

Université de Lille
École doctorale **ED 104 - Sciences de la Matière, du Rayonnement et de l'Environnement**
Spécialité **Milieux dilués et optique**
Unité de recherche **Laboratoire PhLAM**

Author:
Yassine Bouchafra
Defense:
25 September 2019

**On the performance of subsystem approaches
to model heavy element species in solution**

Supervisors: André SEVERO PEREIRA GOMES and Valérie Vallet

Protractors:	Emmanuel FROMAGER	Associate Professor	
		Université de Strasbourg	
	Pina ROMMANIELLO	CNRS Researcher	
		Université Paul Sabatier	
Examiners:	Laurent DELEVOYE	CNRS Researcher Director	Jury President
		Université de Lille	
	Regis GAUTHIER	Professor	
		Université de Rennes	

Cette thèse a été préparée au

Laboratoire PhLAM

Laboratoire PhLAM

CNRS UMR 8523

Université Lille Nord de France

Bâtiment P5

59655 Villeneuve d'Ascq

France

☎ (33)(0)3 20 43 44 84

✉ marc.douay@univ-lille.fr

Site <http://phlam.univ-lille.fr>



Résumé

Afin de comprendre les mesures spectroscopiques, il est important de comprendre les processus physiques se déroulant à l'échelle microscopique à cause de la relation qui les relie au comportement des électrons (et noyaux) dans le système. Le traitement des particules similaires exige une description quantique des atomes et des molécules formant le système d'intérêt. Cela signifie que les simulations théoriques, si le système contient des éléments lourds, sont des tâches particulièrement difficiles à cause des effets relativistes. Cette situation motive le développement de plusieurs approches théoriques qui vise à simplifier le traitement d'une partie du système totale au moins.

Dans cette thèse nous investiguons l'utilisation de l'approche Frozen Density Embedding (FDE) pour calculer les propriétés des systèmes complexes. La FDE est formellement une méthode exacte qui nous permet de séparer un système complexe en sous-systèmes et choisir la méthode théorique appropriée pour chaque sous-système. Avec cette séparation nous concentrons l'effort computationnel sur un sous-système où plus en le traitant avec les méthodes de structure électronique relativistes qui incluent le couplage spin-orbit, tant que l'effet des sous-systèmes restant (environnement) sur le système d'intérêt est traité avec des méthodes suffisamment précises.

Notre premier système d'intérêt concerne le calcul quantique des énergies d'ionisation pour des agrégations moléculaires des halogénures microsolvates. La sensibilité de ces énergies par rapport aux changements structurels autour les halogénures et dans les molécules d'eau à été explorée ainsi que l'évolution de ces énergies avec la taille de l'aggrégation. Nos résultats démontrent que la combinaison de EOM-CC relativiste pour le système actif et la DFT pour l'environnement garantie par la FDE donnent des valeurs comparables à ce qui est trouvé dans les expériences.

De même, nous avons exploré la performance de la FDE pour la description des effets de solvant sur les propriétés magnétiques (Tenseurs de couplage spin-spin indirecte et le shielding RMN) pour le complexe $\text{PtTl}(\text{CN})_5$ qui contient une liaison métal-métal entre les centres lourds (Pt, Tl), cette fois-ci avec un traitement DFT relativiste pure. Pour le couplage spin-spin, comme les résultats théoriques précédentes, l'inclusion de la première couche de solvataion est requise pour arriver à un accord semi-quantitative avec l'expérience. Tandis que pour le shielding NMR, la FDE nous permet de réduire significativement le nombre de molécules à inclure dans le sous-système actif. Ceci ouvre la perspective sur l'utilisation de la FDE avec les méthodes de structure électronique pour ce genre de propriétés dans ces cas compliqués.

Abstract

In order to understand spectroscopic measurements, it is important to understand the physical processes taking place at a microscopic scale, since these are related to the behaviour of the electrons (and nuclei) in the system. The treatment of such particles requires one way or another a quantum mechanical treatment of the atoms and molecules that make up a given system of interest. This means that in order to achieve that we must perform theoretical simulations and, if such systems contain heavy elements, this is a particularly difficult task, since we not only have to deal with the large number of particles but also include relativistic effects. These difficulties have motivated the development of several theoretical approaches that simplify the treatment of at least part of the total system.

This thesis investigates the use of the Frozen Density Embedding (FDE) approach to the calculation of molecular properties of complex systems. FDE is a formally exact method with which we can separate a complex molecular system into subsystems and choose the most suitable electronic structure approach to treat each of these. With this separation, we can focus the computational effort into one or a few subsystems of interest and treat them very accurately with relativistic electronic structure methods that include spin-orbit coupling, while the effect of the remaining subsystems (environment) on the system of interest is treated at a sufficiently high level of accuracy.

Our first interest was in the quantum mechanical description of ionisation energies for molecular aggregates of microsolvated halides, such as found in water droplets. We have explored the sensitivity of these energies to structural changes around the halides and among the waters, and how these energies evolve with the size of the aggregate, with our results being in quantitative agreement with experimental data, and we have predicted the ionisation energies of the heaviest of halides, astatide, which is of interest as a radiotherapeutic agent. Our results demonstrate that with the combination of relativistic EOM-CC for the active subsystem and DFT for the environment, afforded by FDE, one can rival with quite sophisticated theoretical approaches based on periodic quasi-particle calculations which are the current state-of-the-art for condensed matter simulations.

We have also explored the performance of FDE for the description of solvent effects on magnetic properties (indirect spin-spin couplings and NMR shielding tensors) for a complex $\text{PtTl}(\text{CN})_5$ containing a metal-metal bond between the heavy centres (Pt, Tl), this time purely at relativistic DFT level. For spin-spin couplings, we have shown that much like prior theoretical results, we require an extensive first hydration shell around the complex, but nevertheless arrive at a semi-quantitative agreement with experiment. For NMR shieldings on the other hand, FDE allows us to significantly reduce the amount of water molecules explicitly added to the active subsystem to the first hydration shell around the Tl atom. This might open up the perspective to employing FDE with more accurate with more accurate electronic structure methods for this property for this class of compounds.

Acknowledgements

First, I want to thank the protractors Emmanuel Fromager and Pina Romaniello the jury Laurent Delevoye and Regis Gauthier who take the time to read, listen and give me a feedback on my work. I would also like to acknowledge my directors for sharing their scientific knowledges, listening, advices and managing skills. Their guidance helped me in all the time of research and writing of this thesis.

I thank my labmates and friends in for the stimulating discussions and for all the fun we have had in the last three years. I am grateful to all the persons who have given me an unconditional support and valuable and timely advice and suggestions for for the completion of my academic background. In particular, I am grateful to Prof. Driss Bria.

Last but not the least, I would like to thank my family: my parents and to my sisters and brother for supporting me spiritually throughout writing this thesis and in my life in general.

"The purpose of (scientific) computing is insight, not numbers."

Richard Hamming

Contents

1	Theoretical chemistry	12
1.1	Schrödinger equation and Born-Oppenheimer approximation	13
1.2	Wave function based methods	14
1.2.1	Hartree Fock approximation	14
1.2.2	Electron correlation	16
1.3	Density functional based methods	21
1.3.1	Kohn-Sham DFT formalism	22
1.3.2	The meaning of orbital energies in KS-DFT	24
1.3.3	The local density approximation	24
1.3.4	The generalised gradient approximation (GGA)	24
1.3.5	Hybrid functionals	25
1.3.6	The statistical average of orbital potentials	26
1.3.7	The DFT compared to the WFT	28
1.4	The molecule in a magnetic field	28
1.4.1	The chemical shift (CA)	29
1.4.2	Spin-spin coupling	31
1.4.3	Magnetic properties in the electronic structure theory	32
1.5	Relativistic effects	33
1.5.1	The Dirac equation	33
1.5.2	Approximations to the Dirac equation	35
2	Embedding methods	38
2.1	Implicit and QM/MM models	38
2.2	Quantum embedding models	44
2.2.1	Frozen Density Embedding (FDE) approach	44
2.2.2	Subsystem DFT	47
2.2.3	Partition DFT	48
2.2.4	FDE extension to wave function/DFT embedding	49
2.2.5	FDE for second-order magnetic properties	50
2.2.6	Challenges	50
3	Ionisation energies of solvated halide Ions with relativistic embedded equation of motion coupled cluster theory	52
3.1	Experimental context	53
4	Further investigations on the electronic structure of halides in water	76
4.1	Droplet size effects on the water ionised bands	77
4.2	Force Field effects on the water and Halide bands	80
4.3	Perspectives	84

5	FDE NMR properties calculation of platinum-thallium bonded complexes.	85
5.1	[(NC) ₅ -Pt-Tl] complexes in the literature	85
5.2	FDE for the (NC) ₅ -Pt-Tl complex	87
5.2.1	Chemical model and relativistic Hamiltonian	88
5.2.2	Computational protocol	89
5.3	Statistics' significance	93
5.4	Final results	94
5.4.1	Spin-spin couplings	94
5.4.2	Shieldings and chemical shifts	95
	Conclusions and perspectives	97
	Bibliography	99

List of Figures

1.1	Schematic representation of the mean-field approach, - refers to a negative charge (an electron).	15
1.2	Examples of dynamical and static correlations inclusion for a twelve-electron system. (Taken from E. Fromager talk in the 2017 summer school [6]).	18
1.3	DFT approaches popularity	26
1.4	Schematic representation of the orbital energy spectrum of the exact Kohn-Sham potential, a typical GGA potential, and the approximate discrete LDA/GGA orbital energy spectra resulting from small and large basis set calculations. Drawn lines for valence orbitals, dashed lines for Rydberg levels. V denotes a valence excitation (HOMO to LUMO transition) and R denotes a Rydberg transition [21].	27
1.5	The trajectories of an electron in the (a) electric field - the trajectory is a parabola (b) magnetic field, perpendicular to the incident velocity the trajectory is a cycloid in a plane perpendicular to the figure	29
1.6	The local effective magnetic field felt by a given nucleus is different from the external one due to the surrounding electrons.	30
1.7	The number of involving bonds in J-coupling is responsible for its sign [57].	31
2.1	Replacement of the solvent molecules with a dielectric continuum	39
2.2	Quantum description could be applied only for a small part in the whole system	41
2.3	Cutting across boundaries is more sophisticated	42
2.4	Capping bonds bring an alternative to cleaving bonds, link atoms (a) and frozen orbitals (b: local self consistent field [95], c: generalized hybrid orbital [96]).	43
2.5	(a): The scaling behaviour of the computational cost with the molecular size (M); (b): The computational cost can be reduced significantly by dividing the whole system into subsystems. (N represents the number of subsystems).	44
2.6	Freeze and thaw cycles procedure	48
3.1	(a) : (Left) Energy level diagram of occupied molecular orbitals of gas-phase H ₂ O [127]. (Right) Orbital pictures of the H ₂ O molecule. Different signs of the wave function are indicated by solid and dashed lines [128, 129]; (b): Photoemission spectra from (top) liquid water, (center) gas-phase water, measured at 60 eV photon energy, and (bottom) difference spectrum [129].	53
3.2	Diagram of lowest electron binding energies of aqueous (a) alkali and (b) halide, comparing experimental (E_{aq}^{PES} , and calculated values ($E_{aq}^{charges}$, E_{aq}^{PCM} , E_{aq}^{thermo}) and experimental E_g and calculated E_g^{calc} values for the gas phase [131].	54

4.1	The stabilisation energies of hydrated iodide as a function of $(n+2)^{-1/3}$ where n is the hydrating water molecules number [135]. The two solid lines (1,2) represent the centrally solvated anion according to the electrostatic model presented in Ref. ([136]) and each one of them extrapolates to a different experimental value for the stabilisation energy from Refs. ([136, 135]).	77
4.2	Snapshots from classical molecular dynamic of configurations centred on the chloride anion with different numbers of water molecules	78
4.3	Electron binding energies spectra for the $[\text{Cl}^-@(\text{H}_2\text{O})_{50}^{\text{tog}}]$ and $[\text{Cl}^-@(\text{H}_2\text{O})_{200}^{\text{sep}}]$ systems from SAOP (SR-ZORA) calculated over 100 snapshots with triple-zeta basis sets.	79
4.4	(Corrected) Electron binding energies spectra for the $[\text{Cl}^-@(\text{H}_2\text{O})_{50}]$ and $[\text{Cl}^-@(\text{H}_2\text{O})_{200}]$ from SAOP (SR-ZORA) calculated over 100 snapshots with triple-zeta basis, compared to experimental results, 11.31 eV from Kurahashi et al. [133] and 11.16 eV from Winter et al. [131].	79
4.5	Comparison of stabilisation energies for Cl^- in water molecules obtained for the first FF16 [139], the corrected FF18 [139] and experimental results [135].	81
4.6	Calculated bands of $\text{Cl}^-@(\text{H}_2\text{O})_{50}^{\text{tog}}$ over 100 snapshots obtained from MDs performed of the two FFs at the level of SAOP (SR-ZORA) with a treatment of the water cluster as one fragment.	82
4.7	Corrected BEs of $\text{Cl}^-@(\text{H}_2\text{O})_{200}^{\text{sep}}$ system from SAOP (SR-ZORA) calculations over FF18 100 snapshots with triple-zeta basis. Vertical lines refer to experimental values [133, 131].	84
5.1	(a-d): compounds I-IV with the formula $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{CN})_n]^{n-}$, $n = 0-3$ and (e): compound V with the formula $[(\text{NC})_5\text{Pt}-\text{Tl}-\text{Pt}(\text{CN})_5]^{3-}$. Pt (red), Tl (green), C (brown) and N (gray).	85
5.2	^{205}Tl NMR spectrum of complex I. Aqueous solution containing 50 mM Tl, 50 mM $\text{Pt}(^{13}\text{CN})_4$, and (a): 100 mM KCN, (b): 100 mM Na^{13}CN . taken from Ref. [148]	86
5.3	Perspective view of the chemical models, (a) : $(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{31}$ and (b) : $(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{64}$	88
5.4	Convergence of FDE (blue square lines) with respect to the supermolecule results (horizontal red lines) for the shieldings (in <i>ppm</i>) for models $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_n]@(\text{H}_2\text{O})_{n-1}$ with increasing n , the number of explicitly treated water molecules.	90
5.5	Convergence of FDE (blue square lines) with respect to the supermolecule results (horizontal red lines) for the J -couplings (in <i>Hz</i>) for the shieldings (in <i>ppm</i>) for models $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_n]@(\text{H}_2\text{O})_{n-1}$ with increasing n , the number of explicitly treated water molecules.	91
5.6	Comparison of the computing time between supermolecule and FDE calculations for models $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_n]@(\text{H}_2\text{O})_{n-1}$ with increasing n , the number of explicitly treated water molecules).	92
5.7	The active systems composed, in addition of $[(\text{NC})_5\text{Pt}-\text{Tl}, (\text{a}) (\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$; (b) $(\text{H}_2\text{O})_{15}]@(\text{H}_2\text{O})_{49}$	92
5.8	$J_{\text{Tl}-\text{Pt}}$ (a) : autocorrelation function and (b) : average evolution, obtained from 256 $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ calculations.	93
5.9	σ_{Tl} (a) : autocorrelation function and (b) : average evolution, obtained from 256 $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ calculations.	93

List of Tables

1.1	Number of Slater determinant that can possibly be formed for a system of $2n$ electrons in $2n$ orbitals.	18
1.2	H_2O energies for different truncated CI wave functions and their weights in the FCI one. (for more details see Table (5.9) in [7].)	19
1.3	Ionisation potentials (in eV) calculated from different ab initio techniques [45] and SAOP methods using Koopmans' and Janack's theorems, and comparison with the experimental ionisation (I_{exp}).	28
1.4	QM scaling behaviour with N (number of electrons) [5].	29
3.1	[131] The water bands are labeled following the convention for a C_{2v} symmetry H_2O molecule. The chloride ion peaks	54
4.1	Electron binding energies (BE, in eV); This work : spin-orbit coupled components of the P states of the hydrated iodide, from EOM and SAOP (DC) calculations on the embedded iodide with triple-zeta basis sets, and water droplet valence bands from SAOP (SR-ZORA) calculations for the $\text{I}^-@(\text{H}_2\text{O})_{50}$ system averaged over MD snapshots from Ref. [134]; Experiment : spin-orbit coupled components of the P states of the solvated iodide and bulk water valence bands from (a) Kurahashi et al. [133], and (b) Winter et al. [131].	76
4.2	Average (Ave) electron binding energies (BE, in eV) of the hydrated chloride from SAOP (SR-ZORA) calculations over 100 snapshots with triple-zeta basis and effect of using separated representation $(\text{H}_2\text{O})_{50}^{sep}$ rather than together one $(\text{H}_2\text{O})_{50}^{tog}$; SD refers to standard deviation and the water bands are labeled following the convention for a C_{2v} symmetry H_2O molecule.	78
4.3	Average (Ave) BE (eV) of the hydrated chloride $\text{Cl}^-@(\text{H}_2\text{O})_{200}$ from SAOP (SR-ZORA) calculations over 100 snapshots with triple-zeta basis and effect of the water droplet size from $[\text{H}_2\text{O}]_{50}$ to $[\text{H}_2\text{O}]_{200}$	80
4.4	Comparison of SAOP (SR-ZORA) BE, SE, and differential stabilisation energies $\Delta_{SE} = SE_n - SE_{n-1}$ for the clusters $[\text{Cl}(\text{H}_2\text{O})_n]^-$, $n=0,7$ water molecules between the first FF16 [139], the corrected FF18 [139] and experimental results [135]. . .	81
4.5	Comparison of the binding energies of $\text{Cl}^-@(\text{H}_2\text{O})_{50}^{tog}$ over 100 snapshots obtained from MDs performed of the two FFs at the level of SAOP (SR-ZORA).	82
4.6	Changes undergone by the chloride 3p orbital of $\text{Cl}^-@(\text{H}_2\text{O})_{50}$ for different level of computation over FF16 100 snapshots; (a) from Kurahashi et al. [133] and (b) from Winter et al. [131]	83

4.7	FF18 Average (Ave (corrected) subtracting 0.17 eV and adding 0.2 eV and 0.4 eV respectively for $P_{3/2}$ and $P_{1/2}$ EOM values) electron BE (eV) of the hydrated chloride from SAOP (SR-ZORA) calculations over 100 snapshots with triple-zeta basis; SD refers to standard deviation. (a) from Kurahashi et al. [133] and (b) from Winter et al. [131].	83
5.1	(unsigned) NMR parameters for Pt–Tl–CN compound. Chemical shift (δ , in ppm) and spin-spin coupling (nJ , in Hz and n is the number of involved bonds). The chemical shifts were referred (in ppm) toward $TlClO_4$, Na_2PtCl_6 , and (c) water-soluble sodium salt of TMS, for ${}^{205}Tl$, ${}^{195}Pt$, and ${}^{13}C$ NMR spectra, respectively. [148]	86
5.2	Size and SO coupling effects on calculated shieldings (ppm) for 1 snapshot for the chemical models $(CN)_5Pt-Tl(H_2O)_n$ ($n=35,64$).	89
5.3	Size and SO coupling effects on calculated spin-spin couplings (Hz) between perturbing (P) and responding (R) element for 1 snapshot for the chemical models $(CN)_5Pt-Tl(H_2O)_n$ ($n=35,64$).	89
5.4	Effects of different models on J -coupling constants obtained from averaging over 256 snapshots.	94
5.5	Effects of different models on the calculated atomic shieldings (σ) obtained from averaging over 256 snapshots.	95
5.6	Effects of different models on the chemical shifts (δ) obtained from averaging over 256 snapshots compared to experimental results [148].	95

Chapter 1

Theoretical chemistry

The tandem of chemistry, as of physics, is towed by theory and experiment. Moreover, with the advent of high computer resources, the computational chemistry which attempts to model all aspects of real chemistry as closely as possible by using calculations rather than experiment, appears as a third tool that provides remarkable advances to chemical problems. It uses the power of computers combined to approaches based upon classical, quantum, and statistical mechanics and other aspects of molecular physics, chemical physics, and physical chemistry [1], to quantitatively model physical and chemical behaviours. This provides insights, that can be very useful to theory and experiment, participating to their progress. In addition, in many times, experiments are difficult or infeasible (nuclear applications, astrophysics, ...) or hard to interpret (insufficient resolution), thus making computational chemistry indispensable.

As a result, a wide variety of models has been developed to obtain information on the structures, properties and energetics of macroscopic systems. In theoretical chemistry and physics we can distinguish between quantum chemical and molecular mechanics models. Classical molecular models which are based, as the name implies, on classical mechanics and, use the concept of force fields [2], where the building blocks are atoms. The electronic energy is written as a parametric function of the nuclear coordinates, in which the parameters are fitted to experimental or higher level computational data. Within this model, structural and dynamical properties can be obtained easily, even for several thousands of atoms, which makes it, the only realistic approach for performing simulations where solvent effects or crystal packing can be studied.

On the other hand, and in order to compute electric and magnetic properties, methods that are based on the solution of Schrödinger or even the Dirac equation, dedicated to deal with relativistic effects, and that treat molecules as collections of nuclei and electrons without any reference to chemical bonds are required. However these equations cannot be solved analytically for systems with more than two particles, and we then need to make approximations. Therefore, the adopted approximations determine the accuracy of the quantum chemical method, the less severe the approximation, the closer will be its results to experiment. This comparison to experiment draws special attention to computed physical observables, but doesn't overlook the computation of other quantities that cannot be observed.

1.1 Schrödinger equation and Born-Oppenheimer approximation

Electrons are very small particles, as such, their behaviour cannot be properly treated by classical mechanics, but rather requires a quantum mechanical description. The many-body time-dependent Schrödinger equation for a system composed of N nuclei with positions $\{\vec{R}_j\}$ and n electrons with positions $\{\vec{r}_i\}$ is given as:

$$\hat{H}\psi_{\text{MB}} = \hat{T}\psi_{\text{MB}} + \hat{V}\psi_{\text{MB}} = i\hbar\frac{\partial\psi_{\text{MB}}}{\partial t}, \quad (1.1)$$

in which \hat{T} is the sum of the electron kinetic energy operator,

$$\hat{T}_e = -\frac{\hbar^2}{2} \sum_{i=1}^n \frac{\nabla_i^2}{m_e}, \quad (1.2)$$

where m_e is the electron mass, and the nuclear kinetic energy:

$$\hat{T}_{\text{nuc}} = -\frac{\hbar^2}{2} \sum_{j=1}^N \frac{\nabla_j^2}{M_j}, \quad (1.3)$$

where M_j is the mass of nucleus j and \hat{V} is the potential energy operator which sums up the electron-nuclei Coulomb potentials,

$$\hat{V}_{e-\text{nuc}} = -\sum_{j=1}^N Z_j \sum_{i=1}^n \frac{e^2}{|r_i - R_j|}, \quad (1.4)$$

and the electron-electron Coulomb repulsions:

$$\hat{V}_{e-e} = \sum_{i=1}^n \sum_{j<i}^n \frac{e^2}{|r_i - r_j|}, \quad (1.5)$$

and the nuclear-nuclear Coulomb repulsions:

$$\hat{V}_{\text{nuc}-\text{nuc}} = \sum_{i=1}^N \sum_{j<i}^N \frac{Z_i Z_j}{|R_i - R_j|}. \quad (1.6)$$

This Hamiltonian may contain additional terms, for example, in the case of presence of external electric or magnetic fields. For a closed system, the conservation of energy makes a separation of time and spatial coordinates possible. The time-independent Schrödinger equation (1.1) can be formulated as an eigenvalue problem:

$$\hat{H}(\vec{r}, \vec{R})\psi(\vec{r}, \vec{R}) = E\psi(\vec{r}, \vec{R}). \quad (1.7)$$

Moreover, the large difference in mass and velocity between electrons and heavy nuclei opens for a further approximation, so that their motions can be decoupled. Thus, the total wave function $\psi(\vec{r}, \vec{R})$ can be written as a product of separate wave functions that correspond to the nuclear and the electronic parts :

$$\psi(\vec{r}, \vec{R}) = \psi_e(\vec{r}, \vec{R})\psi_{\text{nuc}}(\vec{R}), \quad (1.8)$$

for which the Hamiltonian can be written as a sum of the two parts:

$$\hat{H} = \hat{H}_e + \hat{H}_{nuc}. \quad (1.9)$$

That enables us to reduce significantly the complexity of the system by focusing only on the electronic part by solving Eq.(1.10):

$$\hat{H}_e \psi_e(\vec{r}, \vec{R}) = \left\{ -\frac{\hbar^2}{2} \sum_{i=1}^n \frac{\nabla_i^2}{m_e} - \sum_{j=1}^N Z_j \sum_{i=1}^n \frac{e^2}{|r_i - R_j|} + \sum_{i<k=1}^n \frac{e^2}{|r_i - r_k|} \right\} \psi_e(\vec{r}, \vec{R}) = E_e \psi_e(\vec{r}, \vec{R}). \quad (1.10)$$

Therefore the electronic wave function $\psi_e(\vec{r}, \vec{R})$ and the corresponding energy eigenvalues, contain nuclei coordinates $\{\vec{R}_i\}$ as parameters.

For most purposes, we are only actually interested in the ground state $\psi_e^0(\vec{r}, \vec{R})$ of the electronic system, that corresponds to the electron ground state energy E^0 . The total energy E_{tot}^0 is then recovered by adding the nuclear-nuclear Coulomb repulsions $V_{nuc-nuc}$,

$$E_{tot}^0 = E^0 + V_{nuc-nuc}. \quad (1.11)$$

Methods involving solutions of Eq.(1.10) are known as electronic structure methods. There are two major groups of electronic structure methods; those based on the wave function, which is the mathematical representation of the quantum state of a given quantum system, and the density functional theory (DFT) based methods, that aim to directly determine the density, which is associated to the probability of finding a particle at a given point in space.

1.2 Wave function based methods

Even with the approximations that are made, solving the electronic Schrödinger equation is still challenging due to the $\sum_{i<k=1}^n \frac{e^2}{r_{i,k}}$ term in Eq.(1.10) which complexifies things as soon as more than two electrons (without accounting for the nuclei) are involved.

1.2.1 Hartree Fock approximation

In order to deal with the difficulty of describing electron-electron interactions, the Hartree-Fock approximation has been introduced. In it, the true electron-electron interaction is replaced by a model interaction in which each electron interacts with the mean field of all the other electrons, called the self-consistent field.

In this representation an electron i feels other electrons via the potential (ν^{HF}) [3] that will be defined in the following:

$$\nu_i^{HF} = \sum_b J_b(i) - K_b(i), \quad (1.12)$$

where J and K are defined below. Figure (1.1) illustrates a schematic representation of the Hartree-Fock approach.

The Hartree-Fock method is based on the representation of the many-body wavefunction on the basis of a product of one electron functions $\phi_i(r_i)$. Moreover, in order to respect the

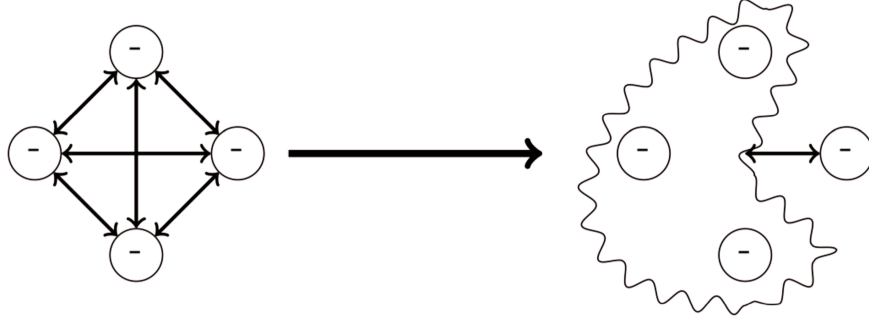


Figure 1.1: Schematic representation of the mean-field approach, - refers to a negative charge (an electron).

fermion nature of the electron, the wave function ψ must to be antisymmetric upon interchange of electron coordinates, giving birth to the Slater determinant concept,

$$\psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_n(2) \\ \dots & \dots & \dots & \dots \\ \phi_1(n) & \phi_2(n) & \dots & \phi_n(n) \end{vmatrix}.$$

Here, ϕ_i is a spin orbital and it equals the product of a spatial function or molecular orbital χ_i , and a spin function α or β leading to enabling that two electrons may occupy the same spatial molecular orbital.

The self-consistent-field (SCF) procedure to minimise the energy of the slater determinant leads a set of n individual electron equations,

$$\hat{F}_i \phi_i(r_i) = \varepsilon_i \phi_i(r_i), \quad (1.13)$$

with:

$$\hat{F}_i(r) = \hat{h}_i(r) + \sum_{j \neq i}^n |\hat{J}_j(r) - \hat{K}_j(r)|, \quad (1.14)$$

is the Fock operator corresponding to the energy of the orbital i :

$$\varepsilon_i = \langle \phi_i | \hat{F}_i | \phi_i \rangle = \langle \phi_i | \hat{h}_i | \phi_i \rangle + \langle \phi_i | \sum_{j \neq i}^n \hat{J}_j(r) - \hat{K}_j(r) | \phi_i \rangle \quad (1.15)$$

where:

$$\hat{h}_i = -\frac{\hbar^2}{2 m_e} \nabla_i^2 - \sum_{k=1}^N \frac{Z_k}{|r_i - R_k|}, \quad (1.16)$$

and:

$$\hat{J}_j | \phi_i(r_i) \rangle = \left\{ \langle \phi_j(r_j) | \frac{1}{|r_i - r_j|} | \phi_j(r_j) \rangle \right\} | \phi_i(r_i) \rangle, \quad (1.17)$$

and:

$$\hat{K}_j |\phi_i(r_i)\rangle = \left\{ \langle \phi_j(r_j) | \frac{1}{|r_i - r_j|} | \phi_i(r_j) \rangle \right\} |\phi_j(r_i)\rangle, \quad (1.18)$$

which represent the Coulomb and exchange operators applied on the orbital $\phi_i(r_i)$. Moreover, it is noticeable that the summation of orbitals energy does not equal the HF total energy because its double counts the Coulomb interaction. The total energy in the HF model is computed as:

$$E_{HF} = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (J_{ij} - K_{ij}). \quad (1.19)$$

Koopmans' theorem and the meaning of orbital energies

In the Slater determinant, only occupied orbitals appear. Each orbital has an energy ε_i . But there is also another solution of the Hermitian Hamiltonian in the complete basis set called virtual or unoccupied orbitals. A good thing about HF potential is that the orbital energies ε_i give approximate ionisation potentials and electron affinities via the Koopmans' theorem. Let us consider the energy of an N-electron system associated with the wave function ψ^0 :

$$\psi_0 = |\phi_1 \phi_2 \phi_3 \dots \phi_n|, \quad (1.20)$$

and the corresponding system with one electron removed from orbital number n associated to the wavefunction ψ^+ :

$$\psi^+ = |\phi_1 \phi_2 \phi_3 \dots \phi_{n-1}|. \quad (1.21)$$

If we assume that the molecular orbitals (MO) are identical for the two systems i.e. there is no relaxation effect on MOs, the subtraction the two corresponding energies yields:

$$\Delta E^I = E_n - E_{n-1} = \langle \psi^0 | H | \psi^0 \rangle - \langle \psi^+ | H | \psi^+ \rangle = \varepsilon_n. \quad (1.22)$$

Thus, the ionisation energy is given as the orbital energy ε_n , a result known as Koopmans' theorem [4]. In a similar way, the electron affinity can be computed as the energy of the $n + 1$ unoccupied orbital of the neutral molecule:

$$\Delta E^{EA} = E_{n+1} - E_n = \langle \psi^- | H | \psi^- \rangle - \langle \psi^0 | H | \psi^0 \rangle = \varepsilon_{n+1}, \quad (1.23)$$

where ψ^- is the wave function of the corresponding anion. However, the lowest unoccupied eigenvalue usually converges to zero, corresponding to a solution for a free electron, described by a linear combination of the most diffuse basis functions. This makes taking unoccupied orbital energies as electron affinities questionable. In contrast approaching ionisation energies as occupied orbital energies is justified [2].

