and computation* 3.3 (2007), pp. 1004–1013.
- [198] Richard FW Bader. “Atoms in molecules”. In: *Accounts of Chemical Research* 18.1 (1985), pp. 9–15.
- [199] ISO. *ISO/IEC 14882:2014 Information technology — Programming languages — C++*. Fourth. Dec. 2014. URL: http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=64029.
- [200] William L. Jorgensen *et al.* “Comparison of simple potential functions for simulating liquid water”. In: *The Journal of chemical physics* 79.2 (1983), pp. 926–935.
- [201] Sewall Wright. “Correlation and causation”. In: *J. agric. Res.* 20 (1921), pp. 557–580.
- [202] Stefan Grimme *et al.* “A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu”. In: *The Journal of chemical physics* 132.15 (2010), p. 154104.
- [203] Joost VandeVondele and Juerg Hutter. “Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases”. In: *The Journal of chemical physics* 127.11 (2007), p. 114105.

- [204] Fred L. Hirshfeld. "Bonded-atom fragments for describing molecular charge densities". In: *Theoretica chimica acta* 44.2 (1977), pp. 129–138.
- [205] M. Tanaka *et al.* "Recommended table for the density of water between 0 °C and 40 °C based on recent experimental reports". In: *Metrologia* 38.4 (2001), p. 301.
- [206] Kenneth R. Harris and Lawrence A. Woolf. "Pressure and temperature dependence of the self diffusion coefficient of water and oxygen-18 water". In: *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* 76 (1980), pp. 377–385.
- [207] I-Chun Lin *et al.* "Structure and Dynamics of Liquid Water from ab Initio Molecular Dynamics Comparison of BLYP, PBE, and revPBE Density Functionals with and without van der Waals Corrections". In: *Journal of chemical theory and computation* 8.10 (2012), pp. 3902–3910.
- [208] Michael W Mahoney and William L Jorgensen. "Diffusion constant of the TIP5P model of liquid water". In: *The Journal of Chemical Physics* 114.1 (2001), pp. 363–366.
- [209] William Sutherland. "LXXV. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 9.54 (1905), pp. 781–785.
- [210] Tianying Yan *et al.* "Molecular dynamics simulation of ionic liquids: The effect of electronic polarizability". In: *The Journal of Physical Chemistry B* 108.32 (2004), pp. 11877–11881.
- [211] Christian Schröder. "Comparing reduced partial charge models with polarizable simulations of ionic liquids". In: *Physical Chemistry Chemical Physics* 14.9 (2012), pp. 3089–3102.
- [212] Lorenzo Gontrani *et al.* "A structural and theoretical study of the alkylammonium nitrates forefather: Liquid methylammonium nitrate". In: *Chemical Physics Letters* 684 (2017), pp. 304–309.
- [213] M Campetella *et al.* "Structural and vibrational study of 2-MethoxyEthylAmmonium Nitrate (2-OMeEAN): Interpretation of experimental results with ab initio molecular dynamics". In: *The Journal of Chemical Physics* 145.2 (2016), p. 024507.
- [214] Junmei Wang *et al.* "Development and testing of a general amber force field". In: *Journal of computational chemistry* 25.9 (2004), pp. 1157–1174.
- [215] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J.C. Berendsen. "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes". In: *Journal of Computational Physics* 23.3 (1977), pp. 327–341.

- [216] Till Cremer *et al.* "Towards a Molecular Understanding of Cation–Anion Interactions—Probing the Electronic Structure of Imidazolium Ionic Liquids by NMR Spectroscopy, X-ray Photoelectron Spectroscopy and Theoretical Calculations". In: *Chemistry—A European Journal* 16.30 (2010), pp. 9018–9033.
- [217] Oldamur Hollóczy *et al.* "On the origin of ionicity in ionic liquids. Ion pairing versus charge transfer". In: *Physical Chemistry Chemical Physics* 16.32 (2014), pp. 16880–16890.
- [218] Dmitry Bedrov *et al.* "Molecular dynamics simulations of ionic liquids and electrolytes using polarizable force fields". In: *Chemical reviews* 119.13 (2019), pp. 7940–7995.
- [219] Carolyn E Lubner *et al.* "Wiring an [FeFe]-hydrogenase with photosystem I for light-induced hydrogen production". In: *Biochemistry* 49.48 (2010), pp. 10264–10266.
- [220] Thomas Nann *et al.* "Water splitting by visible light: a nanophotocathode for hydrogen production". In: *Angewandte Chemie International Edition* 49.9 (2010), pp. 1574–1577.
- [221] Genevieve Blondin and Jean Jacques Girerd. "Interplay of electron exchange and electron transfer in metal polynuclear complexes in proteins or chemical models". In: *Chemical Reviews* 90.8 (1990), pp. 1359–1376.
- [222] Edward Uhler Condon, EU Condon, and GH Shortley. *The theory of atomic spectra*. Cambridge University Press, 1951.
- [223] Arno Bohm and Mark Loewe. *Quantum mechanics: foundations and applications*. Vol. 3. Springer New York, 1986.
- [224] Lawrie B. Skinner *et al.* "Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range". In: *The Journal of chemical physics* 138.7 (2013), p. 074506.
- [225] Francesco Cappelluti *et al.* "F-RESP polarizable force field and its application on liquid water". In: (in preparation).
- [226] M Campetella *et al.* "Physical-chemical studies on putrescine (butane-1, 4-diamine) and its solutions: Experimental and computational investigations". In: *Journal of Molecular Liquids* 322 (2021), p. 114568.
- [227] O Palumbo *et al.* "New insights into chloromethyl-oxirane and chloromethyl-thiirane in liquid and solid phase from low-temperature infrared spectroscopy and ab initio modeling". In: *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 247 (2021), p. 119061.

- [228] Nicolas Nieuwjaer *et al.* "IRMPD spectroscopy and quantum chemistry calculations on mono-and bi-metallic complexes of acetylacetonate ligands with aluminum, iron, and ruthenium ions". In: *The Journal of Chemical Physics* 153.23 (2020), p. 234303.
- [229] Marco Campetella, Francesco Cappelluti, and Lorenzo Gontrani. "Medium range interactions evidences in compounds with aliphatic lateral chain: 1-pentanoic acid, 1-pentanol and pentylammonium nitrate as test cases". In: *Chemical Physics Letters* 734 (2019), p. 136738.