Moreover, the difference $\varepsilon_{n+1} - \varepsilon_n$ is not to be used to approximate an excitation energy from the n to the $n + 1$ orbital, since that the electron in $n + 1$ orbital has to feel $n - 1$ other electrons, whereas it feels n electrons in this case.

1.2.2 Electron correlation

The Hartree-Fock approximation assumes that each electron moves in the average field created by all of the other electrons. The resulting single-determinantal wavefunction corresponds to the resulting lowest possible energy [5]. But this representation engenders an error that

results in a difference between the HF energy and the exact energy $\Delta E = E - E_{HF}$ called the correlation energy, which is mainly due to ignoring the correlated motion of each electron with every other ones. Thus, we need to go beyond the mean-field description of HF by considering many more Slater determinants up to the full interaction.

When speaking of electron correlation, in general one can distinguish between two kinds. The first deals with the instantaneous aspect of electron-electron interaction (dynamical correlation), which can be divided, in turn, into two types. The Fermi correlation, already taken into account in the HF model, which arises from the Pauli antisymmetry principle of the wave function, and the Coulomb correlation, that is not included into the HF approximation, which refers to the fact that the probability of finding two electrons at the same point in space is zero as the repulsion becomes infinite.

Second, we have what is referred to as static or non-dynamic correlation, which arises when we attempt to represent the many-body wavefunction by just one determinant while, in fact, two determinants or more are really needed. Such cases are found for nearly degenerate electron configurations, as in excited states or bond dissociation processes.

In order to include electron correlation back into the calculation there is a variety of methods so-called post-Hartree-Fock methods that could be employed to improve the energetics and the wave functions. Post-Hartree-Fock methods can be distinguished according to their ability to include dynamical or static correlation. The Configuration Interaction (CI) model, which has been successfully applied in quantum chemistry due to its formal and conceptual simplicity, can incorporate both of them.

Configuration Interaction

The most common way to improve the wave function beyond the single determinant:

$$\psi_{HF} = |\phi_1\phi_2\phi_3\dots\phi_n|, \quad (1.24)$$

is to use trial wave functions of the CI form

$$\Psi_{CI} = c_0\psi_{HF} + c_1\psi_1 + c_2\psi_2 + \dots \quad (1.25)$$

where the coefficients c_i reflect the weights of each determinant ψ_i in the expansion and also for ensuring normalisation.

In a general expansion, it is not necessary to use the HF wave function in the CI wave function, but it represents a good starting point for "single-reference" systems, that means, systems for which only a single determinant is enough to describe correctly the corresponding state, and therefore only the incorporation of dynamical correlation effects is sought. In these cases, the "single reference" CI wave function is generated from the HF reference by the application of a linear combination of spin-orbital excitation operators.

$$\Psi_{CI} = \left(c_0 + \sum_{i,m} c_i^m \hat{X}_i^m + \sum_{i,j,m,n} c_{i,j}^{m,n} \hat{X}_{i,j}^{m,n} + \dots \right) \psi_{HF}, \quad (1.26)$$

where,

$$\begin{aligned} \hat{X}_i^m \psi_{HF} &= a_m^\dagger a_i \psi_{HF} = \psi_i^m, \\ \hat{X}_{i,j}^{m,n} \psi_{HF} &= a_m^\dagger a_n^\dagger a_i a_j \psi_{HF} = \psi_{i,j}^{m,n}. \end{aligned} \quad (1.27)$$

are the single and double excitation operators. These operators denote respectively, the creation of an electron in the spin-orbital ϕ_m and the removal of an electron from the spin-orbital ϕ_i to generate the single excitations, and the double excitations from ϕ_i and ϕ_j to ϕ_m and ϕ_n spin-orbitals. Higher excitations relative to the Hartree-Fock state can also be included until the so-called Full Configuration Interaction (FCI) is reached. Figure (1.2a) shows an example of the ground state (in red) and one of their singly and doubly excited states (in green) for a twelve-electron system for which the excitations is done only from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO).

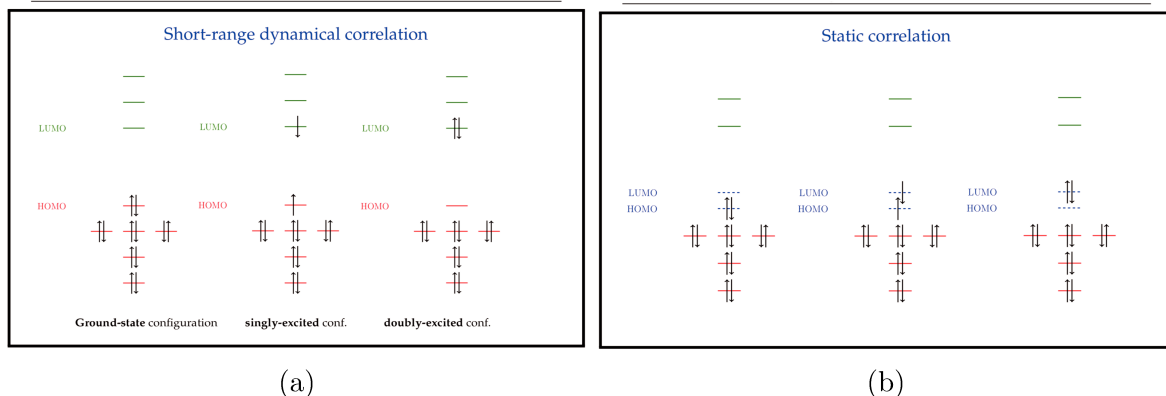


Figure 1.2: Examples of dynamical and static correlations inclusion for a twelve-electron system. (Taken from E. Fromager talk in the 2017 summer school [6]).

The FCI method yields, in a given one electron basis, the exact solution to the Schrödinger equation in the Hilbert space. It is constructed, in addition to the reference wave function, from all the possible excited determinants which may be characterised, in addition to single (S) and double (D) excitations, as triple (T), quadruple (Q), quintuple (5), sextuple (6) and higher possible excitations. Therefore, a drawback of the FCI is its factorial dependence of the number of Slater determinant on the number of electrons and total spin-orbitals as it is clear in Table (1.1), in which the number of many determinants can be formed when one distributes n electrons among $2n$ orbitals ($4n$ spin-orbitals).

$2n$	N_{det}
2	4
4	36
6	400
8	4900
10	63504
12	853776
14	11778624
16	165636900
18	2363904400
20	34134779536

Table 1.1: Number of Slater determinant that can possibly be formed for a system of $2n$ electrons in $2n$ orbitals.

The number of Slater determinants in the FCI expansion is given by:

$$N_{det} = \left(\frac{M!}{N!(M-N)!} \right)^2, \quad (1.28)$$

and quickly makes calculations impossible or impractical. This is one reason why we use in general a truncated CI, in addition to the fact that, in most cases, the lower-order excitations are more important than those of higher orders [7] that can be evaluated with perturbation theory (PT). In Table (1.2) we report the energies of truncated CI wave functions relative to the FCI energy for the water molecule and the weights of CI functions in the FCI one. It is demonstrated that the CISD wave function is enough to recover more than 98 % of the FCI energy with a weight (means here the overlap between the used and the exact wavefunctions $\langle \Psi^{used} | \Psi^{exact} \rangle$) larger than 99% .

method	$E - E_{FCI}$	weight
RHF	0.217822	0.941050
CISD	0.012024	0.998047
CISDT	0.009043	0.998548
CISDTQ	0.000327	0.999964
CISDTQ5	0.000139	0.999985
CISDTQ6	0.000003	1

Table 1.2: H₂O energies for different truncated CI wave functions and their weights in the FCI one. (for more details see Table (5.9) in [7].)

The FCI which is a variational method, meaning that the energy obtained by minimisation of the expectation value of the Hamiltonian represents an upper bound to the exact ground state energy, is very useful when it is applicable. It is used to compare the performance of other methods for which the aim is to get around the FCI inapplicability using some approximations as truncated CI expansions.

The FCI has also the possibility via its wave function construction (Eq.(1.25)) to recover static correlation effects and overcome the limitations of the "single-reference" CI method. The "multi-reference" CI (MRCI) is based on the idea of a reference space that contains all the necessary determinants needed to describe a given physical process (Figure (1.2b) shows a case when three determinant are needed in the reference space), in addition to all possible corresponding excitations, for every determinant obtained in the same way as for a single reference CI.

Coupled-Cluster

In the Coupled Cluster (CC) theory, the way that the wave function is expressed is intrinsically different. This theory uses a single Slater determinant as starting point. So if the studied problem is not well described by a single determinant, one probably should not use coupled cluster theory. The CC wave function can be expressed as:

$$\Psi_{cc} = \exp(\hat{T})\psi_0. \quad (1.29)$$

where Ψ_0 is a single determinant used in the SCF process to generate a set of spin-orbitals (typically a HF determinant). The cluster operator \hat{T} sums up several excitation operators:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots = \sum_{i,m} t_i^m \hat{X}_i^m + \sum_{i,j,m,n} t_{i,j}^{m,n} \hat{X}_{i,j}^{m,n} + \dots, \quad (1.30)$$

where the excitation operators (\hat{X}) were defined in the previous section. Moreover one can include more excitations in the \hat{T} operator.

Slater determinants excited more than n times contribute to the wave function because of the non-linear nature of the exponential function. Therefore, a coupled cluster expansion terminated at \hat{T}_n usually recovers more correlation energy than CI truncated at n excitations. The classification of traditional coupled-cluster methods (like for the CI) rests on the highest number of excitations allowed in the definition of the operator \hat{T} . The abbreviations for coupled-cluster methods usually begin with the letters CC (for coupled cluster) followed by S for single excitations, D for double excitations, T for triple excitations or Q for quadruple excitations. Thus the operator in CCSDT has the form:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3. \quad (1.31)$$

Terms with round brackets indicate that these terms are calculated based on perturbation theory. For example, the CCSD(T) approach simply means that it includes singles and doubles fully while triples are calculated with perturbation theory.

Coupled cluster (CC) method, especially the CCSD(T), by its size-extensivity¹, which is not available in truncated CI methods, has become the gold-standard of quantum chemistry. CC theory, which describes very well the dynamical correlation, was poised to describe essentially all the quantities of interest in chemistry, and has now been shown numerically to offer the most predictive, widely applicable results in the field for "single-reference" systems. Although this method is not variational, its computational cost is very high which makes it, in practice, limited to relatively small systems. In fact application range from small molecular systems with less than 20 electron, for which large basis set leads to highly accurate results, to systems like benzene dimers, naphthylene diradicals or nucleic acid bases, and the like.

Equation-of-motion coupled-cluster method (EOM-CC)

CC theory as presented above is a single reference method so it's primarily dedicated to ground-state problems. To obtain other states, one can use approaches such as the equation-of-motion coupled-cluster (EOM-CC) method [8, 9]. This extension which is very similar to the Configuration Interaction (CI) scheme is particularly useful to compute excited, ionised, and electron-attached states [10, 11, 12, 13, 14].

The formalism starts by considering the Schrödinger equation for two different states, the ground state Ψ_{cc} defined by:

$$\hat{H}\Psi_{cc} = E_{cc}\Psi_{cc}, \quad (1.32)$$

and the target state which could be an excited, ionised, or electron-attached state,

$$\hat{H}\Psi_K = E_K\Psi_K, \quad (1.33)$$

where the target state is written as:

¹ size-extensivity: is a more mathematically formal characteristic which refers to the correct (linear) scaling of a method with the number of electrons

$$\Psi_K = \hat{R}_K \Psi_{cc}, \quad (1.34)$$

where \hat{R} (R for right-handed) is an excitation operator,

$$\hat{R}_K = r_0 + \sum_{i,m} r_i^m a_m^\dagger a_i + \sum_{i,j,m,n} r_{i,j}^{m,n} a_m^\dagger a_n^\dagger a_i a_j + \dots, \quad (1.35)$$

for excited states. If we want an ionised state, then the \hat{R} operator takes the form,

$$\hat{R}^I = \sum_i r_i a_i + \sum_{i,j,n} r_{i,j}^n a_n^\dagger a_i a_j + \dots \quad (1.36)$$

In the reverse case, in which an electron is attached, the \hat{R} operator is written as,

$$\hat{R}^A = \sum_m r^m a_m^\dagger + \sum_{j,m,n} r_j^{m,n} a_m^\dagger a_n^\dagger a_j + \dots \quad (1.37)$$

In the EOM-CC formalism, the problem of solving the electronic Schrödinger equation is reformulated in terms of \hat{T} and \hat{R} operators which are both excitation operators from the same reference. The Schrödinger equation is written therefore as:

$$\hat{H}\hat{R}e^{\hat{T}}|\Psi_{HF}\rangle = E\hat{R}e^{\hat{T}}|\Psi_0\rangle. \quad (1.38)$$

Multiplying both sides of this equation by $e^{-\hat{T}}$, and using the commutation property of \hat{R} and \hat{T} , we arrive at the following equation:

$$\bar{H}\hat{R}|\Psi_{HF}\rangle = E\hat{R}|\Psi_0\rangle, \quad (1.39)$$

where $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$.

The EOM-CC approach is a useful electronic-structure tool that allows one to treat a variety of multi configurational problems within a single reference formalism. It enables rather than total energies, the direct calculation of energy differences [15]. It is always energy differences between the states of the system which are observed experimentally in spectroscopy or chemistry. These energy differences are many orders of magnitude smaller than total energies. This fact makes their computation a challenge because small errors in total energies may result in very large errors in energy differences [16]. A good example for the performance of this formalism can be found in the comparison between CCSD, experimental and "multi-reference"(MR) CI excitation values of CH₂Cl. In this, the EOM-CCSD value for the first excitation energy (6.33 eV) is closer to the observed maximum in the absorption spectrum (6.20 eV) than the MR CI value of 6.51 eV [17].

1.3 Density functional based methods

In the previous section, we showed how complicated it is to compute the correlation energy. In this case we know the Hamiltonian and we struggle to obtain a satisfying wave function. It turns out that there is another way that does not take into account configurations except a single one. This is the premise of Density Functional Theory (DFT) based methods that have as a main goal computing calculations of the total energy of the system and the ground state

electron density distribution without using the wave function of the system explicitly.

The DFT is based on two important theorems [18] presented in 1964 by Hohenberg and Kohn who established DFT as a rigorous quantum chemical methodology. The first theorem demonstrates that we can write the energy of the system as a functional of the electron density ρ , as same as we can do for a wave function that gives the same density. The second theorem proves the DFT variational principle analogue. For a given number of electrons (N) and external potential ν_{n-e} (nuclei-electrons attraction potential), the ground-state density ρ_0 that minimises the energy functional as:

$$E_{\nu_{nuclei}}^{HK}[\rho] \geq E_{\nu_{nuclei}}^{HK}[\rho_0] = E_0, \quad (1.40)$$

where E_0 is the minimal value for an unknown functional $E_{\nu_{nuclei}}^{HK}[\rho]$ defined as,

$$E_0 = E_{\nu_{nuclei}}^{HK}[\rho_0] = \min_{\rho} \left\{ \int \nu_{n-e}(r)\rho(r)d^3r + F[\rho(r)] \right\}, \quad (1.41)$$

with,

$$F[\rho(r)] = \min_{\Psi \in \mathcal{E}_A(\rho)} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle, \quad (1.42)$$

is the universal functional containing all the missing terms to the electronic energy, where, \mathcal{E}_A is the ensemble of antisymmetric wavefunctions giving ρ

In order to proceed toward the unknown universal functional $F[\rho(r)]$, the Kohn-Sham (KS) DFT formalism [19] was introduced in 1965.

1.3.1 Kohn-Sham DFT formalism

The universal functional $F[\rho(r)]$ is difficult to approach, notably for the kinetic energy part. To deal with this, Kohn and Sham proposed to consider a fictitious system of non-interacting electrons where the ground state density is the same as the density of the real system. Moreover, the associated Hamiltonian to this fictitious system is a sum of one-electron operators, where the proper eigenfunctions can be Slater determinants of the individual one-electron eigenfunctions whereas the eigenvalues sum up the one-electron eigenvalues. Thus, the non-interacting kinetic energy T_s is written as:

$$T_s[\rho(r)] = \min_{\Psi \in \mathcal{E}_A(\rho)} \langle \Psi | \hat{T} | \Psi \rangle \quad (1.43)$$

The universal functional can, therefore, be expressed as:

$$F[\rho(r)] = T_s[\rho(r)] + E_H[\rho(r)] + E_{xc}[\rho(r)], \quad (1.44)$$

where $E_H[\rho(r)]$ is the Hartree energy functional that is written in the form:

$$E_H[\rho(r)] = \iint \frac{\rho(r')\rho(r)}{|r-r'|} dr dr', \quad (1.45)$$

and which represents the classical electron-electron repulsion that recovers the majority of the exact electron-electron repulsion. The last term,

$$E_{xc}[\rho(r)] = \Delta T + \Delta V_{ee}, \quad (1.46)$$

represents the exchange-correlation energy which contains all the missing terms in the energy, where $\Delta T = T - T_s$ is the correction to the kinetic energy due to the instantaneous interaction of electrons, and ΔV_{ee} contains all non classical corrections to electron-electron repulsion energy.

In most cases, DFT implementation rely on Kohn-Sham theory, and boil down to finding trial densities by determining energies in the same spirit as the SCF procedure for HF. The ground-state energy E_0 , given, still without any approximations as:

$$E_0 = \min_{\Psi \in \mathcal{E}_A(\rho)} \left\{ \langle \Psi | \hat{T} | \Psi \rangle + \int \nu_{n-e}(r) \rho(r) d^3(r) + E_H[\rho(r)] + E_{xc}[\rho(r)] \right\} \quad (1.47)$$

The minimisation is then, performed following the Euler-Lagrange equation (ELE) as:

$$\frac{\delta E_0}{\delta \rho} - \mathcal{E}_{KS} = 0 \quad \Rightarrow \quad (\hat{T} + \hat{V}_{KS}[\rho_{\Psi}]) | \Psi \rangle = \mathcal{E}_{KS} | \Psi \rangle. \quad (1.48)$$

Here, \mathcal{E}_{KS} is a Lagrange multiplier and \hat{V}_{KS} is the Kohn-Sham local external potential, that is written as:

$$\hat{V}_{KS} = \int \{ \nu_{n-e}(r) + \nu_H(r) + \nu_{xc}(r) \} \rho(r) d^3r, \quad (1.49)$$

including the Hartree potential:

$$\nu_H(r) = \left. \frac{\delta E_H[\rho]}{\delta \rho(r)} \right|_{\rho=\rho_{\Psi}}, \quad (1.50)$$

and the exchange-correlation potential:

$$\nu_{xc} = \left. \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \right|_{\rho=\rho_{\Psi}}. \quad (1.51)$$

Coming now to the exchange-and correlation energy functional, that is decomposed into two parts:

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho], \quad (1.52)$$

where, the correlation can be also seen as the sum of two contributions, ΔT defined above, and ΔU_c which recovers the missing part of the exact electron-electron repulsion energy V_{e-e} as:

$$\Delta U_c[\rho] = V_{e-e}[\rho] - (E_x[\rho] + E_H[\rho]) = E_c - \Delta T. \quad (1.53)$$

As a consequence, the ν_{xc} can be split into two parts:

$$\nu_x = \frac{\partial E_x[\rho]}{\delta \rho(r)} \quad ; \quad \nu_c = \frac{\delta E_c[\rho]}{\delta \rho(r)}. \quad (1.54)$$

They represent, respectively, the exchange and the correlation potentials, that are needed, in practice, to be approximated. In the following we present some of the common approaches.

1.3.2 The meaning of orbital energies in KS-DFT

The counterpart in KS DFT of the Koopmans' theorem is the so-called Janak's theorem [20], which states that, the derivative of the total energy with respect to the occupation number n_i of an occupied orbital equals the energy ϵ_i of this orbital,

$$\frac{\partial E}{\partial n_i} = \epsilon_i. \quad (1.55)$$

Applying this theorem to a HOMO orbital for a fractional number $N' = N - \delta$ with $\delta \rightarrow 0^+$ leads to,

$$\left(\frac{\partial E_0^{N'}}{\partial N'} \right)_{N-\delta} = \epsilon_H^{N-\delta} = \epsilon_H^N. \quad (1.56)$$

And by knowing that $N - 1 < N' < N$, this variation of energy is:

$$\left(\frac{\partial E_0^{N'}}{\partial N'} \right)_{N-\delta} = -I_N, \quad (1.57)$$

where I_N is the ionisation energy of the N electron system. In the limit of an exact exchange-correlation potential, this theorem leads to the exact ionisation energy. Moreover, it is noticeable that the KS HOMO-LUMO gap in molecules is very different of the Hartree-Fock one. It represents the lowest excitation energy, rather than difference between the ionisation energy and the electron affinity [21].

1.3.3 The local density approximation

The simplest approach to calculate the exchange-correlation energy is the so-called local density approximation (LDA). This approximation relies on the fact that the local density, within a small volume, can be assumed as homogeneous, and that the contribution in the total exchange-correlation energy could be calculated as the product of the small volume and the exchange-correlation energy density from the homogeneous gas theory that is calculated inside the small volume [19, 22]. The exchange-correlation energy in this approximation for a closed-shell system may be written therefore as:

$$E_{xc} = \int d^3r \{ \epsilon_x[\rho(r)] + \epsilon_c[\rho(r)] \}, \quad (1.58)$$

with ϵ_x is the local exchange potential of homogeneous gas and ϵ_c is the local correlation potential fit to Monte Carlo calculations [23, 24] for homogeneous gas. Thus everything is decided locally with only density dependency. The LSDA enjoyed early success in physics in giving bond lengths and thus geometries of molecules and solids typically with an remarkable accuracy of $\sim 1\%$ [25]. However, there are systematic errors in computing molecular atomisation energies, where the tendency of the LDA is to over bind by 20-30% [26]. Bulk modulus and vibrational frequencies come out with about 19% error [26].

1.3.4 The generalised gradient approximation (GGA)

LDA, as it approximates the energy of the true density by the energy of a local constant density, fails in situations where the density undergoes rapid changes, such as in molecules and

solids. An improvement can be made by also considering the gradient of the electron density, the so-called generalised gradient approximation (GGA):

$$E^{GGA} = E^{LSDA} (1 + \mu s^2 + \text{higher-order}), \quad (1.59)$$

where μ is a sort of theoretical correction term [27] and:

$$s = \text{constant} * \frac{|\nabla\rho|}{\rho^{4/3}}, \quad (1.60)$$

is the reduced or dimensionless density gradient. Most GGA functionals that depend on both the density and the gradient of the density are constructed with the correction being a term added to the LDA functional:

$$\varepsilon_{x/c}^{GGA}[\rho(r)] = \varepsilon_{x/c}^{LSDA}[\rho(r)] + \Delta\varepsilon_{x/c} \frac{|\nabla\rho(r)|}{\rho^{4/3}(r)}. \quad (1.61)$$

For which x/c means the same functional for either exchange or correlation and the correction term depends on the dimensionless reduced gradient.

The most popular GGA exchange functional to date (44,000+ citations) is the Becke 1988 (B) [28] which corrects the asymptotic behaviour at long range for the energy density and incorporates a single empirical parameter for which the value is optimised by fitting to the exactly known exchange energies of six noble gas atoms (from He to Rn). We can also use other exchange functionals similar to Becke such as PW [29], FT98 and mPW [30]. Alternative GGA exchange functionals have been developed based on a rational function expansion of the reduced gradient. They contain no empirically optimised parameters like B86 [31], or PBE [32]. On the other side, correlation energy was also subject of many development either to add corrections to LDA functionals such as P86 [33] and PW91 [34]. Other functionals do not correct the LDA expression but compute the full correlation energy like the LYP (Lee, Yang, Parr 1988) correlation [35] which contains four empirical parameters adjusted to helium atom. It is a correlation functional that provides an exact cancellation of the self-interaction error² in one-electron systems. In any case, all of these different functionals are developed from different viewpoints with the idea to satisfying different constraints. Benchmark tests will have the last word before deciding which one is most effective in practice.

The next logical step in this Taylor-function-expansion correction is to include second derivative of the density under the name of Meta-GGA. This idea is concretised by Becke and Rousell on the exchange functional (BR) and by Proynov, Salahub, and co-workers for the correlation functional (Lap). In general the Meta-GGA are more accurate than GGA with a cost comparable to that for GGA [5].

1.3.5 Hybrid functionals

The main idea of hybrid functionals [37] introduced in 1993 by Becke is to mix GGA functionals with Hartree-Fock exchange for the reason that the HF exchange cancels self-interaction

² The self-interaction error is related to the spurious interaction of an electron with itself. In Hartree-Fock, self-interaction is explicitly and exactly cancelled, which is why it appears to work. With the LDA, self-interaction is not cancelled [36]

in V_{ee} in the Eq.(1.44). The first proposed functional considering this mixing is the so-called HH [38] (for half and half) with:

$$E_x = \frac{1}{2}(E_x^{HF} + E_x^{DFT}). \quad (1.62)$$

However, it turned out that adding more parameters in the hybrid functional formula can bring more accuracy. In the same year Becke developed a 3-parameter hybrid functional (B3) [39] using GGA exchange, correlation functionals B and PW91 as:

$$E_{xc}^{hyb} = E_{xc}^{LSDA} + a_0(E_x^{HF} - E_x^{LSDA}) + a_x(E_x^{GGA} - E_x^{LSDA}) + a_c(E_c^{GGA} - E_c^{LSDA}). \quad (1.63)$$

The parameters $a_c = 0.20$, $a_x = 0.72$ and $a_0 = 0.81$ were determined by fitting to a data set of measured atomisation energies. This model was modified to use LYP instead of PW91 [40] by Stephens et al. leading to the most popular functional [41] to date B3LYP as shown in the Figure (1.3).

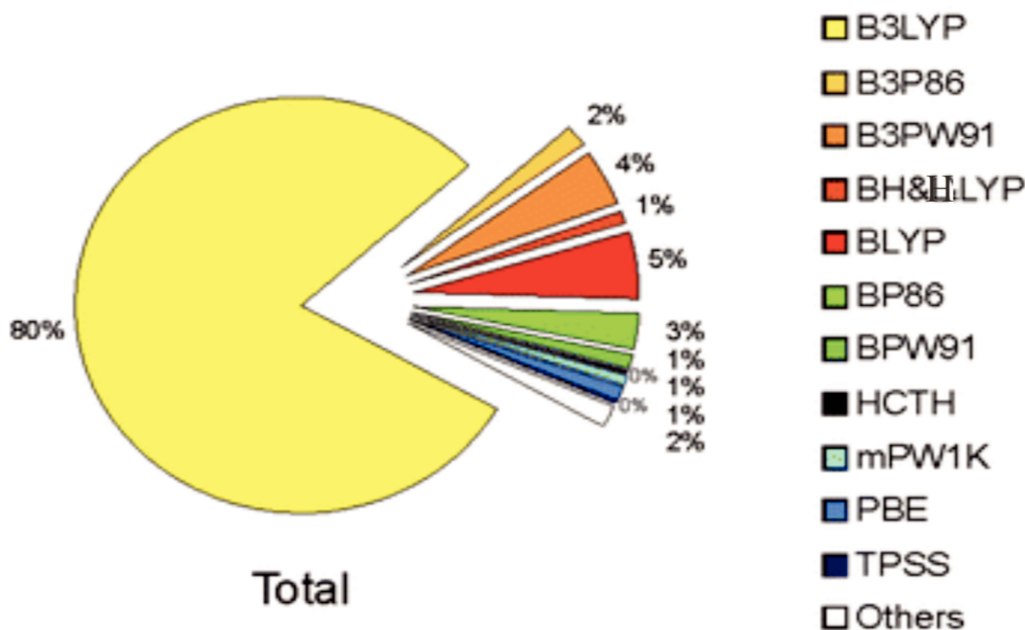


Figure 1.3: DFT approaches popularity

1.3.6 The statistical average of orbital potentials

It was shown in the subsection (1.3.2) that, in the limit of exact exchange correlation potentials, the HOMO and LUMO energies correspond to the first vertical ionisation energy and electron affinity respectively. However, an accurate Kohn-Sham potential exhibits appealing properties of the KS orbitals, including virtual-occupied orbital energy gaps and the realistic shapes of virtual orbitals, leading to straightforward interpretation of most excitations as single

orbital transitions [21]. Figure (1.4) shows the upshift of orbitals due to the use of LDA and GGA potentials.

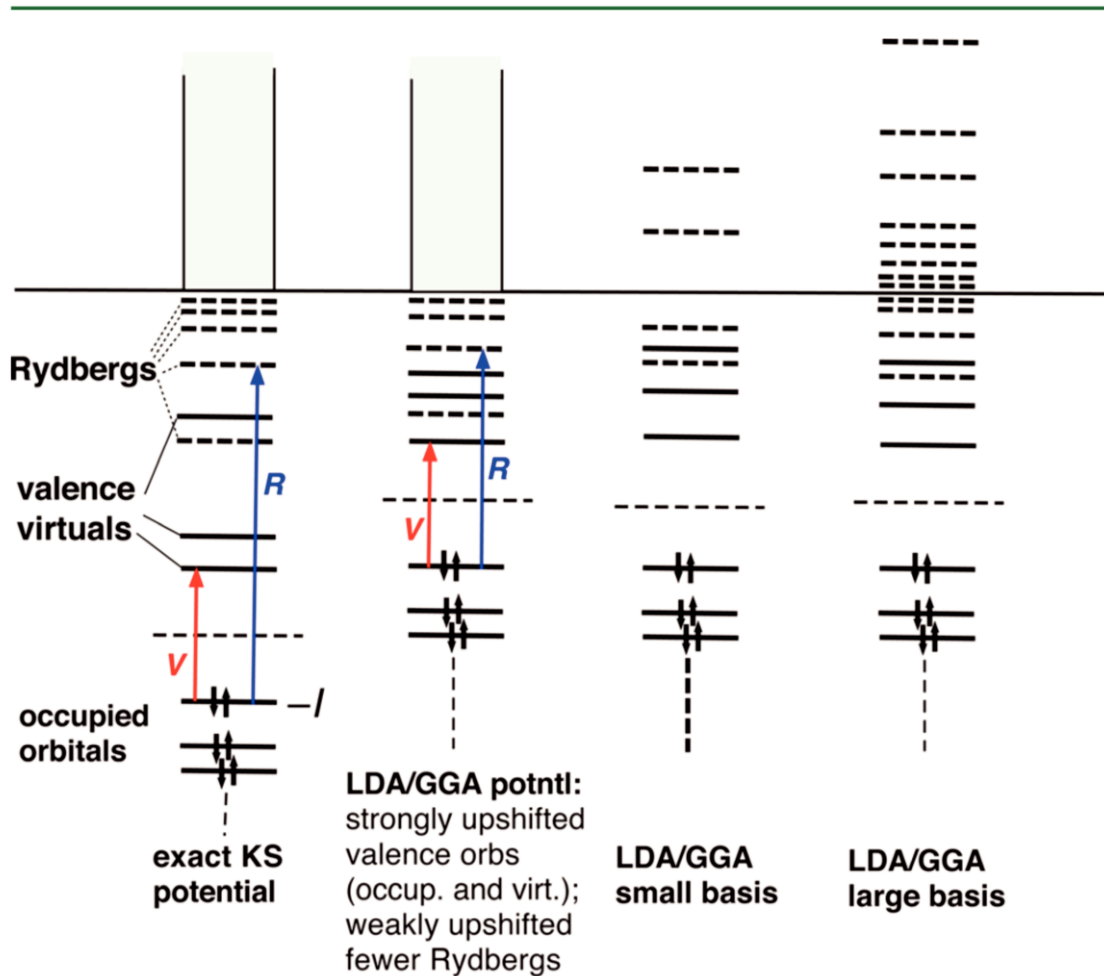


Figure 1.4: Schematic representation of the orbital energy spectrum of the exact Kohn-Sham potential, a typical GGA potential, and the approximate discrete LDA/GGA orbital energy spectra resulting from small and large basis set calculations. Drawn lines for valence orbitals, dashed lines for Rydberg levels. V denotes a valence excitation (HOMO to LUMO transition) and R denotes a Rydberg transition [21].

However, an adequate approximation of the Kohn-Sham exchange-correlation potential remains a challenging problem in DFT. In this context, the statistical average of orbital potentials (SAOP) is proposed as a way to model the Kohn-Sham potential [42]. It combines a model potentials ν_{xc}^{Ei} , based on the exponential integral with a proper behaviour in the outer region, for the highest occupied KS orbital ϕ_N as:

$$\nu_{xc}^{Ei}(r) = \nu_{xc}^{LDA}(r) - \frac{\rho^{1/3}(r)\beta x^2(r)}{1 + 3\beta x(r)E_1[1/x(r)]} \quad (1.64)$$

where $\beta = 0.02$, x is the dimensionless gradient argument and $E_1(z) = \int_z^\infty \frac{\exp(-t)dt}{t}$, with another model potential ν_{xc}^{GLB} [43, 44], that has realistic features in the inner region, for other occupied orbitals leading to:

$$\nu_{xc}^{SAOP}(r) = \nu_{xc}^{GLB}(r) \sum_{i=1}^{N-1} \frac{|\phi_i(r)|^2}{\rho(r)} + \nu_{xc}^{Ei}(r) \frac{|\phi_N(r)|^2}{\rho(r)}. \quad (1.65)$$

With this interpolation between ν_{xc}^{Ei} and ν_{xc}^{GLB} , SAOP allows reproducing good orbital energies [42, 21] and therefore obtaining ionisation energy more accurate than the Hartree-Fock as shown in Table (1.3) [45, 46, 47, 48].

Molecules	I_{exp}	I_{HF}	I_{MP2}	I_{SAOP}
HF	16.00	17.41	17.36	16.30
N ₂	5.58	17.00	16.41	15.50
F ₂	15.34	17.73	18.20	14.98
CO	14.00	14.97	15.08	13.20
CH ₄	13.6	14.91	14.86	14.10
HCl	12.75	13.01	13.01	12.58
H ₂ O	12.62	13.61	13.59	13.02
SO ₂	12.3	13.39	13.46	12.32
Cl ₂	11.49	12.30	12.37	11.59
CCl ₄	11.5	12.68	12.69	11.71
CH ₃ Cl	11.2	11.91	11.90	11.27
CH ₂ O	10.9	11.88	11.98	10.51
CH ₃ CHO	10.2	11.41	11.56	9.91

Table 1.3: Ionisation potentials (in eV) calculated from different ab initio techniques [45] and SAOP methods using Koopmans' and Janack's theorems, and comparison with the experimental ionisation (I_{exp}).

1.3.7 The DFT compared to the WFT

The DFT method is qualitatively different from wave function based method because it optimises the electron density rather than the wave function. Therefore, molecular properties are expressed in terms of electron density, making WF based methods of a broader utility since there are more well characterised operators (to obtain properties) than there are generic property functionals of the density. Moreover DFT lacks the systematic improvability, in contrast of WFT that defines well path to the exact solution (Full CI with infinite basis).

However, by its formulation, and despite of taking correlation energy into account, DFT based approaches are not expensive. Its cost is comparable to Hartree-Fock (see Table 1.4) method and therefore the same computer power enables exploring much larger systems than with post Hartree-Fock methods.

1.4 The molecule in a magnetic field

Molecular properties with or without an external electric or magnetic field change and a challenge is to determine these changes from isolated molecule properties and with respect to the applied field. Particularly molecules respond to the application of a magnetic field (Figure (1.5)) not by changes in the orientations of the nuclear magnetic moments, because

Scaling behaviour	Method(s)
N^4	HF, DFT
N^5	MP2
N^6	MP3, CISD, MP4SDQ, CCSD, QCISD
N^7	MP4, CCSD(T), QCISD(T)
N^8	MP5, CISDT, CCSDT
N^9	MP6
N^{10}	MP7, CISDTQ, CCSDTQ

Table 1.4: QM scaling behaviour with N (number of electrons) [5].

the perturbation is too small, but with modifications in the internal electric currents and local magnetic fields.

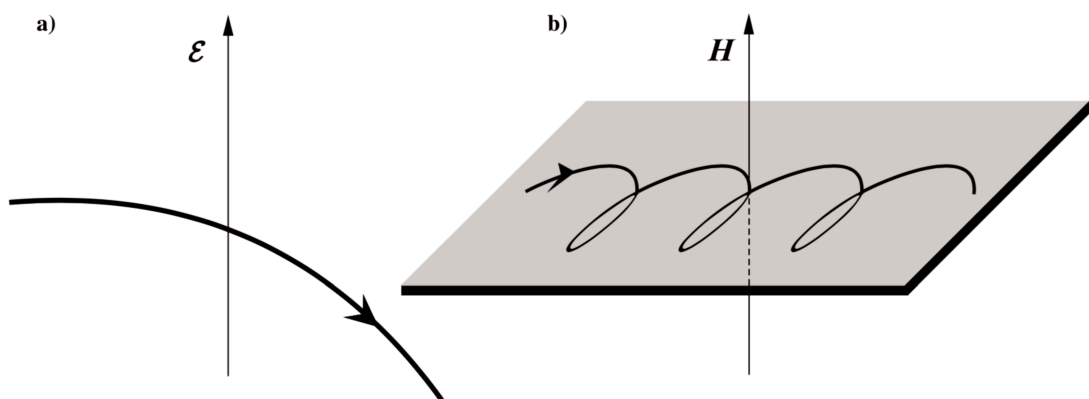


Figure 1.5: The trajectories of an electron in the (a) electric field - the trajectory is a parabola (b) magnetic field, perpendicular to the incident velocity the trajectory is a cycloid in a plane perpendicular to the figure

This dependence on local molecular structure modulates the interaction between nuclear spins (j-coupling) or between the nuclear spin and an external magnetic field (shielding). These interactions lead to transitions between energy levels that can be recorded in a Nuclear Magnetic Resonance (NMR) spectrum [49, 50] making NMR spectroscopy of great utility as it allows one to probe the local electronic structure without damaging it.

1.4.1 The chemical shift (CA)

In the presence of an external magnetic \vec{B}_0 , charged particles will move in loops creating an induced magnetic field \vec{B}_{ind} (Lenz' law) in the opposite direction of the external one. The locally experienced magnetic field (projected on the z axis) is written therefore as:

$$\vec{B}_{local} = \vec{B}_0 - \vec{B}_{ind} = (1 - \sigma)\vec{B}_0. \quad (1.66)$$

The change in the magnetic field experienced by the nucleus, will cause slight changes in the energy levels (Figure 1.6) and this is defined as the shielding (σ), which is linked to the so-called chemical shift if subtracted to a reference shielding value [51].

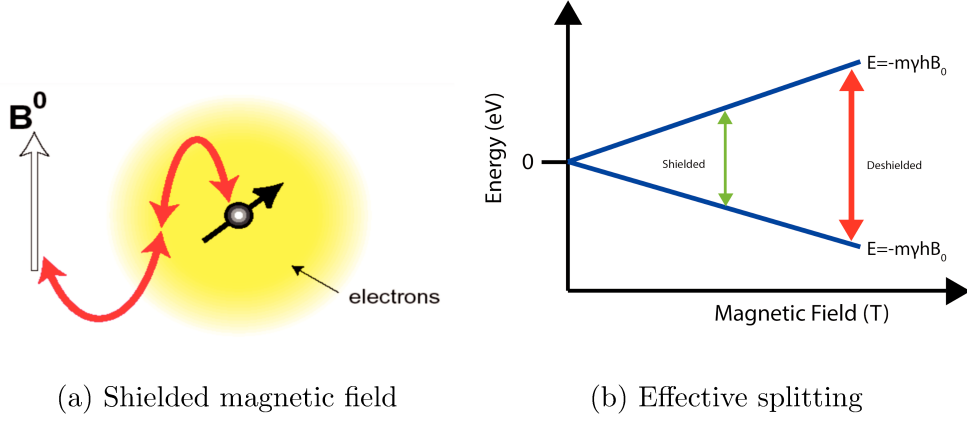


Figure 1.6: The local effective magnetic field felt by a given nucleus is different from the external one due to the surrounding electrons.

The Hamiltonian that describes the shielding effects is:

$$\hat{H}_{CS} = -\gamma\hbar\vec{I}\cdot\vec{B}_{local} = -\gamma\hbar\vec{I}\cdot\vec{B}_0(1 - \sigma), \quad (1.67)$$

where \vec{I} is the nuclear magnetic moment, $\gamma = \frac{\mu}{I}$ is the gyromagnetic ratio, a proportionality constant between the magnetic dipole moment μ and I , and σ is a shielding tensor which can be decomposed into a symmetric and asymmetric parts $\sigma = \sigma_s + \sigma_{as}$. The asymmetric part is associated the relaxation phenomenon and it is found that it does not participate, (in liquids because of the rotational motion of the nuclei) or weakly to the chemical shifts [52, 53], and therefore the shielding tensor can be reduced to its symmetric contribution. The later can be diagonalised then as:

$$\sigma = \begin{pmatrix} \sigma_{1,1} & 0 & 0 \\ 0 & \sigma_{2,2} & 0 \\ 0 & 0 & \sigma_{3,3} \end{pmatrix}. \quad (1.68)$$

For which the isotropic part is:

$$\sigma_{iso} = \frac{1}{3}(\sigma_{1,1} + \sigma_{2,2} + \sigma_{3,3}). \quad (1.69)$$

The isotropic (independent of molecular orientation with respect to the magnetic field) chemical shift (δ), which is controlled by ΔE (spin-flip energy), is determined by the magnitude of the local magnetic field at the nucleus; it is the difference between the frequency of the nucleus spin-flip of the concerned nucleus and the nucleus spin-flip of a reference molecule divided by this last frequency. In other words, it is the resonant frequency of a nucleus compared to a reference. This definition leads to:

$$\delta_{iso} = \frac{\nu_{iso} - \nu_{ref}}{\nu_{ref}} 10^6 = \frac{\sigma_{ref} - \sigma_{iso}}{1 - \sigma_{ref}} 10^6 = (\sigma_{ref} - \sigma_{iso}) ppm, \quad (1.70)$$

with:

$$\nu = \frac{\Delta E}{h} = \frac{\gamma B}{2\pi} = \frac{\gamma B_0(1 - \sigma)}{2\pi}. \quad (1.71)$$

Many factors influence the chemical shift. In general the higher the electron density, the larger the shielding and therefore the lower the observed chemical shift. The electronegativity difference between concerned atom and its surrounding atoms also impacts the chemical shift.

Indeed, if the concerned atom is less electronegative than the surrounding ones, the electron density around it is lower, thus inducing a higher chemical shift [54, 55].

1.4.2 Spin-spin coupling

The spin-spin interaction splits into two types, the direct dipole-dipole coupling which represents the direct magnetic interaction between nuclear spins and the interaction between nuclear spins through the involvement of bonding electron orbitals which is called the indirect dipole-dipole, scalar, or J coupling. The Hamiltonian describing the spin-spin interaction between nucleus A and nucleus B is written in the form:

$$\hat{H}_{SS}^{AB} = \frac{1}{2} \hbar \gamma_A \vec{I}_A^T \cdot (D_{AB} + J_{AB}) \cdot \hbar \gamma_B \vec{I}_B, \quad (1.72)$$

where D_{AB} is the 3x3 classical dipolar coupling tensor describing the direct dipole-dipole interaction [56] through space and J_{AB} is also 3x3 tensor which represents the indirect dipole-dipole (J) coupling.

The direct coupling is non-zero in solids, but it vanishes in isotropic liquids because of the motional averaging where molecules tumble randomly in all orientations. On the other hand, the J-coupling resists to molecular orientation due to its dependency to bonds. It thus makes it as an excellent indicator of the presence of bonding interactions.

The J-coupling, which concerns us in this thesis, is independent of the external field B_0 . It results from the interaction between two nuclear moments mediated by the electron density, so it is important for regions with bonds through it. It can also extend through space, over more than one bond. In the simplest case, for coupling through one bond, the origin mechanism is based on two interactions as illustrated in Figure (1.7).

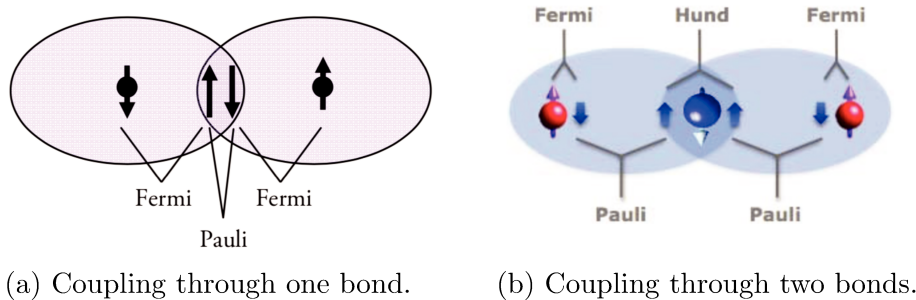


Figure 1.7: The number of involving bonds in J-coupling is responsible for its sign [57].

The magnetic dipole moment of the electron belonging to atom A creates a non-zero magnetic field at the A nucleus because of the Fermi contact interaction.³ The second electron participating to the bond and belonging to atom B has an opposite magnetic moment to the A electron one due to the Pauli principle. The B electron is involved also via the Fermi contact interaction with the B nucleus. The unification of these two interactions results in the J-coupling (Figure (1.7b)). It is noticed also that an odd number of bonds leads to positive J-coupling values, while an even number leads to negative values. i.e. for an even number of bonds the

³The Fermi contact interaction is the magnetic interaction between an electron and an atomic nucleus when the electron is inside that nucleus (Fermi contact interaction shows that an electron can "enter" nucleus). It is said also that this interaction is the main cause of hyperfine structure.

minimum energy configuration is obtained when both nuclear spins are parallel, meaning that for the Hamiltonian to be minimal J must be negative (Figure 1.7b) [57].

1.4.3 Magnetic properties in the electronic structure theory

Due to the complexity of such systems, theoretical investigations can bring important insights to the interpretation of experimental results. As discussed before, the (internal or external) magnetic field yields small effects on the energy. As a consequence, the use of perturbation theory in the calculation of NMR parameters is justified. This perturbation results only in even-order changes in the total energy [58, 59]:

$$E(\boldsymbol{\varepsilon}) = E_0 + \frac{1}{2!} \frac{d^2 E}{d\boldsymbol{\varepsilon}_1 d\boldsymbol{\varepsilon}_2} \boldsymbol{\varepsilon}_1 \boldsymbol{\varepsilon}_2 + \frac{1}{4!} \frac{d^4 E}{d\boldsymbol{\varepsilon}_1 \dots d\boldsymbol{\varepsilon}_4} \boldsymbol{\varepsilon}_1 \dots \boldsymbol{\varepsilon}_4 + \dots, \quad (1.73)$$

where E_0 is the energy at zero field and $\{\boldsymbol{\varepsilon}\}$ are the field strengths of the applied perturbations collected in vector $\boldsymbol{\varepsilon}$. With the perturbed energy expressions at hand, an analogy with the form of the Hamiltonian \hat{H}_{CS} in eq.(1.67) and \hat{H}_{SS}^{AB} in eq.(1.72) states that the shielding (σ) of the nucleus A and the scalar coupling tensor (J) of nuclei A and B are obtained as derivatives of the perturbed energy terms:

$$\sigma_A = \left. \frac{d^2 E(\boldsymbol{\mu}_A, \mathbf{B}_0)}{d\boldsymbol{\mu}_A d\mathbf{B}_0} \right|_{\substack{\boldsymbol{\mu}_A=0 \\ \mathbf{B}_0=0}}, \quad (1.74)$$

and

$$\mathbf{J}_{A,B} = \frac{\hbar}{2\pi} \gamma_A \gamma_B \left. \frac{d^2 E(\boldsymbol{\mu}_A, \boldsymbol{\mu}_B)}{d\boldsymbol{\mu}_A d\boldsymbol{\mu}_B} \right|_{\substack{\boldsymbol{\mu}_A=0 \\ \boldsymbol{\mu}_B=0}}. \quad (1.75)$$

In the case of static perturbations with strengths $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$, and assuming that the energy is optimised with respect to variational parameters at all field strengths, $\frac{\partial E}{\partial \boldsymbol{\kappa}_{pq}} = 0$, the second-order molecular property can be written as:

$$\left. \frac{d^2 E}{d\boldsymbol{\varepsilon}_1 d\boldsymbol{\varepsilon}_2} \right|_{\boldsymbol{\varepsilon}=0} = \left. \frac{\partial^2 E}{\partial \boldsymbol{\kappa}_{pq} \partial \boldsymbol{\varepsilon}_2} \frac{\partial \boldsymbol{\kappa}_{pq}}{\partial \boldsymbol{\varepsilon}_1} \right|_{\boldsymbol{\varepsilon}=0} + \left. \frac{\partial^2 E}{\partial \boldsymbol{\varepsilon}_1 \partial \boldsymbol{\varepsilon}_2} \right|_{\boldsymbol{\varepsilon}=0}, \quad (1.76)$$

The first contribution is determined perturbatively, with the first-order orbital rotation amplitudes, $\frac{\partial \boldsymbol{\kappa}_{pq}}{\partial \boldsymbol{\varepsilon}_1}$, obtained from the linear response (LR) equations:

$$0 = \frac{d}{d\boldsymbol{\varepsilon}_1} \left(\frac{\partial E}{\partial \boldsymbol{\kappa}_{pq}} \right) \Big|_{\boldsymbol{\varepsilon}=0} = \left(\frac{\partial^2 E}{\partial \boldsymbol{\kappa}_{pq} \partial \boldsymbol{\varepsilon}_1} + \frac{\partial^2 E}{\partial \boldsymbol{\kappa}_{pq} \partial \boldsymbol{\kappa}_{rs}} \frac{\partial \boldsymbol{\kappa}_{rs}}{\partial \boldsymbol{\varepsilon}_1} \right) \Big|_{\boldsymbol{\varepsilon}=0}, \quad (1.77)$$

that can be rewritten, in a matrix form, as follows [60, 58]:

$$0 = \mathbf{E}_{\boldsymbol{\varepsilon}_1}^{[1]} + \mathbf{E}_0^{[2]} \mathbf{X}_{\boldsymbol{\varepsilon}_1}, \quad (1.78)$$

where $\mathbf{E}_{\boldsymbol{\varepsilon}_1}^{[1]}$ is the property gradient, $\mathbf{E}_0^{[2]}$ is the electronic Hessian and $\mathbf{X}_{\boldsymbol{\varepsilon}_1}$ represents the solution vector giving $\{\boldsymbol{\kappa}_{rs}^{\boldsymbol{\varepsilon}_1}\}$. After determining $\mathbf{X}_{\boldsymbol{\varepsilon}_1}$, the static linear response function can be written as:

$$\langle\langle \boldsymbol{\varepsilon}_1; \boldsymbol{\varepsilon}_2 \rangle\rangle = \mathbf{E}_{\boldsymbol{\varepsilon}_1}^{[1]\dagger} \mathbf{X}_{\boldsymbol{\varepsilon}_2} = -\mathbf{E}_{\boldsymbol{\varepsilon}_1}^{[1]\dagger} \left(\mathbf{E}_0^{[2]} \right)^{-1} \mathbf{E}_{\boldsymbol{\varepsilon}_2}^{[1]}. \quad (1.79)$$

This leads to final forms of the shielding when $\boldsymbol{\varepsilon}_1 = \mathbf{B}_0$ and $\boldsymbol{\varepsilon}_2 = \boldsymbol{\mu}$ and the J-coupling in the case of $\boldsymbol{\varepsilon}_1 = \boldsymbol{\mu}_A$ and $\boldsymbol{\varepsilon}_2 = \boldsymbol{\mu}_B$.

1.5 Relativistic effects

The Schrödinger equation, is strictly valid, only for non-relativistic particles. When electrons move at velocities close to the speed of light, relativistic effects come into play. First, due to the relativistic mass, which increases with the velocity as:

$$m_{rel} = \frac{m_e}{\sqrt{1 - \left(\frac{v_e}{c}\right)^2}}, \quad (1.80)$$

where m_{rel} is the relativistic mass, m_e is the non relativistic one and v_e is the electron velocity. The effective Bohr radius $a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e^2}$ will decrease for large average speed. For 1s shell, at the non-relativistic limit this average speed is $Z a.u.$ Thus the 1s electron of $At(Z = 85)$ has a ratio $\frac{v}{c} (c = 137a.u.)$ of $\frac{85}{137}$ leading to a shrinkage of 22% ($m_e = 0.78m_{rel}$). This makes the neglect of relativistic effects not be acceptable unless for electrons with small kinetic energies which is the case for light atoms up to the early 3rd row of the periodic table. This assumption stops to be reasonable when heavy elements are present because the relativistic effects on properties are proportional to Z^4 [61]. The direct consequence of the relativistic effect on the electronic structure are the contraction and stabilisation of s (and p shells) and the splitting of the p, d, f, and higher shells. These result in, a screening of the nuclear charge for the outer shells which leads to a decreased effective nuclear charge and expansion and destabilisation of the outer valence d and f shells. The manifestations of these atomic effects are pronounced on molecular properties such as for NMR which is sensitive to the electron density near the nucleus. Moreover, the changes induced on valence orbitals can affect bond lengths, frequencies, valence spectra and ionisation energy, etc ... [62].

1.5.1 The Dirac equation

The relativistic energy of a free particle with a mass m_e and momentum \vec{p} is written as:

$$E = c\sqrt{\vec{p}^2 + m_e^2c^2}. \quad (1.81)$$

The passage from a classical equation to a quantum one is done by applying the associated quantum operators:

$$E = i\hbar\left(\frac{\partial}{\partial t}\right) \quad ; \quad \vec{p} = -i\hbar\vec{\nabla}. \quad (1.82)$$

The direct quantisation of Eq.(1.81) yields to a relativistic Hamiltonian that contains the square root of operators. This equation also contains the first derivative to time and the second derivative to space coordinates which does not satisfy the invariance with respect to the space-time Lorentz transformation.⁴

To fix this, Dirac assumed that the relativistic Hamiltonian, for one electron, should have a form that is similar to the one used in the Schrödinger equation,

$$\hat{h}_D\psi = E\psi, \quad (1.83)$$

in which the Dirac Hamiltonian is a linear operator with respect to momentum

$$\hat{h}_D = c(\vec{\alpha}\cdot\vec{p}) + \beta m_e c^2, \quad (1.84)$$

⁴The Lorentz transformations are a one-parameter family of linear transformations from a 4-coordinate (x,y,z,t) frame in space time to another frame that moves at a constant velocity, the parameter, within the former.

where $\vec{\alpha} = \{\alpha_x, \alpha_y, \alpha_z\}$ and β are the quantities that have to be determined, and $p = -i\hbar\{\nabla_x, \nabla_y, \nabla_z\}$. Therefore, Eq.(1.81) imposes that:

$$\beta^2 = I_4 \quad ; \quad \alpha_k\beta + \beta\alpha_k = 0 \quad ; \quad \alpha_k\alpha_l + \alpha_l\alpha_k = 2\delta_{kl}, \quad (1.85)$$

These conditions are fulfilled only if α and β are at least of dimension four (I_4 is the 4x4 identity matrix), which gives:

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad ; \quad \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad ; \quad \alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad (1.86)$$

and

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (1.87)$$

Thus, the solution to the Dirac equation (1.83) is a four 4-component vector referred to as a spinor:

$$\psi_{p=1,2,3,4} = \begin{pmatrix} \psi^L(r, t) \\ \psi^L(r, t) \\ \psi^S(r, t) \\ \psi^S(r, t) \end{pmatrix}, \quad (1.88)$$

associated to $E_{p=1,2,3,4}$ and where L and S are respectively the large and the small components of the wave function. The four solutions, two fold degenerate, E_+ and E_- energetically separated by $2mc^2$ yielding two Dirac equation solution continua for electrons and positrons. However a free electron has a positive energy which leads to select the two positive values E_+ associated to ψ_+ :

$$\psi_+ = \begin{pmatrix} \psi^L \\ \psi^S \end{pmatrix} = \begin{pmatrix} \psi_{\uparrow}^L \\ \psi_{\downarrow}^L \\ \psi_{\uparrow}^S \\ \psi_{\downarrow}^S \end{pmatrix}, \quad (1.89)$$

where, \uparrow and \downarrow denote the degree of freedom of the electron spin. Thus makes the relativistic Dirac formulation a more complete formulation with respect to the non-relativistic Schrödinger theory in predicting electron spin in lieu of postulating it a posteriori.

The addition of an external potential \hat{V} to the Dirac Hamiltonian leads to the hydrogen-like atoms Hamiltonian:

$$\hat{h}_D = c(\vec{\alpha} \cdot \vec{p}) + \beta' m_e c^2 + \hat{V} I_4 \quad ; \quad \beta' = \beta - I_4, \quad (1.90)$$

where the relativistic energy scale is aligned with the non-relativistic one. To extend the Dirac Hamiltonian to many-electron systems, one needs to add to the previous one-electron operator, the electron-electron repulsion. However a fully invariant description is not straightforward. Indeed electrons that propagate at the speed of light interact with a retardation effect. The quantum electron dynamics, with perturbation theory yields the two body expression:

$$\begin{aligned}
\hat{g}_{ij} &= \left[\frac{I_4}{r_{ij}} \right] - \left[\frac{\alpha_i \alpha_j}{2r_{ij}} \right] - \left[\frac{(\alpha_i \cdot r_{ij})(\alpha_j \cdot r_{ij})}{r_{ij}^2} \right] \\
&= \hat{g}_{ij}^{Coulomb} + \hat{g}_{ij}^{Gaunt} + \hat{g}_{ij}^{gauge} \\
&= \hat{g}_{ij}^{Coulomb} + \hat{g}_{ij}^{Breit},
\end{aligned} \tag{1.91}$$

where $\hat{g}_{ij}^{Coulomb}$ and \hat{g}_{ij}^{Breit} represent the electrostatic and magnetic interactions between electrons. As $c\alpha$ is identified to the velocity operator, the Gaunt term \hat{g}_{ij}^{Gaunt} account for current-current interactions, and is thus, gauge-dependent. Taking into account the difficulty to implement it, in practise 4-component calculations will use, either $\hat{g}_{ij} = \hat{g}_{ij}^{Coulomb}$, defining the Dirac-Coulomb (DC) Hamiltonian, or $\hat{g}_{ij} = \hat{g}_{ij}^{Coulomb} + \hat{g}_{ij}^{Gaunt}$ defining the Dirac-Coulomb-Gaunt (DCG) Hamiltonian.

1.5.2 Approximations to the Dirac equation

Working with 4-component wave functions in lieu of one for non relativistic representations makes these methods very demanding in terms of computing time and required memory [63, 64, 65] Thus, in order to push the domain of application of the Dirac equation, approximations can be done. Most methods aim at obtaining 2-components wave functions based on the eliminations of the small components which yields not only to simplify the 4-component formalism but also to reduce the computational costs [66, 67, 68]. The 2-component wave function can be obtained by applying a unitary block diagonalisation to the (usually the one-electron) Dirac Hamiltonian [69, 70, 68].

$$U^\dagger \begin{pmatrix} h_{D;LL} & h_{D;LS} \\ h_{D;SL} & h_{D;LL} \end{pmatrix} U = \begin{pmatrix} h_{++} & 0 \\ 0 & h_{--} \end{pmatrix}, \tag{1.92}$$

where U depends on the decoupling operators R unique to each approximate method,

$$U^\dagger \begin{pmatrix} 1 & -R^\dagger \\ R & 1 \end{pmatrix} \begin{pmatrix} (1 + R^\dagger R)^{-1/2} & 0 \\ 0 & (1 + RR^\dagger)^{1/2} \end{pmatrix}, \tag{1.93}$$

and where R is written in function of energy E:

$$R(E) = \frac{c(\vec{\sigma} \cdot \vec{p})}{E + 2mc^2 - V} = \frac{c(\vec{\sigma} \cdot \vec{p})}{2mc^2} K(E), \tag{1.94}$$

which is the solution of the equation:

$$\Psi^S = R\Psi^L, \tag{1.95}$$

and with K(E) having the form:

$$K(E) = \left(1 + \frac{E - V}{2mc^2} \right)^{-1}. \tag{1.96}$$

This term is subject of different approximations that yield different approximate Hamiltonians.

Pauli Hamiltonian

The Pauli Hamiltonian is based on the approximation of the coupling R by a Taylor expansion of the $K(E)$ term to the zero order and keeping only the terms with $\Theta(c^{-2})$:

$$R(E) \sim \frac{(\boldsymbol{\sigma} \cdot \mathbf{p})}{2mc} \quad ; \quad K \sim 1. \quad (1.97)$$

Thus, the Pauli Hamiltonian has the following form:

$$\hat{h}_i^{Pauli} = \frac{p_i^2}{2m_e} + V - \underbrace{\frac{p_i^4}{8m^3c^2}}_{\text{mass-velocity}} + \underbrace{\frac{1}{4m^2c^2} \vec{\sigma} \cdot [(\vec{\nabla} V) \wedge \vec{p}]}_{\text{SO}} + \underbrace{\frac{1}{8m^2c^2} (\nabla^2 V)}_{\text{Darwin}}.$$

This Hamiltonian [61] is of little use in practice, but it is useful to identify the different relativistic contributions. The mass-velocity and Darwin terms represent the scalar relativistic corrections, while the SO term is the spin-orbit one.

ZORA Hamiltonians

Singularity problems encountered with Pauli Hamiltonian (in particular in the mass-velocity term) [71] lead to reviewing the approximation made with the coupling R . A second possibility consist in writing $K(E)$ differently as:

$$K(E) = \left(1 + \frac{E - V}{2mc^2}\right)^{-1} = \frac{2mc^2}{2mc^2 - V} \left(1 + \frac{E}{2mc^2 - V}\right)^{-1} \quad (1.98)$$

The Taylor expansion of $(1 + \frac{E}{2mc^2 - V})^{-1}$ to the zero and the first order yields to approximations known respectively as the zeroth order regular approximation (ZORA) and the first order regular approximation (FORA). In ZORA [71, 72] which is considered as one of the most important approximation to Dirac Hamiltonian in electronic spectroscopy of heavy elements, the Hamiltonian is written as:

$$\begin{aligned} \hat{h}^{ZORA} &= \hat{V} + (\vec{\sigma} \cdot \vec{p}) \frac{c^2}{2mc^2 - V} (\vec{\sigma} \cdot \vec{p}) \\ &= \hat{V} + \vec{p} \frac{c^2}{2mc^2 - V} \vec{p} + i\sigma \cdot \left[\vec{p} \frac{c^2}{2mc^2 - V} \times \vec{p} \right]. \end{aligned} \quad (1.99)$$

This Hamiltonian includes no mass-velocity term, parts of the Darwin term and the full SO Hamiltonian leading to a two-component ZORA wave function. It is also possible to reduce the computational cost by neglecting spin-orbit effects, when we use the scalar relativistic ZORA Hamiltonian [73] which has the form (obtained by neglecting SO coupling):

$$\hat{h}^{SR} = V + p \frac{c^2}{2mc^2 - V} p. \quad (1.100)$$

Exact Two-Component Relativistic Hamiltonians

The eXact 2-Component (X2C) relativistic Hamiltonian [74] is based on the idea that it is easy to construct and diagonalise a matrix representation of a one-electron operator. The resulting eigenvectors are used, therefore, in a transformation that decouples exactly electrons energies from positrons ones. In the X2C approach based on molecular mean-field (X2Cmmf) [75],

the transformation is performed after a mean-field SCF four-component procedure (HF or DFT).

The X2C Hamiltonians are popular, in the computational chemistry communities [70], due to their particularity in treating spin-orbit interaction which makes them as exact 2-components Hamiltonians, in comparison with the Douglas-Kroll-Hess type [76], that requires more derivation. It is shown also that the X2Cmmf in the framework of the equation-of-motion for ionisation potentials (EOM-IP) yield results nearly indistinguishable from those obtained with the DC Hamiltonian [77].

It is noticeable that, while the transformations of Dirac Hamiltonian can be seen as straightforward for one-electron, it becomes of much more complexity for realistic systems for realistic systems creating the so-called picture-change errors that require also transforming properties Hamiltonians in order to prevent unreliable results [78, 70].

Chapter 2

Embedding methods

In the majority of cases, the quantum mechanical description of large molecular systems, for which experiments are made [79, 3]. is still challenging [80]. Some properties such as NMR parameters require high-level correlated methods in a relativistic framework. Wave function based methods which are the most accurate ones, are very demanding in terms of memory and computing time especially when heavy elements are present. The big challenge here is what we call the curse of dimension [81, 5], yielding to the unfavourable scaling with the system-size as shown in Table (1.4). On the other hand, density functional based methods known for their success for relatively extended systems offer in many cases an exit way to bypass the bottleneck. However, in many other cases, they fail in yielding good results since there is no systematic way to improve the functionals of the density and there is no reason that a successful functional for a given problem will succeed for another.

A compromise between the accuracy of the used theory and the size of the studied system must to be done. However in many cases, the desired information relates to a small part of the whole system and the surrounding medium is considered as a source of a small perturbation, therefore, applying a quantum mechanical description to the entire system may be inefficient. This fact pushes theoreticians to look for approaches that overcome this obstacle.

Early works were developed over the years in order to reconcile QM methods for large and very large systems by introducing possible simplifications and approximations that should be employed to make the QM approaches more practical. The idea behind these methods is to represent the environment with less accuracy. One can describe it either implicitly, (e.g. by electric charges, polarisabilities or atomic potentials) a polarisable continuum dielectric as PCM [82] or COSMO [83], or representing it by an embedding operator.

2.1 Implicit and QM/MM models

Modelling environmental effects and in particular solvent effects is not trivial. As one desire to model a condensed-phase system, many questions come to mind. The first step is to surround the solute with solvent molecules, but there are many critical questions. How many molecules shall one take into account? How many solvation shells are necessary? How large is the effect of solute charge on distant solvent molecules? Are there any hydrogen bonds and how many? The trivial answer is that we need an enormous system for which a quantum mechanical description is expensive or even impossible. Adding to this, the statistical nature of the interaction between solute and solvent molecules, requires an averaging over the phase

space [84] in order to include the effect of the temperature. This renders the result of a single geometry calculation of limited value and then hundreds geometries calculations are needed to achieve meaningful average values.

To prevent the huge cost of such calculations, continuum implicit solvation models come up with the idea of removing the large number of solvent molecules and modelling the emptied space as a continuous dielectric medium with the appropriate dielectric constant (ϵ). The molecular organisation of the solvent is totally ignored. The embedded system is not larger system than the gas phase solute in detriment of understanding the solvent. The solvation effect is introduced via an added perturbation to the gas-phase molecular Hamiltonian.

$$\hat{H}(r_m) = \hat{H}^{gas}(r_m) + \hat{V}(r_m). \quad (2.1)$$

The perturbation part \hat{V} represents the interaction between the solute and the solvent. It can be decomposed into four components which account for an electrostatic contribution, cavitation, dispersion and repulsion:

$$\hat{V} = \hat{V}_{electrostatic} + \hat{V}_{cavitation} + \hat{V}_{dispersion} + \hat{V}_{repulsion}. \quad (2.2)$$

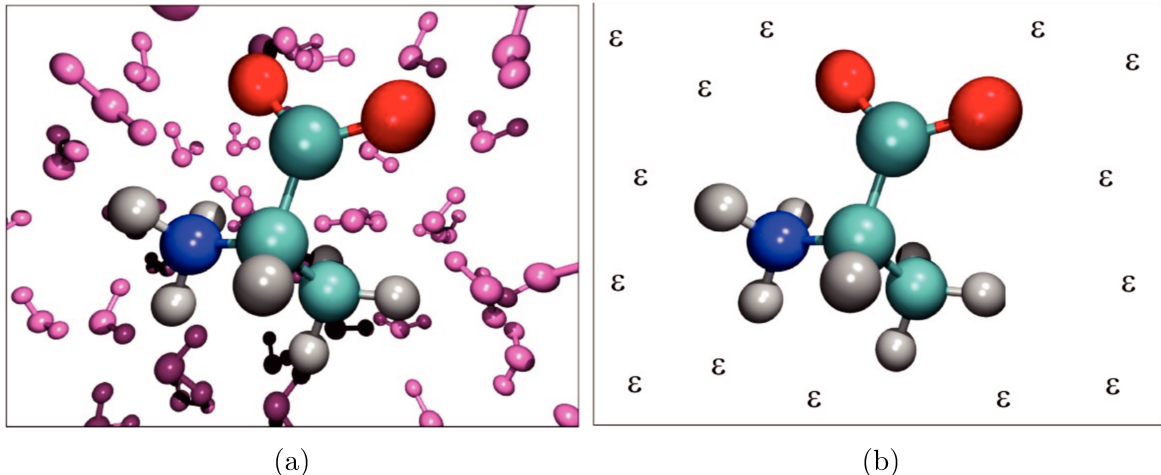


Figure 2.1: Replacement of the solvent molecules with a dielectric continuum

The electrostatic part accounts for the polarisation of the solute and solvent. The second term accounts for the energy needed to create cavity which is the equivalent volume used to encapsulate the solute. It represents the energy spent to compress the solvent in order to create the solute equivalent void in the solvent. The last two terms represent the quantum mechanical dispersion and repulsion energies.

Modelling these different parts could be realised by following different methods. The electrostatic part is described by the Poisson equation valid for a surrounding medium with linear response. The Poisson equation expresses the electrostatic potential as a function of the charge density and the dielectric constant (ϵ).

$$\nabla^2 \phi(r) = -\frac{4\pi\rho(r)}{\epsilon}. \quad (2.3)$$

This equation is valid if there is no ionic strength in the solution. Otherwise, in the case of ionic solutions and in order to respect the electroneutrality principle we will have displaced solvent charges toward the cavity surface which create a layer of surface charges and counter

ions or a double layer due to the thermal motion of ions. Thus, we obtain two regions or more in our solvent which means that the dielectric constant ϵ is no more uniform. Eq(2.3) is then written as in Eq(2.4) called the Poisson-Boltzmann (PB) equation for which the Poisson equation is a special case:

$$\nabla\epsilon(r).\nabla^2\phi(r) = -4\pi\rho(r). \quad (2.4)$$

This expansion is the base of the widely used Polarizable Continuum Model (PCM) which is by its turn, the origin of many variants [85]. The COSMO solvation model [86] is another popular implicit solvation model, in which scaled conductor boundary condition is used, which is a fast and robust approximation to the exact dielectric equations and reduces the outlying charge errors [87] as compared to the PCM model. Its efficiency lead to develop the COSMO-RS extension of COSMO which is designed to predict various thermodynamic quantities, including the free energy of solvation, for uncharged solutes in any organic solvent as well as solvent mixtures [88, 89].

Implicit continuum models are justified in liquids where the potential energy present several minima and therefore, the required properties need an averaging over hundreds of solvent conformations. These models are mostly used in biological applications such as folding or conformational transitions of proteins, DNA and RNA [90]. They are also used in studies aiming evaluating solvent effect on NMR parameters [85] especially for capturing long-range effects. But they cannot be used when trying to learn about solving dynamics or kinetics, which are obviously influenced by the discrete nature of solvents and possible existing hydrogen bonds. It is also to notice that implicit solvent models can have a difficult time with charged systems which are arguably more complicated to model and which interpellate some researchers, in many cases, to combine implicit descriptions with explicit simulation of ions [90] to remedy some encountered issues.

Understanding the solvent structural properties is one of the situations where an explicit representation is required. In fact, for relatively large systems neither quantum nor classical methods, applied separately to the whole system, can solve the problem. A priori, QM can solve any problem, but with increasing numbers of atoms it becomes impractical and combined QM/MM methods are unavoidable. The idea behind such methods, is that the whole system is split into two regions or more, a region where a quantum description is necessary while the other regions could be treated by classical methods. Thus, the hybrid QM/MM methods presented in 1976 by Warshel and Levitt [91] propose a compromise between the accuracy of quantum methods and the speed of classical ones in order to boost applications on systems with chemical processes in solution and in proteins.

Before proceeding to calculations, one should keep in mind some requirements. When dividing the whole system there should be a theoretical background behind it, in order to avoid losing main electronic features by the adopted fragmentation scheme. Moreover, as one obtains a QM/MM interface, a smooth transition has to be guaranteed in the description of near neighbour fragments zone which can raise significant difficulties. An equivalent Hamiltonian can be written in an hybrid way as in:

$$\hat{H}^{full} = \hat{H}^{QM} + \hat{H}^{MM} + \hat{H}^{QM/MM}, \quad (2.5)$$

where the \hat{H}^{QM} represents the description of the important quantum part while the second

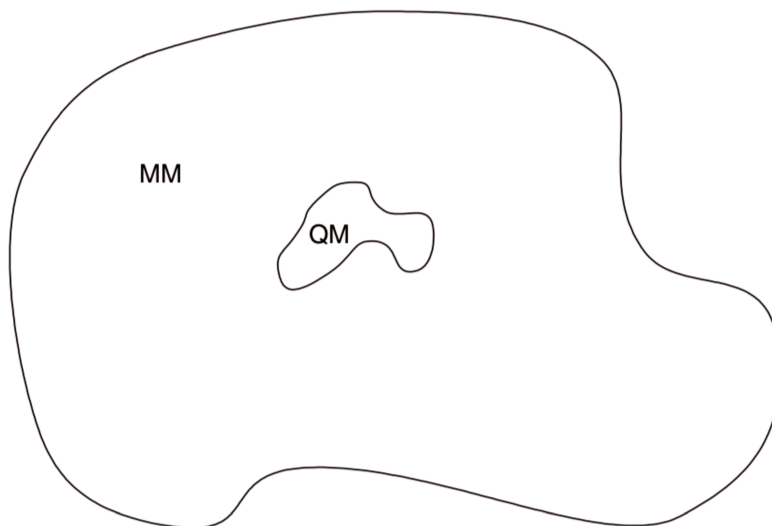


Figure 2.2: Quantum description could be applied only for a small part in the whole system

term \hat{H}^{MM} deals with the classically described region and the $\hat{H}^{QM/MM}$ manages the interaction between the quantum and the classical regions. Methods for the first two terms could be applied separately for systems for which these methods are convenient, while the third term needs more attention. Many formalism are developed based on this scheme and the main distinction is whether or not the separation between the areas cuts across any chemical bonds. If the boundaries are through space the combination could be realised with a given degree of simplicity.

The way the interaction energy between the different regions is described, determines the cost and the quality of the method. In general there are three types of coupling which are distinguishable by whether or not they include medium polarisation. The polarisation or not of the classically treated part is determined by the used force field (FF) which is the main ingredient of molecular mechanics and that describes all the interactions between different atoms and molecules. Force fields can be categorised in two principal types, polarisable and non-polarisable force fields. The simplest QM/MM scheme is the one based on an unpolarised coupling where we sum up calculated energies for different parts with the interaction energy. One of the issues in MM is that developed force field can not be at the same time general and suitable for all kind of molecules. Therefore it doesn't appear trivial how to model interactions between the classical area and given solute with some molecules.

In 1998, Kaminski and Jorgensen [92] proposed a simple scheme which they refer to as AM1/OPLS/CM1 (AOC). In their approach Monte-Carlo calculations are applied for the solute, and the OPLS force field for the solvent while the interaction energy is computed in the classical way for non bonded systems as in:

$$\hat{H}^{QM/MM} = \sum_m^{solute} \sum_n^{solvent} \left[\frac{\alpha q_m^{CM1} q_n}{r_{mn}} + 4\varepsilon_{mn} \left(\frac{\sigma_{mn}^{12}}{r_{mn}^{12}} - \frac{\sigma_{mn}^6}{r_{mn}^6} \right) \right], \quad (2.6)$$

where ε and σ are the Lennard-Jones parameters [93] and α is a parameter to deal with charged versus neutral molecules to treat solvent polarisation in the CM1 charge model applied to the

AM1 wave function [5]. But it still unpolarised interaction and this parameter controls the way that the solvent sees the solute if it is charged or neutral.

The next level of complexity or precision is categorised under the name of Polarized QM/Unpolarized MM where the relaxation of the quantum wave function induced by the environment is taken into account explicitly.

$$\hat{H}^{QM/MM} = \sum_{m:electrons}^{solute} \sum_{n:atoms}^{solvent} \frac{q_m q_n}{r_{mn}} + \sum_{p:nuclei}^{solute} \sum_{n:atoms}^{solvent} \left[\frac{Z_p q_n}{r_{pn}} + 4\epsilon_{mn} \left(\frac{\sigma_{mn}^{12}}{r_{ij}^{12}} - \frac{\sigma_{mn}^6}{r_{mn}^6} \right) \right]. \quad (2.7)$$

By this eq.(2.7) we allow the solvent to polarise the quantum mechanical part by separating the electrostatic interaction in eq.(2.6) into an operator acting on the QM electrons and a classical term for the interaction of the solvent atoms with the solute nuclei.

The final scheme in the QM/MM methods with bounds through the QM/MM frontier is the Fully Polarised Interactions for which a polarisable force field is used for the molecular mechanical part where each atom type carries a polarisability tensor α and we allow the induced dipoles to interact with the QM part.

So far, we discussed only systems for which the boundary between regions is everywhere characterised by a very low electron density, thus the $\hat{H}^{QM/MM}$ term is restricted to non-bonded interactions. The situation becomes more complicated if we try to cut regions across a chemical bond. Systems like big bulky groups surrounding the system of interest which are widely abundant in realistic molecules.

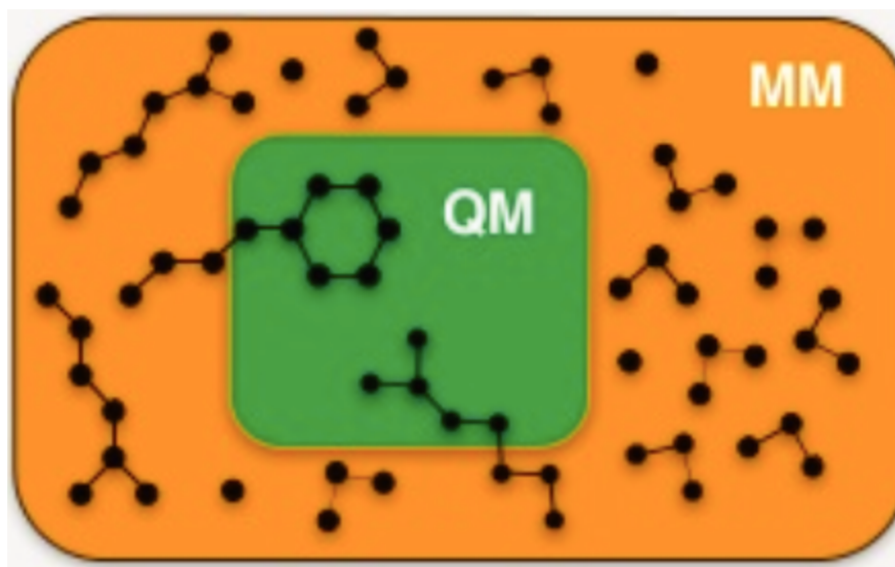


Figure 2.3: Cutting across boundaries is more sophisticated

In such situations it is undesirable to truncate the QM system by cleaving bonds because that yields to one or more unpaired electrons in the QM subsystem which is an unrealistic fragmentation. Moreover, the interaction between the two regions, QM and MM, should be small in the sense of perturbation theory which means that the two regions can be regarded as fully separated in the zeroth-order approximation. This limitation may be very hard when

bond breaking is involved in the QM/MM separation. A common solution to this artefact is to cap the cut bonds [94].

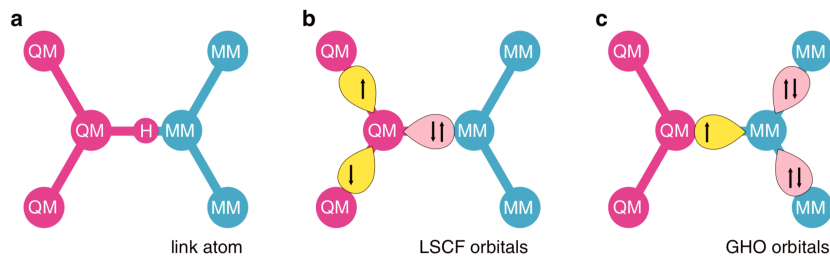


Figure 2.4: Capping bonds bring an alternative to cleaving bonds, link atoms (a) and frozen orbitals (b: local self consistent field [95], c: generalized hybrid orbital [96]).

The easy solution is to introduce a monovalent linked atom at an appropriate position along the bond between the two regions. Here, there is no restriction on the used linking atom, but a hydrogen atom is often used as cap in order to avoid the inclusion of bulky groups in the QM region. Localised orbitals represent also a popular alternative [91] which is proven to be more accurate but at the price of more parametrisation steps [97, 98]. The third category is characterised by adding a third auxiliary region giving birth to three partitions; the usual MM and QM systems in addition to a new area characterised by nuclei having their normal nuclear charges, and electron density expressed in some reduced sets of basis functions. As a result eq.(2.5) becomes:

$$\hat{H}^{full} = \hat{H}^{QM} + \hat{H}^{MM} + \hat{H}^{aux} + \hat{H}^{QM/aux} + \hat{H}^{aux/MM}. \quad (2.8)$$

The three new terms are, \hat{H}^{aux} which gathers the classical electrostatic interaction of the frozen density and its nuclei with themselves. The $\hat{H}^{aux/MM}$ which represents the classical interaction between the new auxiliary region and the MM one and the $\hat{H}^{QM/aux}$ which is similar to the $\hat{H}^{QM/MM}$ unless the molecular orbitals describing the new area are orthogonal to the QM part ones.

Finally, to avoid the question of how to deal with cut bond, one may assume that the electronic structure of the QM region will be of similar quality with either the non-polar, bulky group as a cap, or with simply hydrogen atoms as caps and all the important quantum effects are captured in the quantum system. In this view one can express the energy as:

$$E_{full} = E_{full}^{MM} + (E_{interest}^{QM} - E_{interest}^{MM}), \quad (2.9)$$

where the calculation for the total system is performed at the MM level in the first step. Then we assume that the interesting part requires a QM description to reach a high accuracy to the final description.

Implicit continuum and hybrid methods have encountered an enormous success in dealing with large systems [99]. However, with the huge expansion of calculation resources it appears feasible to treat extended systems with quantum mechanics all the way.

2.2 Quantum embedding models

In order to push the limits of the molecular size that can be treated in a quantum framework, one hopes that a linear scaling of computational cost with the molecular size can be achieved. If $S = A + B$ is the whole system, the computational effort to deal with S is much more greater than that dealing with the two subsystems A and B taken individually. The computational cost can be reduced by splitting S in more fragments in $S = A_1 + A_2 + B_1 + B_2$ as it is clear in Figure (2.5).

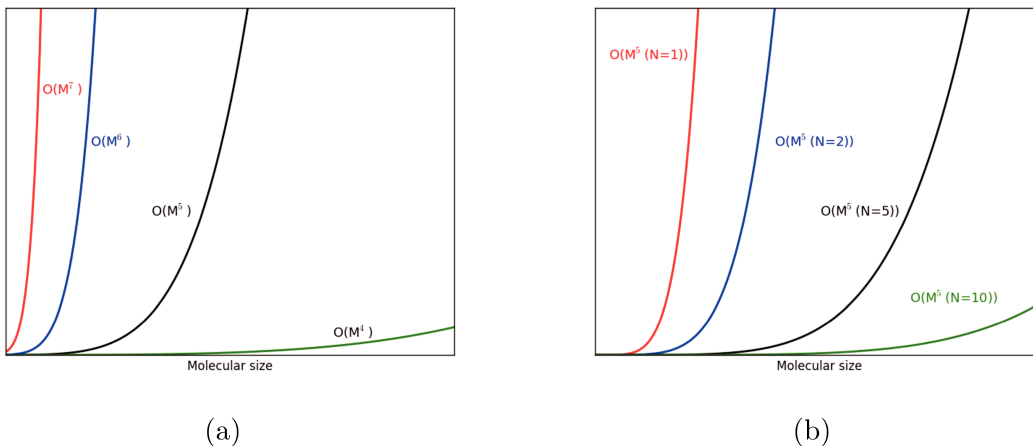


Figure 2.5: (a): The scaling behaviour of the computational cost with the molecular size (M); (b): The computational cost can be reduced significantly by dividing the whole system into subsystems. (N represents the number of subsystems).

In 1993 Wesolowski and Warshel [100] proposed a formulation of a quantum embedding approach. The Frozen Density Embedding (FDE) theory is a framework in which the local potential associated to the environment is not approximated. In contrast, it is constructed in an exact way and it is associated to total electron density of the whole system. The embedded wavefunction is an auxiliary quantity to optimise the partial embedded electron density. This scheme could be seen as an analogy of the atomic pseudopotential theory [101] in which the core orbitals are frozen and replaced by an effective core potential in order to focus the computational effort on optimising the valence orbitals that are active in different chemical and physical processes. The target in this analogy is shifted from valence orbitals to an embedded electron density, and from frozen core orbitals to frozen density. The FDE could also be seen as a universal formal framework besides any empirical QM/MM type of methods since the classical description of the environment can be considered as a special case of a quantum description.

2.2.1 Frozen Density Embedding (FDE) approach

The FDE approach is based on dividing the total system into two subsystems or more. Its formalism starts from dividing, in real space, the whole electron density $\rho(r)$ into the electron densities of subsystem of the interest $\rho_I(r)$ and of the environment $\rho_{II}(r)$ i.e.:

$$\rho(r) = \rho_I(r) + \rho_{II}(r), \quad (2.10)$$

and:

$$\rho_{II}(r) = \sum_n \rho_{II}^{(n)}(r). \quad (2.11)$$

With the condition that the partial densities integrate to the total number of electrons and their sum equals the total density over the whole space. It enables us to choose exactly the size of the system to be studied. Thus we can make the best compromise between the size of the interesting part and the needed cost to perform a calculation. In the following, we will restrict ourselves to two subsystems only, with the environment being treated as one subsystem.

Interaction energy and embedding potential

Within the DFT framework, partitioning the whole density into two subsystem densities leads us to write the DFT total energy as a bifunctional of ρ_I and ρ_{II} :

$$\begin{aligned} E[\rho_{tot}] &= E[\rho_I, \rho_{II}] = E_{NN} + T_s[\rho_{tot}] + J[\rho_{tot}] + V_{Ne}[\rho_{tot}] + E_{xc}[\rho_{tot}] \\ &= E_{NN} + T_s[\rho_I] + T_s[\rho_{II}] + T_s^{na}[\rho_I, \rho_{II}] + \frac{1}{2} \int \frac{(\rho_I(r) + \rho_{II}(r))(\rho_I(r') + \rho_{II}(r'))}{|r - r'|} d^3r d^3r' \\ &\quad + \int (\rho_I(r) + \rho_{II}(r))(\nu_{nuc}^I(r) + \nu_{nuc}^{II}(r)) d^3r + E_{xc}[\rho_I] + E_{xc}[\rho_{II}] + E_{xc}^{na}[\rho_I, \rho_{II}], \end{aligned} \quad (2.12)$$

where E_{NN} is the nuclear repulsion energy, ν_{nuc}^I and ν_{nuc}^{II} are the electrostatic potentials of the nuclei in subsystems I and II, respectively, E_{xc} is the exchange-correlation energy functional, and T_s is the kinetic energy of a reference system of noninteracting electrons with density ρ_{tot} . The non-additive terms are defined as follows:

$$T_s^{na}[\rho_I, \rho_{II}] = T_s[\rho_I + \rho_{II}] - T_s[\rho_I] - T_s[\rho_{II}], \quad (2.13)$$

and:

$$E_{xc}^{na}[\rho_I, \rho_{II}] = E_{xc}[\rho_I + \rho_{II}] - E_{xc}[\rho_I] - E_{xc}[\rho_{II}]. \quad (2.14)$$

The total energy as expressed in Eq (2.12) can be decomposed as the sum of the energies of two subsystems and an interaction energy:

$$E_{tot} = E_I + E_{II} + E_{int}, \quad (2.15)$$

where:

$$E_I = E_{NN}^I + T_s[\rho_I] + \frac{1}{2} \int \frac{\rho_I(r)\rho_I(r')}{|r - r'|} d^3r d^3r' + \int \rho_I(r)\nu_{nuc}^I(r) d^3r + E_{xc}[\rho_I], \quad (2.16)$$

and:

$$E_{II} = E_{NN}^{II} + T_s[\rho_{II}] + \frac{1}{2} \int \frac{\rho_{II}(r)\rho_{II}(r')}{|r - r'|} d^3r d^3r' + \int \rho_{II}(r)\nu_{nuc}^{II}(r) d^3r + E_{xc}[\rho_{II}]. \quad (2.17)$$

This allows us to use well suited functionals for each subsystems in order to obtain the best accuracy. The interaction energy that is written as:

$$\begin{aligned} E_{int}[\rho_I, \rho_{II}] &= E_{nuc}^{I,II} + T_s^{na}[\rho_I, \rho_{II}] + \int \int \frac{\rho_I(r)\rho_{II}(r')}{|r - r'|} dr dr' \\ &\quad + \int \rho_I(r)\nu_{nuc}^{II}(r) dr + \int \rho_{II}(r)\nu_{nuc}^I(r) dr + E_{xc}^{na}[\rho_I, \rho_{II}]. \end{aligned} \quad (2.18)$$

represents the exact interaction between the two subsystems. It contains repulsion terms between the two individual electron densities and the two nuclear distributions. It also contains attraction terms between the electron density of one subsystem and the nuclear potential of the other subsystem. Other terms are here to recover the non-additive parts of the energy.

Since we are not going to do direct calculations for the whole system, there is no available representation of ρ_{tot} in the canonical Kohn-Sham orbitals. Thus, the determination of non-additive terms of both kinetic and exchange-correlation energy, as expressed in eq.(2.13) and (2.14) respectively, necessitates in practical implementations, a numerical calculation using approximated functionals. The accuracy of the resulting energies and properties will depend on the way that we treat these non-additive terms.

At the beginning, the density of non-interacting isolated subsystems is different from the interacting subsystems density mainly at the boundaries. Each subsystem will change its electron density in the presence of the second one. This presence will be manifested by the addition of an embedding potential and each part of a given subsystem will be affected by the other subsystems.

In the case of Kohn-Sham formalism, the one-electron Kohn-Sham-like equation will be modified as follow:

$$\left[\hat{T}_i + \nu_{eff}^{KSi}[\rho_I](r) + \nu_{eff}^{emb}[\rho_1, \rho_{II}](r) \right], \phi_i^{(I)}(r) = \epsilon_i \phi_i^{(I)}(r) \quad ; \quad i = 1, \dots, N_I/2, \quad (2.19)$$

where $\nu_{eff}^{emb}[\rho_1, \rho_{II}](r)$ is the effective embedding potential which collects all the other terms arising from the interaction of the first subsystem with the frozen electron density and nuclei of the second subsystem. We can obtain the form of the embedding potential from the derivative of the interaction energy over the density of the influenced subsystem(I).

$$\begin{aligned} \nu_{eff}^{emb}[\rho_1, \rho_{II}](r) = \nu_{int}^I(r) = \frac{\delta E_{int}[\rho_I, \rho_{II}]}{\delta \rho_I} = \frac{\delta T_s^{na}}{\delta \rho_I} \Big|_{\rho_{tot}} - \frac{\delta T_s^{na}}{\delta \rho_I} \Big|_{\rho_I} + \int \frac{\rho_{II}(r')}{|r - r'|} d^3 r' \\ + \nu_{nuc}^{II}(r) + \frac{\delta E_{xc}^{na}}{\delta \rho_I} \Big|_{\rho_{tot}} - \frac{\delta E_{xc}^{na}}{\delta \rho_I} \Big|_{\rho_I}. \end{aligned} \quad (2.20)$$

Thanks to this potential, the interaction energy is exact with the assumption that the used functionals are exact. Therefore, we can obtain the exact density of the subsystem of interest at a minimal cost. This embedding potential can be calculated in a first step and then exported to perform more accurate DFT calculations, including relativity, or highly demanding wave function based calculations.

Nonadditive terms approximations

The main idea in embedding is to avoid supermolecule calculations, i.e. calculations for which the whole system is treated globally a unique uniform level of theory. Thus it is of great importance to treat with high attention all the components of the interaction energy. Nonadditive terms of equations (2.13) and (2.14) represent well-defined examples. In conventional KSDFT, as explained in chapter 1, the exchange and correlation energy is approximated and

it contains a part representing the complement of the non-interacting kinetic energy T_s which is obtained from a minimisation over orbitals.

$$T_s[\rho_n] = -2 \sum_{i=1}^{i=N/2} \langle \phi_i^n | A | \phi_i^n \rangle \quad (2.21)$$

The nonadditive kinetic energy as described in eq.(2.13) requires the $T_s[\rho_1 + \rho_{II}]$ term which, in turn, necessitates canonical KS orbitals representing the total density ρ_{tot} , which are in general no available and need to be numerically approximated [102]. Many approximations are already presents in the context of dealing with exchange-correlation energy (see section 1.3.1) such as the Thomas-Fermi used in the local density approximation (LDA).

There are many investigations aiming at finding the best functional approximating the kinetic energy [102, 103, 104]. These works show that the PW91k functional, also known under the name of GGA97, yields the most accurate description of the interaction energy. It reads:

$$T_s[\rho] = 2^{2/3} C_F \int \rho^{5/3}(r) F(s(r)) dr \quad (2.22)$$

where $C_F = \frac{3}{10}(3\pi^2)^{2/3}$ is the Thomas-Fermi constant and $F(s(r))$ is written as:

$$F(s(r)) = \frac{1 + A_1 s(r) \sinh^{-1}(A s(r)) + (A_2 - A_3 \exp(-A_4 s^2(r))) s^2(r)}{1 + A_1 s(r) \sinh^{-1}(A s(r)) + B_1 s^4(r)} \quad (2.23)$$

where $A, A_1, A_2, A_3, A_4, B_1$ are constants and $s(r)$ has been already defined in section (1.3.4).

The FDE scheme derives from an exact theoretical development and its accuracy mostly depends, as for DFT, on used functionals. However the interaction energy could be an additional source of errors since subsystem densities are sometimes poorly described with functionals. Therefore, subsystem density relaxations could be important in particular for charged system.

2.2.2 Subsystem DFT

In most cases, the approximate determination of environment can result in complementary ρ_I density that is not $\nu - representable$ ¹ and not positive everywhere. To remedy this deficiency, the so-called Freeze-and-Thaw (FnT) cycles procedure, in which the role of the frozen subsystem is interchanged between the subsystem of interest and the environment, is used. The new element here is that the total energy is calculated by a minimisation over all subsystems densities ρ_i . Thus the total energy has to satisfy, in the case of two subsystems, the following condition:

$$dE = \left(\frac{\delta E}{\delta \rho_1} \right) \delta \rho_1 + \left(\frac{\delta E}{\delta \rho_2} \right) \delta \rho_2 = 0, \quad \forall \delta \rho_1, \delta \rho_2. \quad (2.24)$$

Therefore eq.(2.19) will be transformed into a set of two coupled equations:

¹ $\nu - representable$ density is the ground state density of a system of N-electrons in the presence of some ν_{ext} and $n - representable$ is the density of a system of N-electrons (thats can be ground state or not. It's more general).

$$\begin{cases} \hat{T}_i + \nu_{eff}^{KS}[\rho_I](r) + \nu_{eff}^{emb}[\rho_I, \rho_{II}] \phi_i^{(I)} = \epsilon_i^I \phi_i^{(I)} & ; i = 1, \dots, N_I/2, \\ \hat{T}_i + \nu_{eff}^{KS}[\rho_{II}](r) + \nu_{eff}^{emb}[\rho_I, \rho_{II}] \phi_i^{(II)} = \epsilon_i^{(II)} \phi_i^{(II)} & ; i = 1, \dots, N_{II}/2, \end{cases} \quad (2.25)$$

where $\nu_{eff}^{emb}[\rho_I, \rho_{II}](r)$ and $\nu_{eff}^{emb}[\rho_{II}, \rho_I](r)$ are the embedding potentials representing subsystems II and I, respectively. In the Freeze-and-Thaw (FnT) formalism, these equations have to be solved iteratively by interchanging the role of the frozen subsystem as schematically illustrated by Figure (2.6):

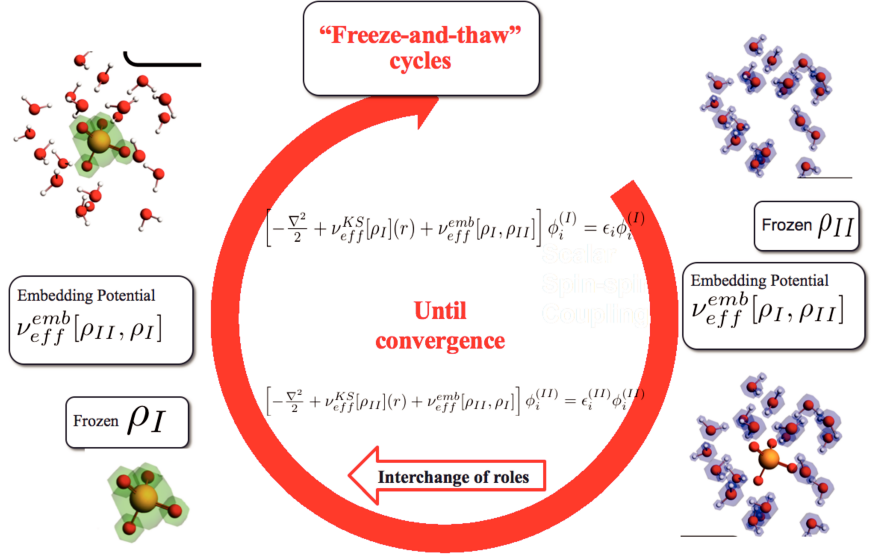


Figure 2.6: Freeze and thaw cycles procedure

The subsystem DFT [105] that provides an efficient alternative to conventional KS-DFT, is for great importance mainly when charged species are present. That is due to the polarisation effect of one subsystem on the other electron density.

2.2.3 Partition DFT

Before continuing on different features of the FDE method, it is convenient to present another embedding method which is closer to a supermolecule treatment. It has the particularity that partitioning the whole system density, and the subsystems could possess fractional charges unless the sum of all partition densities equals to the total system density. This method is called Partition DFT (PDFT) [106] and it is based on Kohn-Sham calculations on isolated fragments. This feature is dint by of the use partition theory (PT) [107, 108] which combines the simplicity of functional minimisation with a density optimisation in the goal of defining fragments in molecules and which has a remarkable success in overcoming limitations of earlier approaches for (chemical) reactivity [109, 110]. Partition theory deals with the problem of dividing a system into localised fragments as in:

$$\nu(r) = \sum_N \frac{Z_N}{|r - R_N|} = \sum_f \nu_f(r), \quad (2.26)$$

so,

$$\rho_{tot}(r) = \sum_f \rho_f. \quad (2.27)$$

As said, there is no condition on ρ_f which can be integrated to a non integer number since a meaningful fragment could possess a non integer charge. Once isolated fragments are defined, the total energy can be calculated using the formula:

$$E_{tot}(r) = \sum_f \min_{\rho_f} E_f[\rho_f] \quad || \quad \rho_{tot}(r) = \sum_f \rho_f. \quad (2.28)$$

Then, for each isolated fragment, one has to solve KS equation:

$$\left\{ -\frac{1}{2}\nabla^2 + \nu_\alpha[n_\alpha, \bar{n}_\alpha](r) \right\} \phi_j(r) = \epsilon \phi_j(r), \quad (2.29)$$

where ν_α is the effective KS potential of the α -th fragment and $n_\alpha = n - \bar{n}_\alpha$ and generating their self-consistent KS potentials, densities $n_\alpha^{(0)}(r)$, and overlapped sum $n^{(0)}(r)$. These quantities are reintroduced in Eq. (2.29) to produce new quantities. The procedure is iterated until convergence is reached. Thus PDFT maps a problem of interacting fragments into an effective isolated fragment problem, just as the KS scheme does for a system of interacting electrons into a noninteracting electron system. By its formulation, PDFT has many advantages, namely the total energy of the system never needs to be calculated directly and the dissociation energy for a given fragment is computed from the difference between its energies in the first isolated calculation $E_f^{(0)}$ and the final iteration E_f . Furthermore, it ensures correct dissociation energies for molecules with simply constraining occupations to be those of the isolated fragments. Moreover, there are no formal difficulties arising from taking density variations within a fixed density, as the trial molecular density is simply the sum of the fragment densities, which are varied freely [106]. However, the fact that there is no constraint on the subsystem charge numbers could be less appealing and useful since most correlated wave function methods require an integer number in calculation [111]. Moreover, potential inversion techniques are known to be expensive [112].

2.2.4 FDE extension to wave function/DFT embedding

In FDE and subsystem DFT, all subsystems are described using DFT. However, in many cases, as discussed before, DFT presents some problems in computing some properties [add citation]. Thus the use of highly accurate wave function based methods is recommended. By its formulation, FDE enables us the use of wave function methods for selected subsystems and that is due to its use of electron density which is accessible at every quantum chemical level [113, 114]. For this purpose, one can replace the subsystem of interest (subsystem I, for example) terms in the energy expression as in Eqs (2.15), (2.16) initially expressed in the DFT framework by the target wave function descriptor.

The subsystem of interest with the density ρ_I is then described by the embedded wave function via the embedded Schrödinger equation [113, 115] as in:

$$[\hat{H}^{wf} + \hat{V}^{emb}] \psi_I = E_I \psi_I, \quad (2.30)$$

where \hat{V}^{emb} is the embedding potential and \hat{H}^{wf} is a wave function based method Hamiltonian. The interaction energy and the embedding potential are still expressed in the KS-DFT formalism and the only trace of the wave function method trace is the subsystem I density ρ_I . This scheme could be also seen as a hybrid approach as two different quantum levels are used to

describe the part of interest and the surrounding medium.

The WFT-in-DFT embedding schemes in the framework of FDE and subsystem DFT are usually used when we first need to improve results coming generally from DFT-in-DFT embedding calculations, or wherever DFT performs poorly or not accurately in computing the desired properties especially for time-dependent one's. The area of applicability encompasses molecules adsorbed on surfaces, impurities in solids, of solvated molecules, mainly when the desired property is localised.

2.2.5 FDE for second-order magnetic properties

In order to evaluate shielding parameters, the FDE was extended in 2006 [116]. However, the use of FDE for calculating second-order magnetic properties is not straightforward, since they depend on the total paramagnetic current which is neglected in nonadditive kinetic energy functionals. It was shown [116] that the FDE can yield good induced chemical shift results, within 2 ppm to the KSDFE supermolecule calculations. Moreover, this approach was generalised to formulate the calculation of nuclear spin-spin coupling constants [117] with the additional approximation of neglecting the contribution of environment to the spin magnetisation density. The FDE performs differently in describing the effect of the environment. It reproduce 94% of the effect of solvent on $\text{NH}_3\text{-H}_2\text{O}$ compared to KSDFE calculation on the whole system [118]. This ratio decreases to less than 80% for strongly interacting complexes, where the underestimation is mainly due to the failure of the approximations used in FDE in accurately describing the solvent-to-metal charge donation [118].

In practice FDE, the electronic Hessian $\mathbf{E}_0^{[2]}$ and the property gradient $\mathbf{E}_{\epsilon_1}^{[1]}$ previously presented in the section (1.4.3) take a subsystem shape:

$$\mathbf{E}_0^{[2]} = \begin{bmatrix} \mathbf{E}_0^{[2];M,M} & 0 \\ 0 & \mathbf{E}_0^{[2];N,N} \end{bmatrix} + \begin{bmatrix} \mathbf{E}_{0,int}^{[2];M,M} & \mathbf{E}_{0,int}^{[2];M,N} \\ \mathbf{E}_{0,int}^{[2];N,M} & \mathbf{E}_{0,int}^{[2];N,N} \end{bmatrix}, \quad (2.31)$$

and:

$$\mathbf{E}_{\epsilon_1}^{[1]} = \begin{bmatrix} \mathbf{E}_{\epsilon_1}^{[1];M} & \mathbf{E}_{\epsilon_1}^{[1];N} \end{bmatrix}^\dagger + \begin{bmatrix} \mathbf{E}_{\epsilon_1,int}^{[1];M} & \mathbf{E}_{\epsilon_1,int}^{[1];N} \end{bmatrix}^\dagger, \quad (2.32)$$

where $M, N \in \{\text{subsystems } I, II\}$ and $M \neq N$ and lead to a system of LR equations:

$$\begin{cases} \left(\mathbf{E}_0^{[2];M,M} + \mathbf{E}_{0,int}^{[2];M,M} \right) \mathbf{X}_{\epsilon_1}^M + \mathbf{E}_{0,int}^{[2];M,N} \mathbf{X}_{\epsilon_1}^N = - \left(\mathbf{E}_{\epsilon_1}^{[1];M} + \mathbf{E}_{\epsilon_1,int}^{[1];M} \right) \\ \mathbf{E}_{0,int}^{[2];N,M} \mathbf{X}_{\epsilon_1}^M + \left(\mathbf{E}_0^{[2];N,N} + \mathbf{E}_{0,int}^{[2];N,N} \right) \mathbf{X}_{\epsilon_1}^N = - \left(\mathbf{E}_{\epsilon_1}^{[1];N} + \mathbf{E}_{\epsilon_1,int}^{[1];N} \right) \end{cases} \quad (2.33)$$

where the response vector has also been split into blocks pertaining to each subsystem,

$$\mathbf{X}_{\epsilon_1} = \begin{bmatrix} \mathbf{X}_{\epsilon_1}^M & \mathbf{X}_{\epsilon_1}^N \end{bmatrix}^\dagger. \quad (2.34)$$

2.2.6 Challenges

The FDE represents a very defined and exact framework to deal with large systems. However, there are some difficulties, especially with the nonadditive kinetic energy T_s^{na} and the corresponding potential in describing some situations such as for system were the embedded

species and the environment are linked by covalent bonds. Approximating T_s^{na} can be bypassed in the projector-based embedding method which uses localised occupied orbitals to divide the system at the price of new challenges coming from approximating E_{xc}^{na} [119].

The failure of the approximations used in FDE to describe the solvent-to-metal charge donation is also challenging making the FDE suffering in reproducing the effect of solvent on properties compared firstly to the whole system calculations and secondly to experimental results [118].

However, it is noticeable as mentioned before, despite that the used approximations can engender some errors, the FDE approach lies on exact formalism, opening the way to improve its performance.

Chapter 3

Ionisation energies of solvated halide Ions with relativistic embedded equation of motion coupled cluster theory

In this work we describe a general computational approach capable of obtaining binding energies for valence states of solvated ionic species in the current example, halide ions from fluoride to astatide in water.

The approach is based on electronic structure calculations accounting for electron correlation (with the equation of motion coupled-cluster method for electron detachment method, EOM-IP-CCSD) relativistic effects (with four-component Dirac-Coulomb Hamiltonian) and the interaction between the anions and the water environment (with the frozen density embedding method, FDE). Furthermore, we incorporate temperature and statistical effects by coupling the electronic structure calculations with classical molecular dynamics with sophisticated polarisable force fields devised to faithfully represent the interaction between water molecules themselves and with ionic species, simulating the solvent as a discrete droplet model.

200,000 hours of computing time (PhLAM cluster and GENCI supercomputers) lead to interesting results. First, it has shown that we can use embedding approaches to couple molecular electronic structure methods (coupled cluster and DFT), and those to molecular dynamics, to obtain very accurate binding energies for the whole system (solute and solvent). This is in contrast to state-of-the-art electronic structure methods based on periodic boundary conditions, which obtain results of similar accuracy than ours only if sophisticated (and computationally very demanding) density functionals and Green's function based methods such as G_0W_0 are used.

Second, our investigation shows the reliability of the SAOP model potential for obtaining very good first approximations to electron binding energies. This method is still not very popular in molecular electronic structure applications, and is largely unexplored by the condensed matter community in spite of the obvious performance gains it could bring about if combined, for instance, to the G_0W_0 . We expect that our work draws attention to its potential for the two communities.

Third, our embedding approach makes it extremely easy to employ different electronic Hamiltonians for different parts of the system. This allows for the rigorous inclusion of spin-orbit

interactions and scalar relativistic effects in the calculations, essential if one wishes to compare directly to experiment for atoms beyond Gallium. To the best of our knowledge, due to the cost of two- or four-component approaches (such as those used here) are generally not applied to condensed matter calculations, and our approach can bridge this gap while periodic four-component approaches are not available.

3.1 Experimental context

The electronic properties of solvated ions are attracting much interest for their role in electrochemical processes occurring at aqueous interfaces, as well as in several atmospheric, biochemical and biological processes [120]. Moreover, in general the importance of liquid water in almost all aspects of life is indisputable. Despite the apparent simplicity of the water molecule structure, the corresponding liquid is very complex, and the intermolecular hydrogen-bonding network is an important factor for understanding its different characteristics. This network leads to complex structural and properties changes in the liquids on ultrafast time scales [121, 122, 123, 124, 125, 126]. Therefore, understanding water electronic structure may provide very helpful insights in various domains where the water is present.

Many investigations on the PE spectra of water yield their orbital energies (liquid and gas) that are reported in Table (3.1) and illustrated in Figure (3.1).

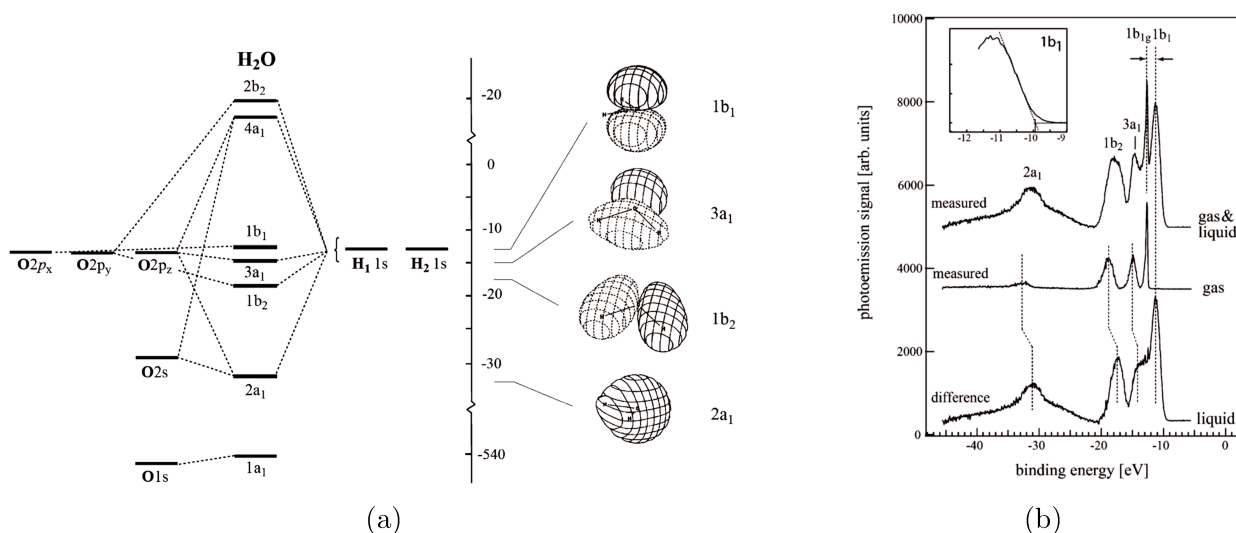


Figure 3.1: (a) : (Left) Energy level diagram of occupied molecular orbitals of gas-phase H₂O [127]. (Right) Orbital pictures of the H₂O molecule. Different signs of the wave function are indicated by solid and dashed lines [128, 129]; (b): Photoemission spectra from (top) liquid water, (center) gas-phase water, measured at 60 eV photon energy, and (bottom) difference spectrum [129].





orbital	binding energy/eV		shift/eV gas to liquid
	gas	liquid	
$1b_1$ 	12.60 [130]	11.16(4)	1.45(5)
$3a_1$ 	14.84(2)	13.50(10)	1.34(12)
$1b_2$ 	18.78(2)	17.34(4)	1.46(6)
$2a_1$ 	32.62(10)	30.90(6)	1.72(16)

Table 3.1: [131] The water bands are labeled following the convention for a C_{2v} symmetry H_2O molecule. The chloride ion peaks

The observed water gas-to-liquid-phase peak shifts water, that are not systematic, are due to different contributions. Electronic polarisation, changes in the orbital structure due to hydrogen bonding and surface dipoles. The last effect which is small, arises from oriented water molecules. The hydrogen bonding effect on the H_2O orbital structure and energies is behind the principal contribution to gas-liquid peak shifts. However, at the present moment, experimental quantification of this effect is not possible.

In contrast, the effect of water on other species can be dramatic, especially when charged elements are present. Alkali Halide salts are an example of aqueous solution where the PE spectra are drastically changed as graphically illustrated in Figure (3.2).

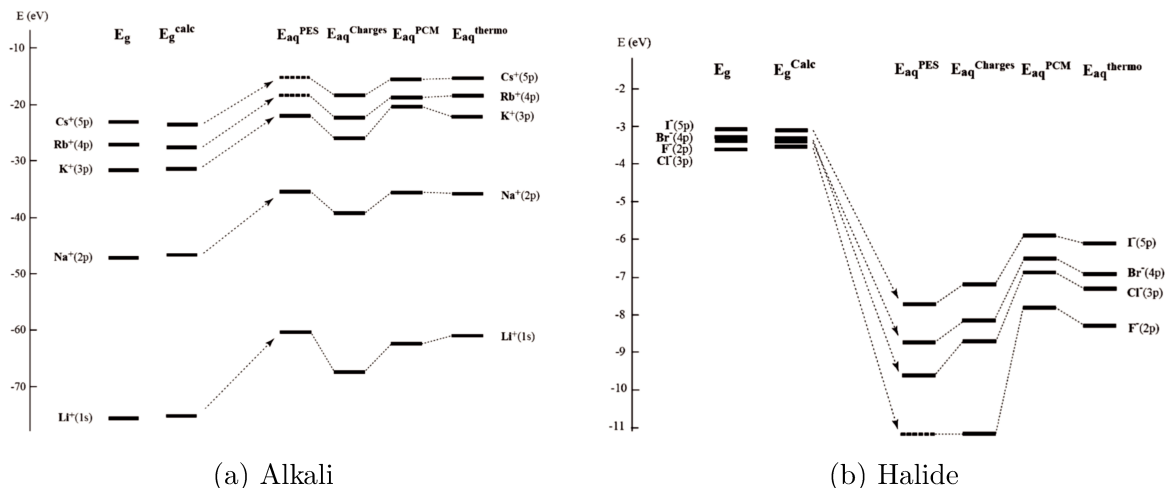


Figure 3.2: Diagram of lowest electron binding energies of aqueous (a) alkali and (b) halide, comparing experimental (E_{aq}^{PES} , and calculated values ($E_{aq}^{charges}$, E_{aq}^{PCM} , E_{aq}^{thermo}) and experimental E_g and calculated E_g^{calc} values for the gas phase [131].

From these two Figures. we establish that E_{aq}^{PES} for cations is smaller than the respective E_g , whereas the opposite is true for aqueous anions. Although the energy difference is considerably larger for alkali than for halide. Moreover, it is observable that the theoretical models nicely reproduce the experimental energies for alkali, whereas the situation for the halide is more complicated. On the contrary, for halide anions, electron detachment results in a significant reorientation of the solvation shell; water molecules prefer to point with hydrogen atoms toward the anion but with oxygen atoms toward neutral halogen atoms. In addition, there is a smaller change in the electronic polarisation of the surrounding water by the formation of the neutral halogen atom. Therefore, for the anions, it is more important to have a vertical

treatment of the nuclear polarisation, which also rationalises the poor performance of adiabatic continuum models.

In the next, we discuss the mutual effect of halide and water on their binding energies.

Predictive Simulations of Ionization Energies of Solvated Halide Ions with Relativistic Embedded Equation of Motion Coupled Cluster Theory

Yassine Bouchafra, Avijit Shee,[†] Florent Réal, Valérie Vallet, and André Severo Pereira Gomes^{*}
Université de Lille, CNRS, UMR 8523—PhLAM—Physique des Lasers, Atomes et Molécules, F-59000 Lille, France

 (Received 6 November 2018; published 28 December 2018)

A subsystem approach for obtaining electron binding energies in the valence region is presented and applied to the case of halide ions (X^- , $X = F - \text{At}$) in water. This approach is based on electronic structure calculations combining the relativistic equation-of-motion coupled cluster method for electron detachment and density functional theory via the frozen density embedding approach, using structures from classical molecular dynamics with polarizable force fields for discrete systems (in our study, droplets containing the anion and 50 water molecules). Our results indicate that one can accurately capture both the large solvent effect observed for the halides and the splitting of their ionization signals due to the increasingly large spin-orbit coupling of the $p_{3/2}$ - $p_{1/2}$ manifold across the series, at an affordable computational cost. Furthermore, owing to the quantum mechanical treatment of both solute and solvent electron binding energies of semiquantitative quality are also obtained for (bulk) water as by-products of the calculations for the halogens (in droplets).

DOI: 10.1103/PhysRevLett.121.266001

Photoelectron (PE) spectroscopy [1] is a particularly powerful technique (now often complemented by electronic structure calculations) to investigate bound states at the valence or inner regions, either to obtain information on the nature of bonding for species in the gas phase [2–4], in solution [5,6], or at interfaces [7–9] or to follow and identify chemical changes in complex media [10–12]. Such techniques have been extensively used to investigate species such as halogens and halogen-containing species [13–15], which are of great importance in atmospheric processes [16,17] such as photochemical reactions leading to ozone depletion, or aerosol formation [18].

The simplest halogenated systems of relevance are the halides, originating mostly from marine aerosols [19], and understanding how these species interact with water is, apart from its intrinsic interest, of importance for better understanding their effects in the environment. Experimental studies on clusters [20] and bulk [21] aqueous solutions have established that there are very large shifts in the PE spectrum of the halides upon solvation, highlighting strong interactions between the anions and the water solvent. Early theoretical studies determined the halides' electron binding energies (BEs) by employing *ab initio* calculations [22–24] or combining these with classical molecular dynamics simulations with periodic boundary conditions [21]. These studies indicate that not including specific interactions (hydrogen bond, etc.) between the halogens and the solvent water molecules leads to a poor description of the halide BEs [21,25], apart from the fact that quantum-classical approaches cannot yield the electronic structure of the solvent.

Currently the most sophisticated theoretical approaches to obtain PE spectra for the whole system quantum mechanically (“full QM”) rely upon density functional theory (DFT) to obtain the ground state for the solvent-solute system (as in Car-Parrinello molecular dynamics (CPMD) [26]), followed by the use of many-body Green's function (MBGF)-based perturbation theories (e.g., *GW* and variants such as G_0W_0 [6,27–31]). MBGF approaches are not without downsides: The first is their high computational cost for fully self-consistent variants, especially if the calculations employ periodic boundary conditions and require large (super)cells. A second, and more serious, issue is the lack of exchange diagrams in self-energy beyond first order. This is particularly a shortcoming in the treatment of molecular systems.

GW-based approaches have been shown to introduce relatively large errors for the calculation of BEs [32,33] compared to reference single-reference coupled cluster [CCSD(T)] or equation-of-motion coupled cluster for electron detachment (EOM-IP-CCSD) [34,35] calculations. Recent benchmarking studies suggest that even lower-scaling, approximate variants to the EOM-CCSD method [36,37] can be competitive in accuracy with *GW* calculations of ionizations and electron affinities, and especially so for G_0W_0 [33].

This Letter presents a full-QM electronic structure approach for obtaining BEs of discrete systems such as water-halide ion (X^- , $X = F - \text{At}$) aggregates, as a cost-effective yet accurate alternative to *GW*-based calculations, by coupling relativistic EOM-IP-CCSD calculations for the halides (since relativistic effects, and in particular

spin-orbit coupling (SOC) [38], on the BEs are increasingly important along the halogen series) and scalar relativistic DFT calculations for the water molecules through the frozen density embedding (FDE) method [39–41].

The key idea of FDE (see Refs. [42–45] for further details and its relationship to other embedding methods) is the partitioning of a system's electron density $n(\mathbf{r})$ into a number of fragments [for simplicity two such fragments are considered here, so $n(\mathbf{r}) = n_I(\mathbf{r}) + n_{II}(\mathbf{r})$] and total energy $E[n(\mathbf{r})]$, which can be rewritten as a sum of subsystem energies ($E_i[n_i(\mathbf{r})]$, $i = I, II$) plus an interaction energy ($E_{(\text{int})}$):

$$E[n] = E_I[n_I] + E_{II}[n_{II}] + E_{(\text{int})}[n_I, n_{II}]. \quad (1)$$

The latter collects the intersubsystem interaction terms,

$$\begin{aligned} E_{(\text{int})}[n_I, n_{II}] = & \int [n_I(\mathbf{r})v_{\text{nuc}}^{II}(\mathbf{r}) + n_{II}(\mathbf{r})v_{\text{nuc}}^I(\mathbf{r})]d\mathbf{r} \\ & + \iint \frac{n_I(\mathbf{r})n_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \\ & + E_{\text{xc}k}^{\text{nadd}}[n_I, n_{II}] + E_{\text{nuc}}^{I,II}, \end{aligned} \quad (2)$$

where v_{nuc}^i is the nuclear potential ($i = I, II$), $E_{\text{nuc}}^{I,II}$ the nuclear repulsion energy between subsystems, and $E_{\text{xc}k}^{\text{nadd}}$ accounts for nonadditive contributions due to the exchange-correlation (xc) and kinetic energy (k) contribution. $E_{\text{xc}k}^{\text{nadd}}$ is defined as

$$\begin{aligned} E_{\text{xc}k}^{\text{nadd}}[n^I, n^{II}] = & E_{\text{xc}}^{\text{nadd}}[n^I, n^{II}] + T_s^{\text{nadd}}[n^I, n^{II}] \\ = & E_{\text{xc}}[n^I + n^{II}] - E_{\text{xc}}[n^I] - E_{\text{xc}}[n^{II}] \\ & + T_s[n^I + n^{II}] - T_s[n^I] - T_s[n^{II}]. \end{aligned} \quad (3)$$

The nonadditive kinetic energy contribution provides a repulsive interaction that offsets the attractive interaction between the nuclear framework of one subsystem and the density of the other [46], which, if not properly matched, can lead to spurious delocalization of the electron density of one subsystem over the region of the other [47] (as seen, for instance, in point-charge or QM–molecular mechanics embedding [48]). For reasons of computational efficiency, the FDE calculations in this Letter employ approximate kinetic energy density functionals [49] which provide good but nevertheless limited accuracy [50] for systems such as those discussed here, which are not covalently bound.

In a purely DFT framework, the density for a subsystem of interest n_I is obtained by minimizing the total energy [Eq. (1)] with respect to variations on n_I while keeping n_{II} frozen, yielding Kohn-Sham-like equations,

$$[T_s(i) + v_{\text{KS}}[n_I] + v_{\text{int}}^I[n_I, n_{II}] - \varepsilon_i]\phi_i^I(\mathbf{r}) = 0, \quad (4)$$

where $v_{\text{KS}}[n_I]$ and $T_s(i)$ are the usual Kohn-Sham potential and kinetic energy (from $\delta E_I[n_I]/\delta n_I$), and

$$v_{\text{int}}^I(\mathbf{r}) = v_{\text{xc}}^{\text{nadd}}(\mathbf{r}) + \left. \frac{\delta T_s^{\text{nadd}}}{\delta n} \right|_{n_I} + v_{\text{nuc}}^{II}(\mathbf{r}) + \int \frac{n_{II}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' \quad (5)$$

is the embedding potential (from $\delta E_{(\text{int})}[n_I, n_{II}]/\delta n_I$), which describes the interaction between subsystems.

FDE provides a formally exact framework that allows DFT to be replaced by wave function theory (WFT)-based treatments for one [51–54] (WFT-in-DFT) or all subsystems [55] (WFT-in-WFT), with the embedding potential being calculated from Eq. (5) irrespective of the level of electronic structure employed, though using the electron densities from the respective methods. Obtaining electron densities for WFT methods in general and coupled cluster in particular is computationally expensive (the latter requiring the solution of the ground state CC Λ -equations [34]), and it has been found that an approximate scheme—where v_{int}^I is obtained from preparatory DFT-in-DFT calculations [53,56] and treated as a (local) one-electron operator added to the Fock matrix in the WFT calculations—works very well in practice. This latter prescription is the one followed here.

In the EOM-IP-CCSD method, BEs are obtained from the solution of the eigenvalue equation [35,57]

$$(\bar{H}R_k^{\text{IP}})_c = \Delta E_k R_k^{\text{IP}} \quad (6)$$

where ΔE_k is the k th ionization energy for the system, $\bar{H} = e^{-T}\hat{H}e^T$ is the (CCSD) similarity transformed Hamiltonian [here including $v_{\text{int}}^I(\mathbf{r})$] and

$$R^{\text{IP}} = \sum_i r_i \{i\} + \sum_{i>j,a} r_{ij}^a \{a^\dagger j i\} \quad (7)$$

the wave operator that transforms the CC ground state to the electron detachment states.

In the preparatory DFT-in-DFT calculations, the statistical average of model orbital potentials (SAOP) [58] has been used. This potential is constructed to yield Kohn-Sham potentials showing proper atomic shell structure and correct asymptotic behavior, and with it calculations have a computational cost equivalent to Kohn-Sham DFT using generalized gradient approximations. The SAOP orbital energies have been shown to provide BEs that are in very good agreement with coupled cluster calculations [59]. Given the evidence in the literature that Kohn-Sham densities obtained with functionals yielding accurate BEs compare quite well to densities obtained with coupled cluster methods [60,61], a v_{int}^I obtained with SAOP densities should provide a good approximation to one obtained with coupled cluster densities, with the advantage that one obtains a representation for the PE spectrum of water at no additional cost.

The FDE calculations were performed on structures obtained with classical molecular dynamics (CMD) simulations on water-halide droplets containing 50 water molecules and constraining the halogen to be fixed at the droplet's center of mass, using the POLARIS (MD) code [62–65] and many-body force fields [66] accounting for both polarization effects and the bonding effects within the water molecules (hydrogen bonds), and between the halide and first-hydration shell water units (strong hydrogen bond). From these, after equilibration of the system, were extracted 200 snapshots, which were verified as uncorrelated for the BEs (see the Supplemental Material [67]). A particular feature of the droplet structures for all halogen species, such as that shown in Fig. 1 for a snapshot of solvated I^- , is that the water distribution around the anion is not spherical but elongated due to strong polarization effects that favor disymmetrized structures, with about six to eight water molecules making up the first solvation shell.

The total system was partitioned into two subsystems, the halide (subsystem I) and the 50 water molecules (subsystem II), corresponding to the simplest partition to calculate the halide BEs (referred to as $[\text{X}^- @ (\text{H}_2\text{O})_{50}]$). This choice is supported by benchmark tests (see the Supplemental Material [67]) as well as prior calculations on small halide-water clusters [23], which show that for Cl^- , the valence ionizations are mostly coming from the halide. For F^- , on the other hand, there are important contributions from both the halogen and the waters (with ionization from the latter being lower in energy than from the halide), and because of this a second model was considered in which the nearest eight water molecules are also included in subsystem I [referred to as $([\text{F}(\text{H}_2\text{O})_8]^- @ (\text{H}_2\text{O})_{42})$].

The DFT-in-DFT v_{int} were obtained over 200 CMD snapshots with the PYADF scripting environment [69], which used the subsystem DFT implementation in the ADF code [70] and employed the scalar relativistic (SR) zero-order regular approximation (ZORA) Hamiltonian [71] and triple-zeta quality basis sets [72] with two polarization functions for all atoms. The nonadditive kinetic energy and exchange-correlation contributions to v_{int} were calculated with the Lembarki-Chermette (PW91k) [49] and Perdew-Burke-Ernzerhof (PBE) [73] density functionals, respectively. Unless otherwise noted,

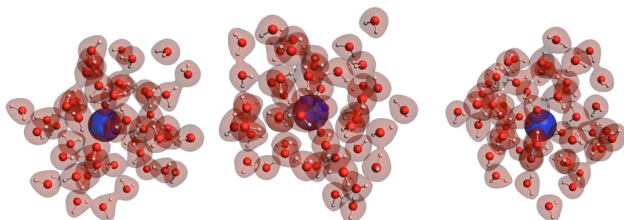


FIG. 1. Views along the (x, y, z) axes for a sample configuration of the CMD simulation for I^- . The (frozen) density for the water subsystem (n_{II}) is superimposed onto the structures [68].

all SR-ZORA DFT-in-DFT calculations reported use the same computational setup. The embedded EOM-IP-CCSD (EOM) calculations were performed over a subset of 100 CMD snapshots from the originally selected 200 snapshots (see the Supplemental Material [67]) with a development version (revisions e25ea49 and 7c8174a) [57] of the DIRAC electronic structure code [74], using the Dirac-Coulomb (DC) Hamiltonian [38,75] and uncontracted augmented triple-zeta quality [76–78] with two additional diffuse functions for the halogens, and the Dunning aug-cc-pVTZ sets [79] for oxygen and hydrogen. Because of constraints in computational resources for the $([\text{F}(\text{H}_2\text{O})_8]^- @ (\text{H}_2\text{O})_{42})$ partition, DFT-in-DFT calculations were performed exclusively using the DC Hamiltonian for F^- . In order to estimate the energies at the complete basis set (CBS) limit calculations with augmented quadruple-zeta basis sets were also performed: for F^- and Cl^- , it was computationally feasible to do so for all snapshots. For the other halides, this was not the case, and estimates for the CBS energies were obtained based on quadruple-zeta calculation on the bare halides. The data set comprising the DFT-in-DFT and CC-in-DFT calculation is available in the Zenodo repository [80].

We start by discussing the trends along the series for the BEs over the 100 snapshots, presented in Fig. 2 as histograms plots, with the area under each rectangle being proportional to the number of BEs found at each energy interval. There is very little variation on the BEs of the water subsystems (the yellow and brown rectangles) upon changing the halogen. For the halogens, one finds, first, the displacement of the first ionization energy peak, which in the presence of SOC corresponds to the $^2P_{3/2}$ halogen atom ground electronic states, towards lower energies as the halogen gets heavier. This results in a clear separation between the halogen and water peaks from Br^- onwards. One can also see, as expected from experiments and prior calculations, that irrespective of the treatment of the first solvation shell of F^- (here carried out only with DC SAOP calculations, as explained above), its electron BEs remain entangled with those of the water cluster. Second, the increasing separation between the $^2P_{1/2}$ and $^2P_{3/2}$ components of the halogen ground state is clearly seen, and for I^- the two peaks are clearly distinguishable from those of the water. It is interesting to note, however, that for At^- the SOC effect is so large (with a $^2P_{3/2} - ^2P_{1/2}$ splitting of ≈ 3.0 eV) that the $^2P_{1/2}$ peak ends up overlapping with that of water.

Table I summarizes the average BEs for the DFT-in-DFT and CC-in-DFT calculations of Fig. 2 (corresponding to peak maxima), while the experimental results are shown in Table II. By their comparison, one sees that, apart from the F^- case, the EOM results agree rather well with the experimental peak maxima for the halides, with differences of about 0.2 eV for Cl^- , and about 0.1 eV for Br^- and I^- . We attribute this relative improvement along the series to a

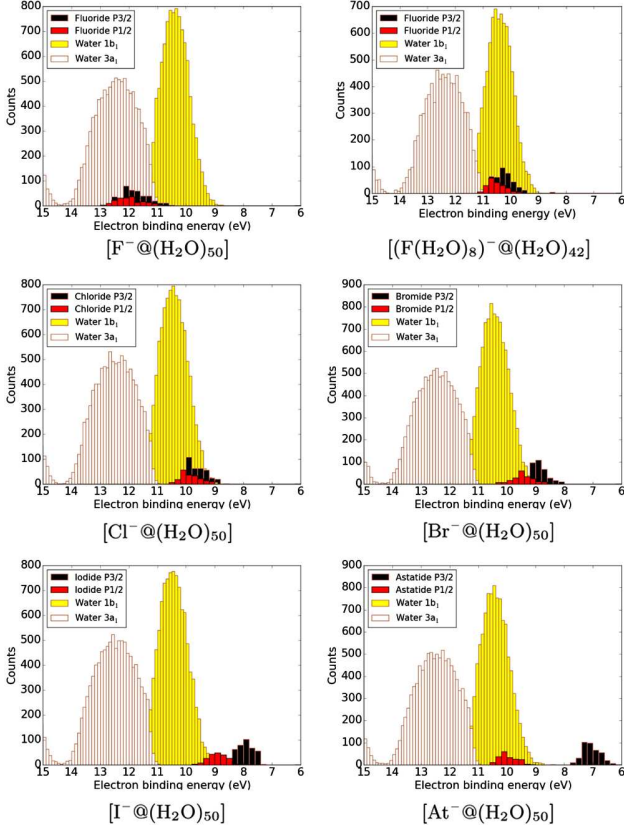


FIG. 2. Electron binding energies spectra for the $[X^- @ (H_2O)_{50}]$ systems over the 100 snapshots. Halides BEs obtained with triple-zeta basis sets from DC EOM [except for $[(F(H_2O)_8)^- @ (H_2O)_{42}]$ obtained with DC SAOP] [68].

decrease in entanglement between the halide and the surrounding water molecules as the halide gets heavier [66], which would make our simple embedding model better represent the physical system. For I^- , the only system for which Kurahashi *et al.* [81] provide the spin-orbit splitting of the 2P state, there is also very good agreement with the experiment for the ionization from the $^2P_{1/2}$ state.

Table I presents results for the halides obtained with triple-zeta base and CBS energy (for F^- and Cl^-) estimates (for Br^- to At^-). A comparison of EOM triple-zeta and CBS results indicates that the latter show a discrete improvement over the former, and in general make our results closer to the experiment. Furthermore, the SAOP results are in rather good agreement with the EOM values, with rather systematic differences on the order of 0.4 eV. This underscored the good performance of SAOP for BEs, especially in view of its modest computational cost, and validates our choice of employing SAOP for the DFT-in-DFT calculations. Additionally, as seen from Table III, SAOP and EOM yield good gas-phase BEs, meaning that the experimental halide BE shifts upon solvation are well reproduced. That said, our embedding model shows what appears to be a systematic underestimation of the water

TABLE I. Average electron binding energies (BE, in eV) for the spin-orbit coupled components of the P states of the hydrated halogens from EOM and SAOP (DC) calculations on the embedded halides with triple-zeta basis sets and the CBS values; and water droplet valence bands from SAOP (SR-ZORA) calculations for the $(H_2O)_{50}$ and $(H_2O)_{42}$ subsystems.

Species	Halogen				Water	
	BE $_{3/2}$		BE $_{1/2}$		BE $_{1b_1}$	BE $_{3a_1}$
	EOM	SAOP	EOM	SAOP	SAOP	
Triple-zeta bases						
F^-	11.8(5)	11.4(5)	12.0(5)	11.5(4)	10.4(5)	12.4(7)
$F(H_2O)_8^-$		10.3(4)		10.5(3)	10.4(5)	12.4(7)
Cl^-	9.7(3)	9.4(4)	9.9(3)	9.5(4)	10.4(5)	12.5(4)
Br^-	9.0(4)	8.7(3)	9.5(4)	9.2(4)	10.4(5)	12.5(4)
I^-	7.9(3)	7.8(3)	8.9(3)	8.6(3)	10.4(5)	12.5(4)
At^-	7.1(3)	7.0(3)	10.0(3)	9.5(3)	10.4(5)	12.5(4)
CBS (F^- , Cl^-) and CBS ^a (Br^- – At^-)						
F^-	11.9(5)	11.4(5)	12.1(5)	11.5(4)		
$F(H_2O)_8^-$		10.3(4)		10.5(3)		
Cl^-	9.9(3)	9.4(4)	10.1(3)	9.5(4)		
Br^-	9.0(4)	8.7(3)	9.5(4)	9.2(4)		
I^-	8.0(3)	7.8(3)	9.0(3)	8.6(3)		
At^-	7.1(3)	7.0(3)	10.1(3)	9.5(3)		

^aEstimates from single quadruple-zeta calculations.

spectra, by roughly 1 eV for the b_1 and a_1 peaks. Part of this discrepancy should originate from using SAOP rather than EOM energies (if errors follow those for the halides discussed above, up to 0.4–0.5 eV). We believe that the other major source of errors is the discrete size of the droplets used since the experimental results are for bulk water, and we intend to investigate this issue in a subsequent publication.

For Cl^- , a comparison to prior theoretical results can be made to the G_0W_0 calculations (without SOC) of Gaiduk *et al.* [28], shown in Table IV, for which the most sophisticated calculation using the self-consistent hybrid (sc-hybrid) density functional places the peak position at 9.89 eV. This is higher than the experimental results by a

TABLE II. Experimental electron binding energies (BE, in eV) for the spin-orbit coupled components of the P states of the solvated halide and bulk water valence bands from (a) Kurahashi *et al.* [81], and (b) Winter *et al.* [21].

Species	Halogen		Water			
	BE $_p$		BE $_{1b_1}$		BE $_{3a_1}$	
	(a)	(b)	(a)	(b)	(a)	(b)
F^-	9.8					
Cl^-	9.5(2)	9.60(7)				
Br^-	9.00(7)	8.80(6)				
		8.1(1)				
I^-	8.03(6) ^a	7.7(2) ^a	11.31(4)	11.16(4)	13.78(7) ^b	13.50(10)
	8.96(7) ^c	8.8(2) ^c				

^a $\Omega = 3/2$.

^bAverage value of the $3a_1$ H and $3a_1$ L bands.

^c $\Omega = 1/2$.

TABLE III. Gas-phase electron binding energies (BE, in eV) for the halides (DC) and the water molecule (SR-ZORA, PBE optimized geometry).

Species	SAOP		EOM		Expt.	
	Triple-zeta	CBS	Triple-zeta	CBS		
F ⁻	BE _{3/2}	3.16	3.16	3.32	3.45	3.40 [82,83]
Cl ⁻	BE _{3/2}	3.41	3.41	3.59	3.77	3.62 [84,85]
Br ⁻	BE _{3/2}	3.23	3.23	3.40	3.48	3.37 [86,87]
I ⁻	BE _{3/2}	3.02	3.02	3.12	3.19	3.06 [88]
At ⁻	BE _{3/2}	2.48	2.48	2.41	2.55	2.40 ^a [89]
H ₂ O	BE _{1b₁}	12.33				12.62 [90]

^aCCSD(T).

little over 0.3 eV. It is also higher than the EOM calculations, even if it is compared to our ²P term value of 9.76 eV. The G_0W_0 -sc-hybrid calculations show very good agreement with experiment for the water peaks, though a comparison to our results would be somewhat biased since the G_0W_0 ones are made for a bulk liquid, and ours are not. It is important to note that the G_0W_0 results do not show very good agreement with the experimental BEs if less sophisticated functionals such as PBE and PBE0 are used—in fact, the DC SAOP results are of slightly better quality than those.

Another relevant comparison is with the electron propagator calculations of Dolgounitcheva *et al.* [23], performed for microsolvated clusters of F⁻ and Cl⁻ and including the effect of outer solvation shells via the PCM. For Cl⁻, the propagator results agree well with each other but are nevertheless 0.7–1 eV higher than the experiment, whereas our results are not more than 0.2 eV higher. For the first ionization of F⁻ to which there are significant contributions from Dyson orbitals on F, the propagator results are closer to each other but again quite far from the experiment. If part

 TABLE IV. Selected theoretical electron binding energies (BE, in eV) from the literatures for solvated F⁻ and Cl⁻ using the G_0W_0 [28] approach, and the Outer-Valence Greens Function (OVGF), Partial third order (P3) and renormalized Partial third order (P3+) propagator approaches combined with PCM (polarizable continuum model) [23] or explicit solvation (PC, point-charge embedding) [91].

Method	Cl ⁻	F ⁻
G_0W_0 -PBE [28]	8.76	
G_0W_0 -PBE0 [28]	9.43	
G_0W_0 -RSH [28]	9.86	
G_0W_0 -sc-hybrid [28]	9.89	
OVGF-PCM [23]	10.53	10.70
P3-PCM [23]	10.32	12.21
P3 + -PCM [23]	10.29	12.02
P3-6H ₂ O [91]	6.95	
P3-6H ₂ O + 60H ₂ O(PC) [91]	9.41	

of the discrepancy comes from differences in treatment of electron correlation between the propagators and EOM (or SAOP) and basis set effects (bases smaller than ours were used), the most significant contribution should be due to the explicit inclusion of the outer solvation shells in our calculations. The importance of this effect is seen in the P3 calculations of Canuto *et al.* [91], which, when considering outer-shell effects via point-charge embedding, recover nearly 2.5 eV with respect to the microsolvated ion, showing an agreement to experiment similar to SAOP.

In conclusion, our results show that FDE is a viable method for obtaining quantitatively accurate electron binding energies (and with that simulate PE spectra) in the valence region for species in solution. For systems not undergoing chemical changes, the combination of CC-in-DFT calculations with CMD simulations with polarizable force fields can yield results which rival much more sophisticated simulation approaches, but at a much smaller computational cost (the embedded EOM calculations take about a day per snapshot on four cores for At⁻, the most expensive calculations). In this sense, the SAOP model potential appears to be a rather interesting alternative to more computationally expensive functionals by itself or, eventually, being combined with many-body treatments based on the *GW* method. Finally, our work was based on droplet simulations, which can be interesting to investigate systems made up by a relatively small amount of water molecules, though monitoring droplet size effects on such properties and their convergence towards the bulk requires further investigation. The FDE calculations are, however, completely agnostic to the nature of the procedure employed to obtain the structures, and they can be equally applied to snapshots from standard (or FDE-based [92]) CPMD calculations (whenever DFT-based interaction potentials are sufficiently accurate [93]) or static band-structure FDE calculations [94] that naturally describe long-range interactions in extended systems.

We acknowledge support from the Labex CaPPA (Chemical and Physical Properties of the Atmosphere, Contract No. ANR-11-LABX-0005-01), CPER CLIMIBIO (European Regional Development Fund, Hauts de France council, French Ministry of Higher Education and Research), the CNRS Institute of Physics (PICS Grant No. 6386), and French national supercomputing facilities (Grant No. DARI A0030801859).

*Corresponding author.

andre.gomes@univ-lille.fr

[†]Present address: Department of Chemistry, University of Michigan, 930 North University, Ann Arbor, Michigan 48109-1055, USA.

[1] J. L. Bahr, *Contemp. Phys.* **14**, 329 (1973).

[2] P. D. Dau, J. Su, H.-T. Liu, D.-L. Huang, J. Li, and L.-S. Wang, *J. Chem. Phys.* **137**, 064315 (2012).

- [3] W.-L. Li, J. Su, T. Jian, G. V. Lopez, H.-S. Hu, G.-J. Cao, J. Li, and L.-S. Wang, *J. Chem. Phys.* **140**, 094306 (2014).
- [4] J. Su, P. D. Dau, H.-T. Liu, D.-L. Huang, F. Wei, W. H. E. Schwarz, J. Li, and L.-S. Wang, *J. Chem. Phys.* **142**, 134308 (2015).
- [5] R. Seidel, B. Winter, and S. E. Bradforth, *Annu. Rev. Phys. Chem.* **67**, 283 (2016).
- [6] T. A. Pham, M. Govoni, R. Seidel, S. E. Bradforth, E. Schwegler, and G. Galli, *Sci. Adv.* **3**, e1603210 (2017).
- [7] P. S. Bagus, E. S. Ilton, and C. J. Nelin, *Surf. Sci. Rep.* **68**, 273 (2013).
- [8] L. Trotochaud, A. R. Head, O. Karslıoğlu, L. Kyhl, and H. Bluhm, *J. Phys. Condens. Matter* **29**, 053002 (2017).
- [9] A. Knop-Gericke, V. Pfeifer, J.-J. Velasco-Velez, T. Jones, R. Arrigo, M. Hävecker, and R. Schlögl, *J. Electron Spectrosc. Relat. Phenom.* **221**, 10 (2017).
- [10] X. Kong, A. Waldner, F. Orlando, L. Artiglia, T. Huthwelker, M. Ammann, and T. Bartels-Rausch, *J. Phys. Chem. Lett.* **8**, 4757 (2017).
- [11] T. Bartels-Rausch, F. Orlando, X. Kong, L. Artiglia, and M. Ammann, *ACS Earth Space Chem.* **1**, 572 (2017).
- [12] A. A. Raheem, M. Wilke, M. Borgwardt, N. Engel, S. I. Bokarev, G. Grell, S. G. Aziz, O. Kühn, I. Y. Kiyani, C. Merschjann, and E. F. Aziz, *J. Biomol. Struct. Dyn.* **4**, 044031 (2017).
- [13] B. P. Tsal, T. Baer, A. S. Werner, and S. F. Lin, *J. Phys. Chem.* **79**, 570 (1975).
- [14] M. K. Gilles, M. L. Polak, and W. C. Lineberger, *J. Chem. Phys.* **96**, 8012 (1992).
- [15] A. F. Lago, J. P. Kercher, A. Bödi, B. Sztáray, B. Müller, D. Wurzelmann, and T. Baer, *J. Phys. Chem. A* **109**, 1802 (2005).
- [16] W. R. Simpson, S. S. Brown, A. Saiz-Lopez, J. A. Thornton, and R. von Glasow, *Chem. Rev.* **115**, 4035 (2015).
- [17] A. Saiz-Lopez, J. M. C. Plane, A. R. Baker, L. J. Carpenter, R. von Glasow, J. C. Gómez Martín, G. McFiggans, and R. W. Saunders, *Chem. Rev.* **112**, 1773 (2012).
- [18] J. C. Gómez Martín, O. Gálvez, M. T. Baeza-Romero, T. Ingham, J. M. C. Plane, and M. A. Blitz, *Phys. Chem. Chem. Phys.* **15**, 15612 (2013).
- [19] L. J. Carpenter and P. D. Nightingale, *Chem. Rev.* **115**, 4015 (2015).
- [20] G. Markovich, S. Pollack, R. Giniger, and O. Cheshnovsky, *J. Chem. Phys.* **101**, 9344 (1994).
- [21] B. Winter, R. Weber, I. V. Hertel, M. Faubel, P. Jungwirth, E. C. Brown, and S. E. Bradforth, *J. Am. Chem. Soc.* **127**, 7203 (2005).
- [22] A. K. Pathak, T. Mukherjee, and D. K. Maity, *Chem. Phys. Lett.* **454**, 17 (2008).
- [23] O. Dolgounitcheva, V. G. Zakrzewski, and J. V. Ortiz, *Int. J. Quantum Chem.* **112**, 3840 (2012).
- [24] Z. He, G. Feng, B. Yang, L. Yang, C.-W. Liu, H.-G. Xu, X.-L. Xu, W.-J. Zheng, and Y. Q. Gao, *J. Chem. Phys.* **148**, 222839 (2018).
- [25] M. P. Coons and J. M. Herbert, *J. Chem. Phys.* **148**, 222834 (2018).
- [26] J. Hutter, *Comput. Mol. Sci.* **2**, 604 (2012).
- [27] C. Zhang, T. A. Pham, F. Gygi, and G. Galli, *J. Chem. Phys.* **138**, 181102 (2013).
- [28] A. P. Gaiduk, M. Govoni, R. Seidel, J. H. Skone, B. Winter, and G. Galli, *J. Am. Chem. Soc.* **138**, 6912 (2016).
- [29] A. P. Gaiduk and G. Galli, *J. Phys. Chem. Lett.* **8**, 1496 (2017).
- [30] A. P. Gaiduk, J. Gustafson, F. Gygi, and G. Galli, *J. Phys. Chem. Lett.* **9**, 3068 (2018).
- [31] A. P. Gaiduk, T. A. Pham, M. Govoni, F. Paesani, and G. Galli, *Nat. Commun.* **9**, 247 (2018).
- [32] X. Blase, P. Boulanger, F. Bruneval, M. Fernandez-Serra, and I. Duchemin, *J. Chem. Phys.* **144**, 034109 (2016); **145**, 169901 (2016).
- [33] M. F. Lange and T. C. Berkelbach, *J. Chem. Theory Comput.* **14**, 4224 (2018).
- [34] R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- [35] R. J. Bartlett, *Comput. Mol. Sci.* **2**, 126 (2012).
- [36] J. J. Goings, M. Caricato, M. J. Frisch, and X. Li, *J. Chem. Phys.* **141**, 164116 (2014).
- [37] A. K. Dutta, N. Vaval, and S. Pal, *Int. J. Quantum Chem.* **118**, e25594 (2018).
- [38] T. Saue, *Chem. Phys. Chem.* **12**, 3077 (2011).
- [39] P. Cortona, *Phys. Rev. B* **44**, 8454 (1991).
- [40] P. Cortona, *Phys. Rev. B* **46**, 2008 (1992).
- [41] T. A. Wesolowski and A. Warshel, *J. Phys. Chem.* **97**, 8050 (1993).
- [42] A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **108**, 222 (2012).
- [43] C. R. Jacob and J. Neugebauer, *Comput. Mol. Sci.* **4**, 325 (2014).
- [44] T. A. Wesolowski, S. Shedge, and X. Zhou, *Chem. Rev.* **115**, 5891 (2015).
- [45] Q. Sun and G. K.-L. Chan, *Acc. Chem. Res.* **49**, 2705 (2016).
- [46] O. Roncero, M. P. de Lara-Castells, P. Villarreal, F. Flores, J. Ortega, M. Paniagua, and A. Aguado, *J. Chem. Phys.* **129**, 184104 (2008).
- [47] C. R. Jacob, S. M. Beyhan, and L. Visscher, *J. Chem. Phys.* **126**, 234116 (2007).
- [48] P. Reinholdt, J. Kongsted, and J. M. H. Olsen, *J. Phys. Chem. Lett.* **8**, 5949 (2017).
- [49] A. Lembarki and H. Chermette, *Phys. Rev. A* **50**, 5328 (1994).
- [50] A. W. Götz, S. M. Beyhan, and L. Visscher, *J. Chem. Theory Comput.* **5**, 3161 (2009).
- [51] N. Govind, Y. A. Wang, and E. A. Carter, *J. Chem. Phys.* **110**, 7677 (1999).
- [52] T. A. Wesolowski, *Phys. Rev. A* **77**, 012504 (2008).
- [53] A. S. P. Gomes, C. R. Jacob, and L. Visscher, *Phys. Chem. Chem. Phys.* **10**, 5353 (2008).
- [54] S. Höfener, A. S. P. Gomes, and L. Visscher, *J. Chem. Phys.* **136**, 044104 (2012).
- [55] S. Höfener and L. Visscher, *J. Chem. Phys.* **137**, 204120 (2012).
- [56] S. Höfener, A. S. P. Gomes, and L. Visscher, *J. Chem. Phys.* **139**, 104106 (2013).
- [57] A. Shee, T. Saue, L. Visscher, and A. S. P. Gomes, *J. Chem. Phys.* **149**, 174113 (2018).
- [58] O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, *Chem. Phys. Lett.* **302**, 199 (1999).
- [59] P. Tecmer, A. S. P. Gomes, U. Ekström, and L. Visscher, *Phys. Chem. Chem. Phys.* **13**, 6249 (2011).

- [60] I. Grabowski, A. M. Teale, S. Śmiga, and R. J. Bartlett, *J. Chem. Phys.* **135**, 114111 (2011).
- [61] D. S. Ranasinghe, A. Perera, and R. J. Bartlett, *J. Chem. Phys.* **147**, 204103 (2017).
- [62] M. Masella, *Mol. Phys.* **104**, 415 (2006).
- [63] M. Masella, D. Borgis, and P. Cuniassé, *J. Comput. Chem.* **32**, 2664 (2011).
- [64] M. Masella, D. Borgis, and P. Cuniassé, *J. Comput. Chem.* **34**, 1112 (2013).
- [65] J. P. Coles and M. Masella, *J. Chem. Phys.* **142**, 024109 (2015).
- [66] F. Réal, A. S. P. Gomes, Y. O. Guerrero Martínez, T. Ayed, N. Galland, M. Masella, and V. Vallet, *J. Chem. Phys.* **144**, 124513 (2016).
- [67] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.121.266001> for atomic calculations, benchmarks calculations with FDE models, time-correlations of properties, basis set convergence tests.
- [68] Y. Bouchafra, A. Shee, F. Real, V. Vallet, and A. S. P. Gomes, <http://dx.doi.org/10.5281/zenodo.1477103>.
- [69] C. R. Jacob, S. M. Beyhan, R. E. Buló, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, *J. Comput. Chem.* **32**, 2328 (2011).
- [70] C. R. Jacob, J. Neugebauer, and L. Visscher, *J. Comput. Chem.* **29**, 1011 (2008).
- [71] E. van Lenthe, E. J. Baerends, and J. G. Snijders, *J. Chem. Phys.* **99**, 4597 (1993).
- [72] E. van Lenthe and E. J. Baerends, *J. Comput. Chem.* **24**, 1142 (2003).
- [73] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996); **78**, 1396 (1997).
- [74] L. Visscher *et al.* (DIRAC Collaboration), computer code DIRAC, 2017 release, 2017, <http://www.diracprogram.org>.
- [75] L. Visscher, *Theor. Chem. Acc.* **98**, 68 (1997).
- [76] K. G. Dyall, *Theor. Chem. Acc.* **108**, 335 (2002).
- [77] K. G. Dyall, *Theor. Chem. Acc.* **115**, 441 (2006).
- [78] K. G. Dyall, *Theor. Chem. Acc.* **135**, 128 (2016).
- [79] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [80] Y. Bouchafra, A. Shee, F. Real, V. Vallet, and A. S. P. Gomes, <http://dx.doi.org/10.5281/zenodo.1477004>.
- [81] N. Kurahashi, S. Karashima, Y. Tang, T. Horio, B. Abulimiti, Y.-I. Suzuki, Y. Ogi, M. Oura, and T. Suzuki, *J. Chem. Phys.* **140**, 174506 (2014).
- [82] H.-P. Popp, *Z. Naturforsch.* **22A**, 254 (1967).
- [83] R. Milstein and R. S. Berry, *J. Chem. Phys.* **55**, 4146 (1971).
- [84] G. Mück and H.-P. Popp, *Z. Naturforsch.* **23A**, 1213 (1968).
- [85] I. S. McDermid and C. R. Webster, *J. Phys. (Paris), Colloq.* **44**, C7-461 (1983).
- [86] R. S. Berry and C. W. Reimann, *J. Chem. Phys.* **38**, 1540 (1963).
- [87] H. Frank, M. Neiger, and H.-P. Popp, *Z. Naturforsch.* **25A**, 1617 (1970).
- [88] C. R. Webster, I. S. McDermid, and C. T. Rettner, *J. Chem. Phys.* **78**, 646 (1983).
- [89] A. Borschevsky, L. F. Pašteka, V. Pershina, E. Eliav, and U. Kaldor, *Phys. Rev. A* **91**, 020501 (2015).
- [90] J. E. Reutt, L. S. Wang, Y. T. Lee, and D. A. Shirley, *J. Chem. Phys.* **85**, 6928 (1986).
- [91] S. Canuto, K. Coutinho, B. J. C. Cabral, V. G. Zakrzewski, and J. V. Ortiz, *J. Chem. Phys.* **132**, 214507 (2010).
- [92] A. Genova, D. Ceresoli, A. Krishtal, O. Andreussi, R. A. DiStasio, Jr., and M. Pavanello, *Int. J. Quantum Chem.* **117**, e25401 (2017).
- [93] M. J. Gillan, D. Alfè, and A. Michaelides, *J. Chem. Phys.* **144**, 130901 (2016).
- [94] J. Tölle, A. S. P. Gomes, P. Ramos, and M. Pavanello, *Int. J. Quantum Chem.* **119**, e25801 (2018).

Supplementary Information
**Predictive simulations of ionization energies of solvated halide ions with relativistic
embedded Equation of Motion Coupled-Cluster Theory**

Yassine Bouchafra

*Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers,
Atomes et Molécules, F-59000 Lille, France Tel: +33-3-2043-4163*

Avijit Shee

*Department of Chemistry, University of Michigan,
930 N. University, Ann Arbor, MI 48109-1055, USA**

Florent Réal, Valérie Vallet, and André Severo Pereira Gomes[†]
*Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers,
Atomes et Molécules, F-59000 Lille, France; Tel: +33-3-2043-4163*

(Dated: December 4, 2018; Revised December 4, 2018)

TABLE I: Gas-phase atomic binding energies (BE, in eV) computed at the SAOP and EOM (DC) levels for triple-zeta, quadruple-zeta basis sets and at the CBS levels.

Species		SAOP			EOM		
		TA	QZ	CBS	TZ	QZ	CBS
F ⁻	BE _{3/2}	3.16	3.16	3.16	3.32	3.39	3.45
	BE _{1/2}	3.21	3.21	3.21	3.37	3.45	3.51
Cl ⁻	BE _{3/2}	3.41	3.41	3.41	3.59	3.69	3.77
	BE _{1/2}	3.51	3.51	3.50	3.70	3.81	3.89
Br ⁻	BE _{3/2}	3.23	3.23	3.23	3.40	3.45	3.48
	BE _{1/2}	3.65	3.65	3.65	3.89	3.93	3.96
I ⁻	BE _{3/2}	3.02	3.02	3.02	3.12	3.16	3.19
	BE _{1/2}	3.87	3.87	3.87	4.09	4.14	4.18
At ⁻	BE _{3/2}	2.48	2.48	2.48	2.41	2.49	2.55
	BE _{1/2}	5.05	5.05	5.05	5.35	5.44	5.51

ATOMIC CALCULATIONS

Complete Basis Set (CBS) values are calculated using the following formula, in which n is the basis set cardinal number:

$$E = E_{cbs} + \frac{A}{n^3} \quad (1)$$

Thus for two cardinal numbers $n_1 = 3$ (triple-zeta) and $n_2 = 4$ (quadruple-zeta), one can write

$$E(n_1) = E_{cbs} + \frac{A}{n_1^3} \quad (2)$$

$$E(n_2) = E_{cbs} + \frac{A}{n_2^3} \quad (3)$$

leading to the CBS extrapolated energy E_{cbs} :

$$E_{cbs} = \frac{E(n_1)n_1^3 - E(n_2)n_2^3}{n_1^3 - n_2^3} \quad (4)$$

CHOICE OF THE EMBEDDING MODEL FOR THE DFT-IN-DFT FNT CALCULATIONS

Molecular orbitals compositions in the $[\text{X}(\text{H}_2\text{O})_{50}]^-$ calculations

From the scalar-relativistic ZORA SAOP calculations on the $[\text{X}(\text{H}_2\text{O})_{50}]^-$ supermolecular systems we have drawn in Figure 1 the percentage contribution of the halide valence p orbitals into each molecular orbital. For all halides heavier than fluoride, the three highest occupied molecular orbitals correspond to the valence p halide orbitals, while for fluoride its 2p orbitals are immersed into the water valence manifold.

Influence of the embedding model on binding energies

The subsystem DFT approach [2–4] invokes calculation of the effective embedding potential, in order to take into account the effect of the environment on the embedded system. The simplest implementation of subsystem DFT is frozen density embedding (FDE) [4], in which the environment subsystem density $n_{\text{II}}(r)$ is kept frozen while the total energy is minimized with respect to changes in the other subsystem density $n_{\text{I}}(r)$. The minimization of the total energy with respect to the supermolecular density can be achieved through freeze-and-thaw (FnT) cycles (typically less than 20), where the roles of the subsystems I and II are iteratively interchanged. The relaxation steps are needed to account for the deformation/polarization of the subsystem’s densities, in the presence of charges.

For this study, we have explored several density partitioning for the case of iodide hydrated by 50 water molecules, referred to as the supermolecule (cf. Figure 2a). The simplest embedded model includes the halide anion as the

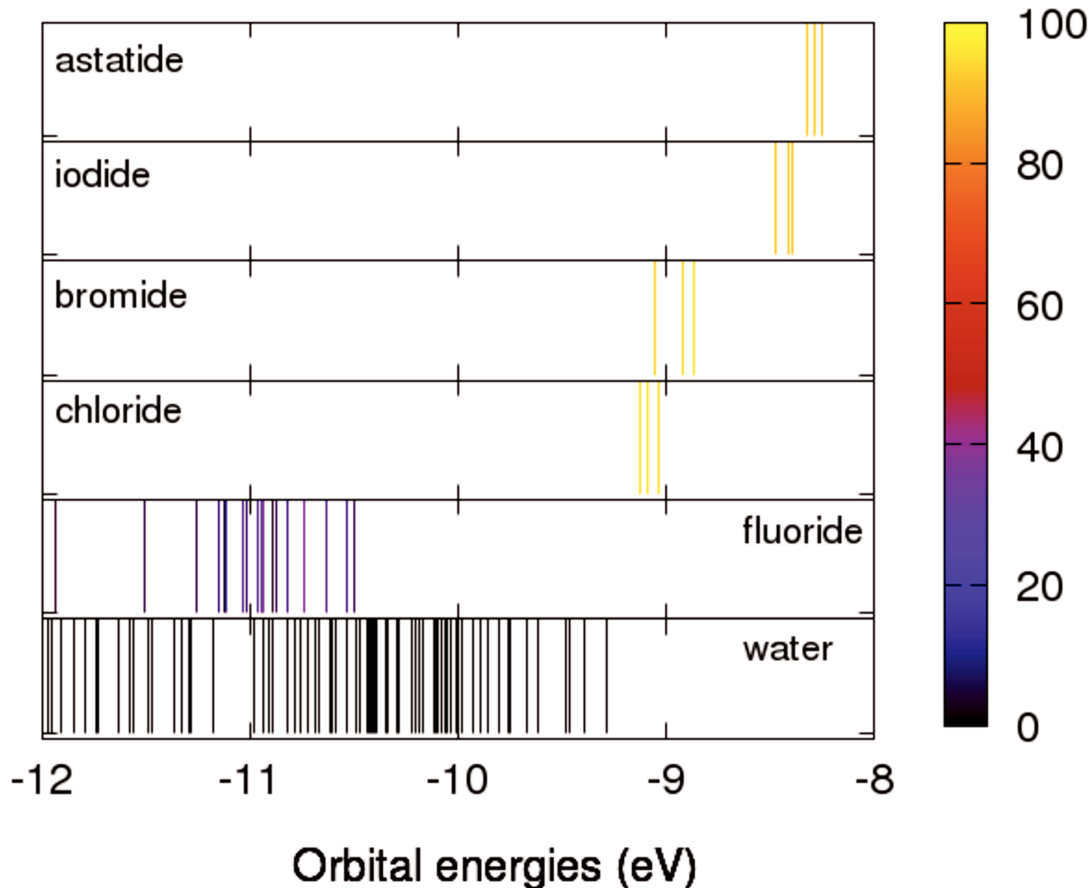


FIG. 1: Contributions of the halide valence p orbitals into the molecular orbitals of $[X(\text{H}_2\text{O})_{50}]^-$ hydrated cluster for a single snapshot, from 0 (black) up to 40% for fluoride, and 100% for the heavier halides. All calculations refer to scalar relativistic ZORA SAOP calculations. [1]

central subsystem, and the solvating water molecules in the environment as depicted in Figure 2b. A midway model places 10 water molecules constituting the iodide’s first hydration shell, the 40 other water molecules being part of the embedding, as illustrated by Figure 2c.

Figure 3 monitors the electron binding energy (SR-ZORA-SAOP) as a function of the density relaxation of the water molecules as a function of their distance from the iodide. For the $[\text{I}^-@(\text{H}_2\text{O})_{50}]$ model, the binding energy (BE) approaches the supermolecule value, when the 50 water molecules are relaxed as individual fragments (red dashed curve). The difference to the supermolecule reference is narrowed down to 0.2 eV when the 50 water molecules are relaxed altogether as a single fragment, certainly because the inaccuracies in the embedding potential due to the approximate kinetic energy functionals are minimized as the number of interacting fragment is kept small. It is noteworthy that including explicitly the first hydration shell of 10 water molecules in the central subsystem (blue curves), helps gaining only 0.1 eV with respect to the supermolecule at the expense of larger computational costs for the DFT-in-DFT FnT steps, and breaking the current computational limits of DC EOM-IP-CC calculations. This, together with the molecular orbital analysis supports the choice for all halogen heavier than fluoride of an embedding model composed of the halogen, embedded in a cluster of 50 water molecules.

Fluoride stands out, as its valence 2p orbitals strongly overlap with water $1b_1$ and $3a_2$ water orbitals. In the $[\text{F}^-@(\text{H}_2\text{O})_{50}]$ model, the FnT relaxation steps cannot account for this strong entanglement, leading to overshoot the average BE of fluoride (11.4(5) eV) by about 1.6 eV, with respect to the experimental value [5]. It thus deemed necessary to treat fluoride with its first hydration shell built from the 8 closed water molecule, embedded in the other 42 molecules, to reach good agreement (10.3(4) eV) with experiment.

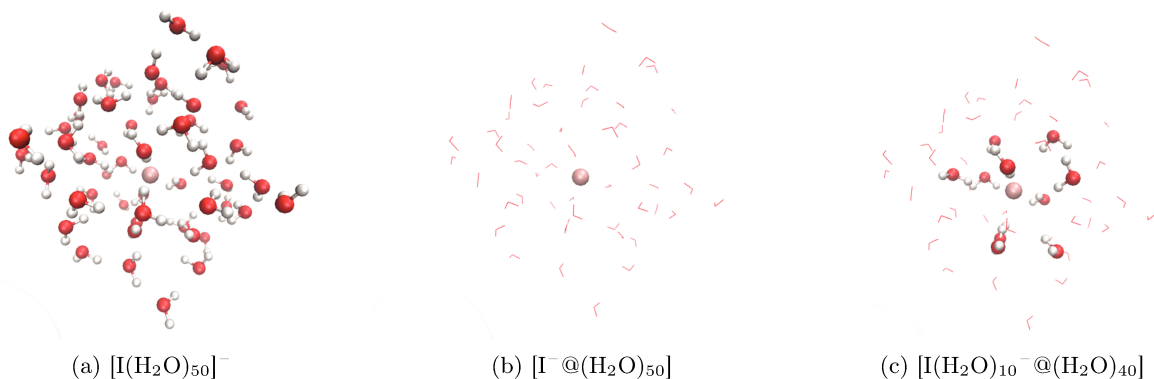


FIG. 2: Perspective views of the three different models, (a) the supermolecule, (b) iodide embedded in 50 water molecules; (c) $\text{I}(\text{H}_2\text{O})_{10}^-$ embedded in 40 water molecules. [1]

TIME AVERAGES AND CORRELATION TIMES

The dynamics of halide droplets with 50 water molecules were obtained from classical molecular dynamics (CMD) simulations performed at $T=300$ K with runs of 5 ns, with the POLARIS(MD) code [6–9] using the polarized force fields previously developed for the halides [10], and were found to reproduce the radial distributions obtained from periodic calculations. We extracted 200 snapshots sampled each 10 ps the last 1 ns of the CMD trajectories. The autocorrelation functions of the halide binding energies shown in Figure 5 die away instantaneously indicating that the 200 snapshots are temporarily uncorrelated, which justifies our present sampling choices. For all halides, the statistical BEs averages are converged with 100 snapshots only as seen from Figure 6. To preserve a statistically uncorrelated sampling, the statistical ensemble comprises 100 snapshots with a twice as long sampling interval of 20 ps.

SAOP DFT-IN-DFT RESULTS FOR $[\text{X}^-(\text{H}_2\text{O})_{50}]$ FNT MODELS

To optimize the computational cost of the accurate DC-EOM-IP-CC calculations, it is relevant to explore the effect of spin-orbit coupling on the binding energies at the DFT-in-DFT level. Figures 7 and 8 report the BEs obtained at the scalar relativistic and spin-orbit levels, respectively. Spin-orbit coupling (SOC) has no effect on the water BE bands, but it increasingly separates the halogen $^2P_{1/2}$ and $^2P_{3/2}$ bands. In astatide, SOC amounts to about 2.5 eV, making the $^2P_{1/2}$ peak overlap the water $1b_1$ band.

DC-EOM-IP-CC RESULTS, BASIS SET EFFECTS

Figure 9 shows the electron BEs spectra obtained with triple-zeta and quadruple-zeta basis sets for F^- and Cl^- . The increase of the basis set quality on the halide yields a slide increase of about 0.1 eV of the halide BEs. This energy shift is constant across the 100 snapshots as illustrated by Figure 11, and is of the same order of magnitude of that observed for bare halides, thus justifying that complete basis set (CBS) extrapolated values can be estimated from the sole atomic calculations reported in Table I. Figure 10 represents the extended PES for the $[\text{I}^-(\text{H}_2\text{O})_{50}]$ with triple-zeta basis sets in which water inner bands are shown.

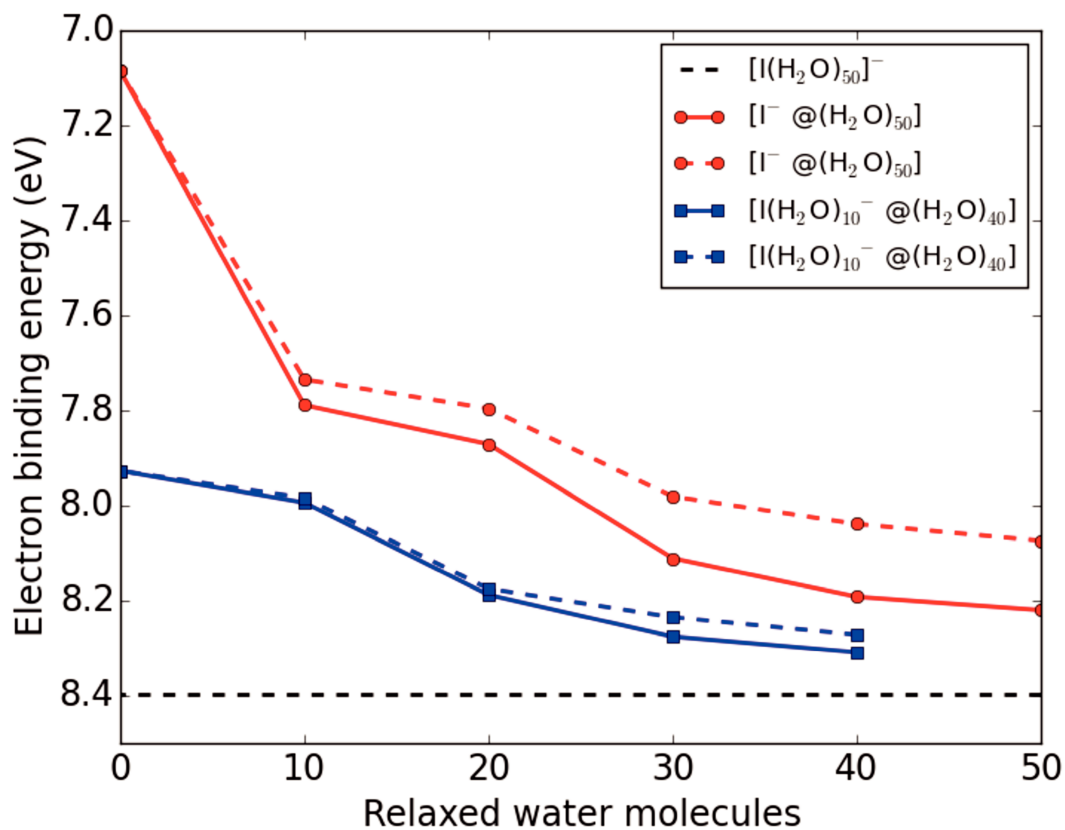


FIG. 3: Variation of the electron binding energy (SR-ZORA SAOP HOMO energy) for iodide embedded in water ($[I^- @ (H_2O)_{50}]$ (in red) and $[I(H_2O)_{10}^- @ (H_2O)_{40}]$ (in blue) as a function of the number of water molecules which have their density relaxed in the DFT-in-DFT procedure. The dashed lines correspond to FnT calculations in which the water molecules are considered individually, while for the plain lines correspond to the calculations in which the whole $(H_2O)_n$ water cluster's density is relaxed. The horizontal black dashed line represents the supermolecule calculation. [1]

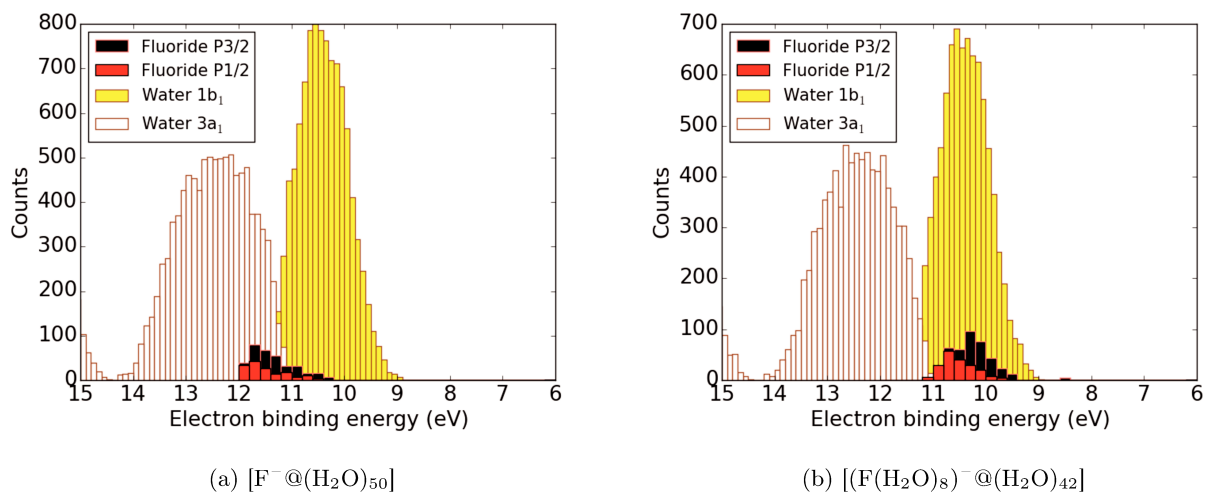


FIG. 4: Comparison of the fluoride binding energies in eV for the two models, $[\text{F}^- @ (\text{H}_2\text{O})_{50}]$ and $[(\text{F}(\text{H}_2\text{O})_8)^- @ (\text{H}_2\text{O})_{42}]$, computed at the DC SAOP level for 100 snapshots. [1]

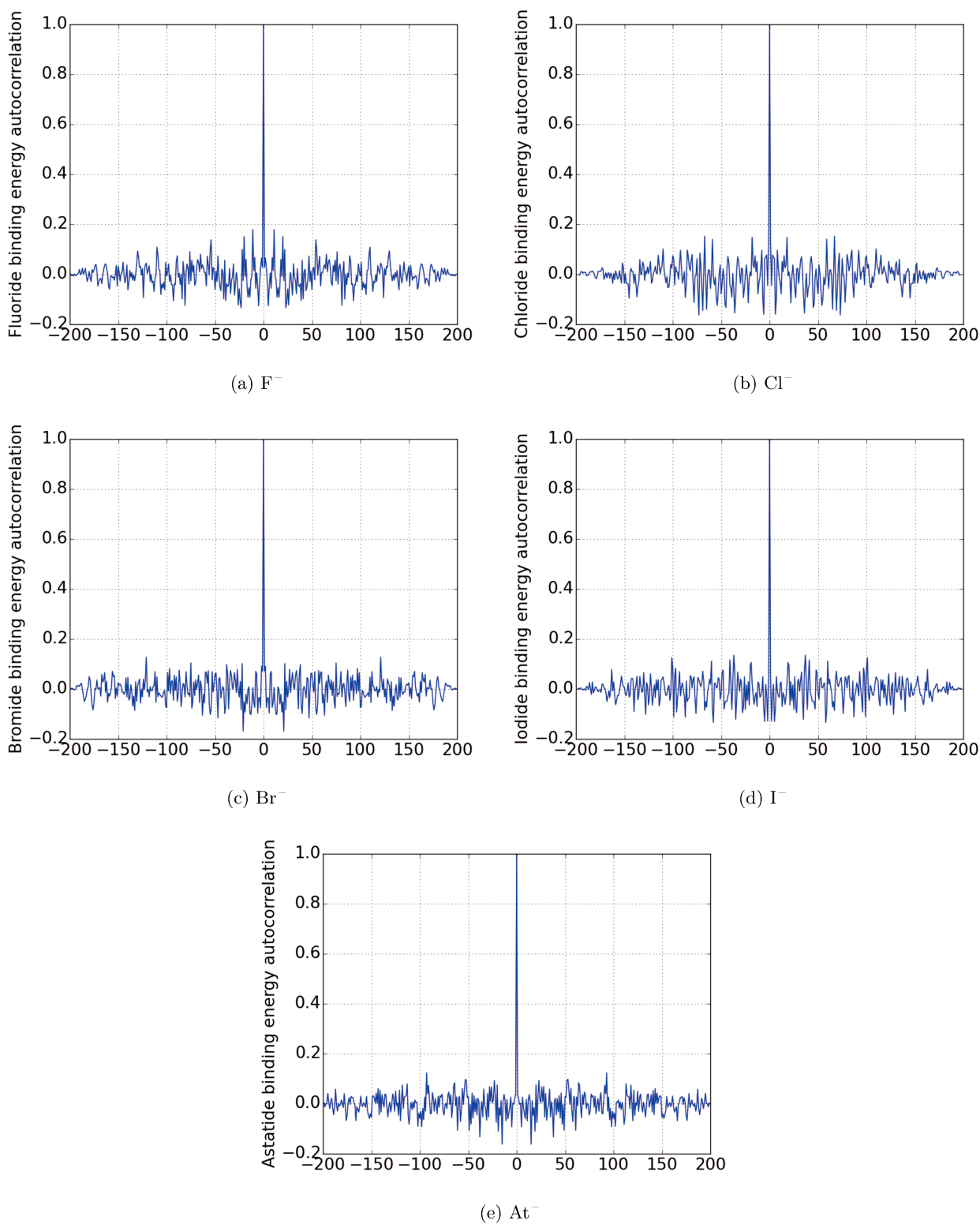


FIG. 5: Autocorrelation functions of the first halide binding energies computed at the DC SAOP level for the $[X^-(H_2O)_{50}]$ FnT models, computed for the 200 selected snapshots. [1]

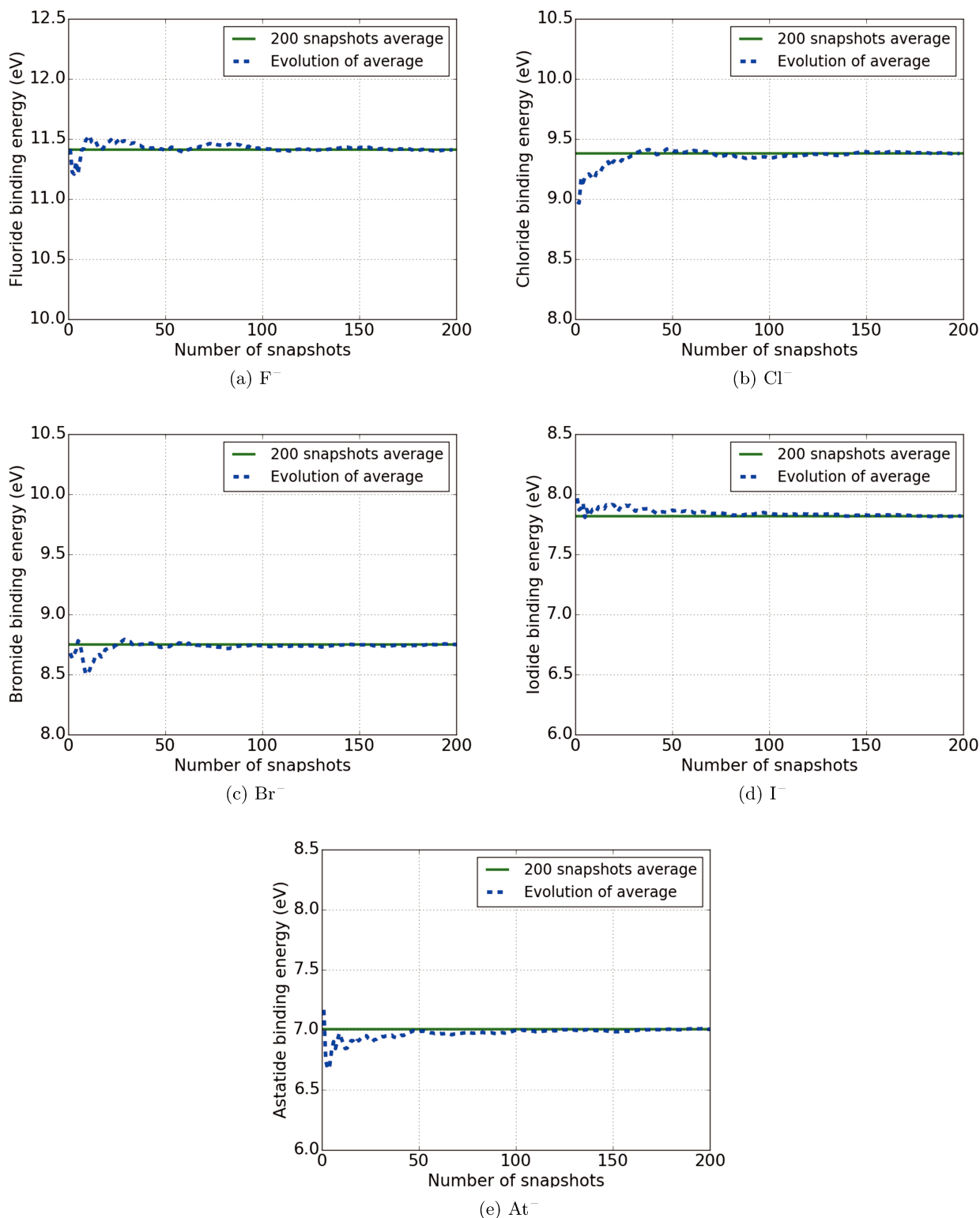


FIG. 6: Evolution of the average halide binding energies computed at the DC SAOP level for the $[X^-@(\text{H}_2\text{O})_{50}]$ FnT models, across the 200 selected snapshots. [1]

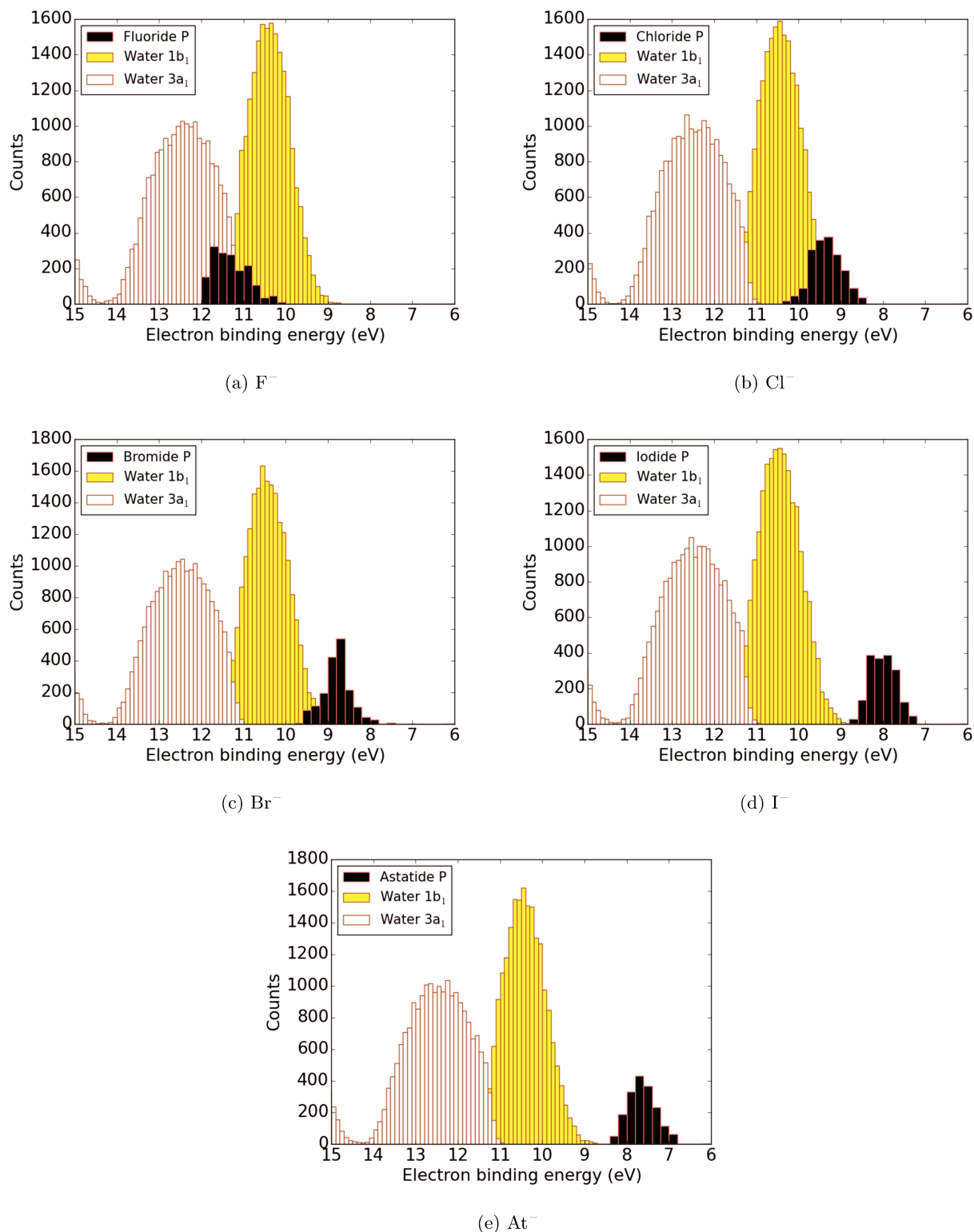


FIG. 7: Electron binding energies spectra in eV for the $[X^-(H_2O)_{50}]$ FnT systems over 200 snapshots, obtained from ZORA SAOP scalar relativistic calculations. [1]

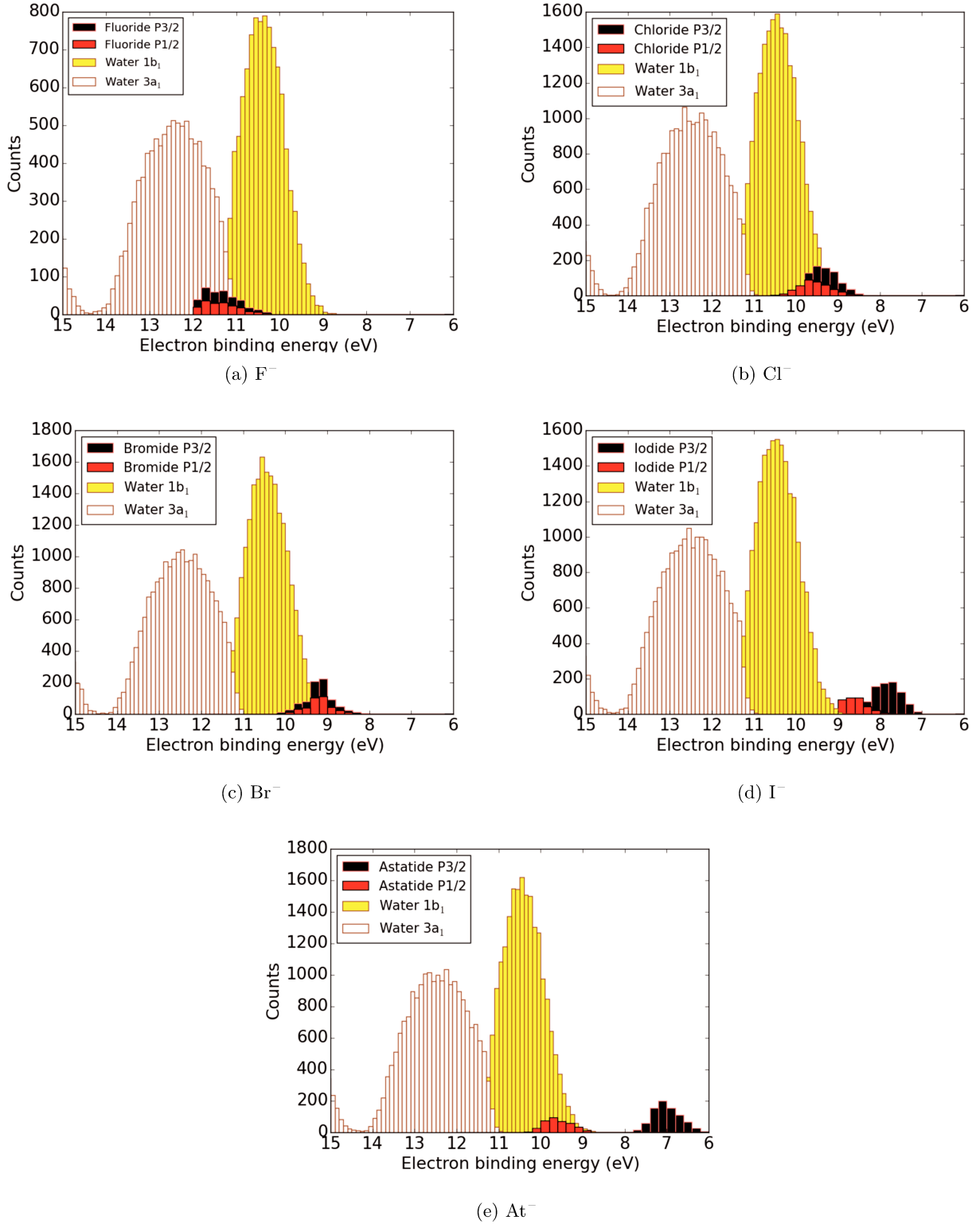


FIG. 8: Electron binding energies spectra in eV for the $[X^-(H_2O)_{50}]$ FnT systems over 200 snapshots, obtained from DC SAOP calculations with triple-zeta basis sets. [1]

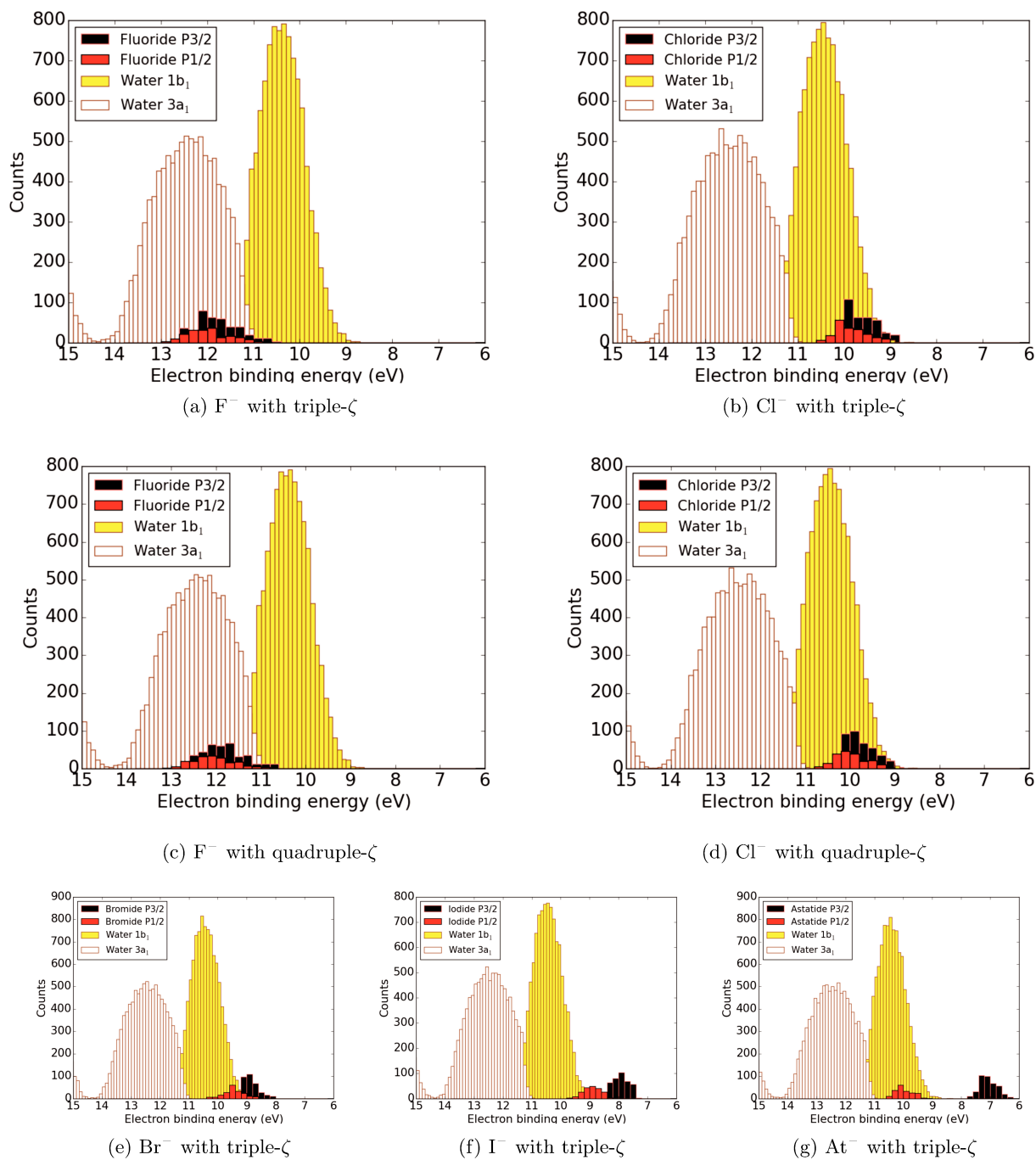


FIG. 9: Electron binding energies spectra in eV for the $[X^-(H_2O)_{50}]$ systems over 100 snapshots. Comparison of halides BEs obtained from DC EOM-IP-CCSD with triple-zeta basis sets, and quadruple-zeta basis sets for F^- and Cl^- . [1]

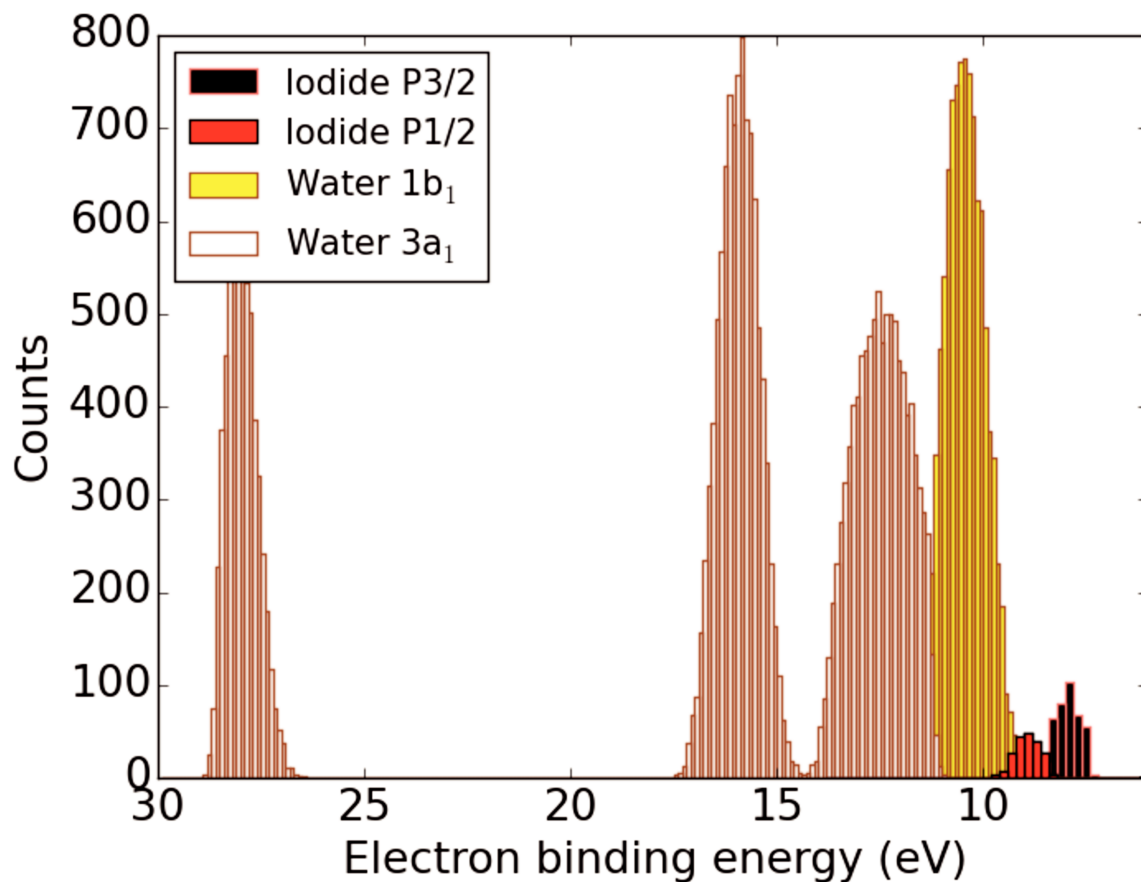


FIG. 10: Extended electron binding energies spectra in eV for the $[I^-@(\text{H}_2\text{O})_{50}]$ with triple-zeta basis sets. [1]

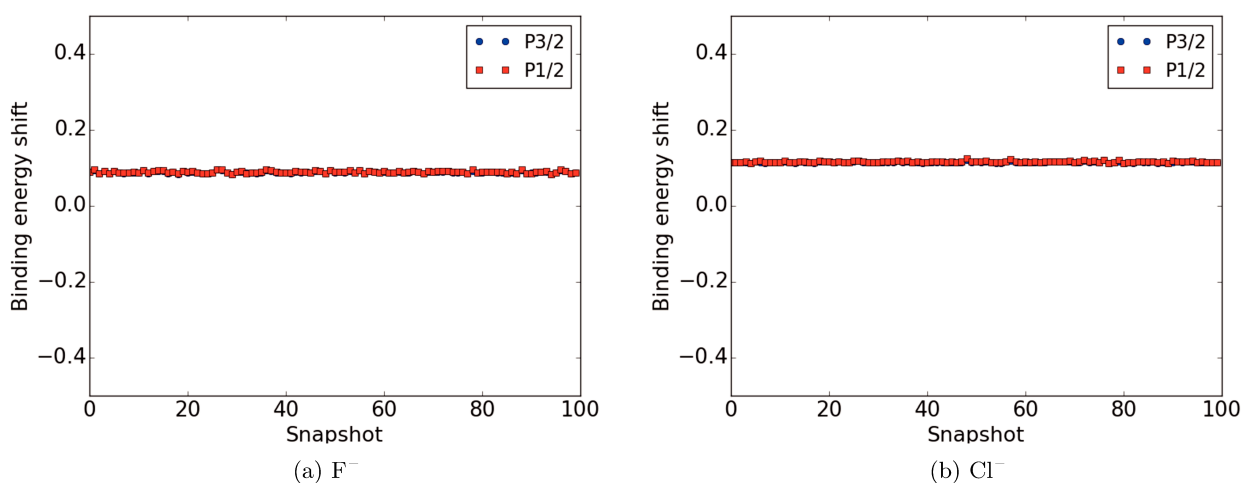


FIG. 11: Triple-zeta to quadruple-zeta shifts of the halide's DC EOM-IP-CCSD binding energies in eV, for 100 snapshots of $[F^-@(\text{H}_2\text{O})_{50}]$ and $[Cl^-@(\text{H}_2\text{O})_{50}]$.

-
- * Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers, Atomes et Molécules, F-59000 Lille, France Tel: +33-3-2043-4163
- † andre.gomes@univ-lille.fr (corresponding author)
- [1] Bouchafra, Y.; Shee, A.; Real, F.; Vallet, V.; Gomes, A. S. P. Predictive simulations of ionization energies of solvated halide ions with relativistic embedded Equation of Motion Coupled-Cluster Theory: Figures (Version 1.0). Zenodo <http://dx.doi.org/10.5281/zenodo.1477103>, 2018.
- [2] Cortona, P. Self-consistently determined properties of solids without band-structure calculations. *Phys. Rev. B* **1991**, *44*, 8454–8458.
- [3] Cortona, P. Direct determination of self-consistent total energies and charge densities of solids: A study of the cohesive properties of the alkali halides. *Phys. Rev. B* **1992**, *46*, 2008–2014.
- [4] Wesolowski, T. A.; Warshel, A. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **1993**, *97*, 8050–8053.
- [5] Kurahashi, N.; Karashima, S.; Tang, Y.; Horio, T.; Abulimiti, B.; Suzuki, Y.-I.; Ogi, Y.; Oura, M.; Suzuki, T. Photoelectron spectroscopy of aqueous solutions: Streaming potentials of NaX (X = Cl, Br, and I) solutions and electron binding energies of liquid water and X⁻. *J. Chem. Phys.* **2014**, *140*, 174506–10.
- [6] Masella, M. The multiple time step r-RESPA procedure and polarizable potentials based on induced dipole moments. *Mol. Phys.* **2006**, *104*, 415–428.
- [7] Masella, M.; Borgis, D.; Cuniasse, P. Combining a polarizable force-field and a coarse-grained polarizable solvent model. II. Accounting for hydrophobic effects. *J. Comput. Chem.* **2011**, *32*, 2664–2678.
- [8] Masella, M.; Borgis, D.; Cuniasse, P. A multiscale coarse-grained polarizable solvent model for handling long tail bulk electrostatics. *J. Comput. Chem.* **2013**, *34*, 1112–1124.
- [9] Coles, J. P.; Masella, M. The fast multipole method and point dipole moment polarizable force fields. *J. Chem. Phys.* **2015**, *142*, 024109–11.
- [10] Réal, F.; Gomes, A. S. P.; Guerrero Martínez, Y. O.; Ayed, T.; Galland, N.; Masella, M.; Vallet, V. Structural, dynamical, and transport properties of the hydrated halides: How do At⁻ bulk properties compare with those of the other halides, from F⁻ to I⁻? *J. Chem. Phys.* **2016**, *144*, 124513–13.

Chapter 4

Further investigations on the electronic structure of halides in water

The results presented in chapter 3 yield interesting insights. The binding energies (BE) of the water valence orbitals are found to be insensitive to the halide (TABLE I) type which confirms the previously observed results [132, 131]. However, the resulting water bands peaks tend to underestimate the experimental values [131, 133] by 0.7 to 1.3 eV for $1b_1$ and $3a_1$ binding energies (BEs). For the hydrated iodide, our multilevel model $I^-@(H_2O)_{50}$ based on droplet snapshots prepared using a polarisable force field (FF) developed in our group by Réal et al. [134], yield the results presented in Table (4.1):

	Iodide				Water			
	BE $_{3/2}$		BE $_{1/2}$		BE $_{1b_1}$		BE $_{3a_1}$	
This work	EOM	SAOP	EOM	SAOP	SAOP		SAOP	
	7.9(3)	7.8(3)	8.9(3)	8.6(3)	10.4(5)		12.5(4)	
Experiment	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
	8.03(6)	7.7(2)	8.96(7)	8.8(2)	11.31(4)	11.16(4)	13.78(7)	13.50(10)

Table 4.1: Electron binding energies (BE, in eV); This work : spin-orbit coupled components of the P states of the hydrated iodide, from EOM and SAOP (DC) calculations on the embedded iodide with triple-zeta basis sets, and water droplet valence bands from SAOP (SR-ZORA) calculations for the $I^-@(H_2O)_{50}$ system averaged over MD snapshots from Ref. [134]; Experiment : spin-orbit coupled components of the P states of the solvated iodide and bulk water valence bands from (a) Kurahashi et al. [133], and (b) Winter et al. [131].

The underestimation of SAOP (SR-ZORA) calculations for water droplet valence bands may have several origins. It could be due to the use of SAOP model functional to describe a large system, as it could come also from the droplet structures generated by the FF. But, given that we compare our $(H_2O)_{50}$ droplet results directly to experimental values deduced from bulk, the limited water droplet size can be a too small model. Moreover, Markovich et al. [135] recorded the photoelectron spectra of I^- solvated in water cluster $(H_2O)_n$, with n varying from 1 to 60. The vertical BE of the solvated I^- are used to extract the solvent electrostatic stabilisation energies (SE_n) of I^- defined as the difference between I^- BE in a water cluster and that of the isolated anion:

$$SE_n = BE_{([X^- @ (H_2O)_n])} - BE_{(X^-)}. \quad (4.1)$$

The plot of SE_n (Figure (4.1)) as a function of n can be fitted by a $(n + 2)^{-1/3}$ linear function:

$$SE_n = SE(\infty) - 5.76 \times (n + 2)^{-1/3}, \quad (4.2)$$

where $SE(\infty)$ is the bulk stabilisation energy. Figure (4.1) suggests that the bulk stabilisation energy is underestimated by 1.26 eV for water cluster, thus suggesting that we need larger water droplets to approach the bulk value.

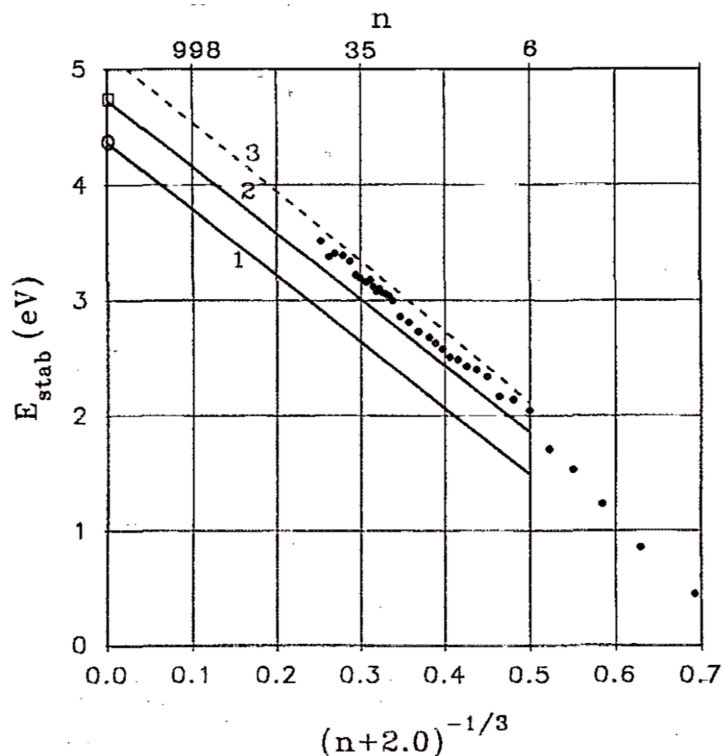


Figure 4.1: The stabilisation energies of hydrated iodide as a function of $(n + 2)^{-1/3}$ where n is the hydrating water molecules number [135]. The two solid lines (1,2) represent the centrally solvated anion according to the electrostatic model presented in Ref. ([136]) and each one of them extrapolates to a different experimental value for the stabilisation energy from Refs. ([136, 135]).

In the next section, and as the water valence orbitals are found to be insensitive to the halide type, we choose to test the effects of the water droplet size on the Cl^- system, since the chloride is much smaller than the iodide allowing a faster computation.

4.1 Droplet size effects on the water ionised bands

In order to probe the size effect we proceed to perform a comparison between the value of water BE of the same system with two different anions concentrations, a droplet of chloride solvated by 50 water molecules ($Cl^- @ (H_2O)_{50}$) and a largest droplet consisting of chloride surrounded by 200 water molecules ($Cl^- @ (H_2O)_{200}$) (see Figure 4.2).

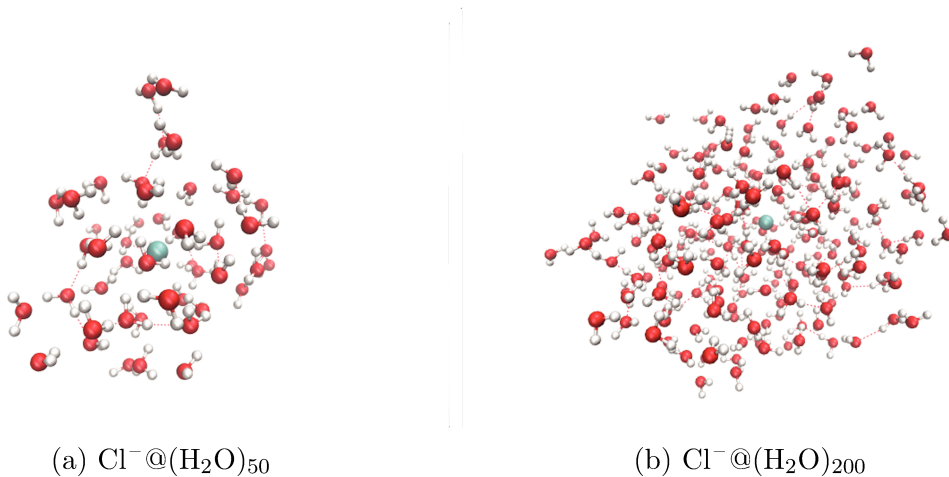


Figure 4.2: Snapshots from classical molecular dynamic of configurations centred on the chloride anion with different numbers of water molecules

As ($\text{Cl}^- @ (\text{H}_2\text{O})_{200}$) is a large system and we want to perform calculations over 100 snapshots, we propose to treat the water molecules not as a single subsystem, as done previously, but as a collection of 200 interacting subsystems (that we refer to, here, as a separate representation). In Table (4.2), we evaluate the effect on Cl^- BE of fragmenting a 50-water cluster, in order to provide an estimation of how this change in the treatment of the waters will affect the BEs.

		binding energy/eV		
orbital		$\text{Cl}^- @ (\text{H}_2\text{O})_{50}^{tog}$	$\text{Cl}^- @ (\text{H}_2\text{O})_{50}^{sep}$	Difference/eV
		Ave \pm SD	Ave \pm SD	
Chloride	$3p$	9.47 ± 0.34	9.64 ± 0.42	-0.17
	$3s$	21.05 ± 0.34	21.22 ± 0.42	-0.17
	$2p$	189.01 ± 0.33	189.19 ± 0.41	-0.18
	$2s$	249.95 ± 0.33	250.14 ± 0.41	-0.19
	$1s$	2750.87 ± 0.32	2751.06 ± 0.41	-0.19
Water	$1b_1$	10.43 ± 0.46	10.83 ± 0.80	-0.40
	$3a_1$	12.49 ± 0.67	12.67 ± 0.89	-0.18
	$1b_2$	15.94 ± 0.48	16.13 ± 0.85	-0.19
	$2a_1$	NP	28.53 ± 0.85	-

Table 4.2: Average (Ave) electron binding energies (BE, in eV) of the hydrated chloride from SAOP (SR-ZORA) calculations over 100 snapshots with triple-zeta basis and effect of using separated representation ($\text{H}_2\text{O})_{50}^{sep}$ rather than together one ($\text{H}_2\text{O})_{50}^{tog}$; SD refers to standard deviation and the water bands are labeled following the convention for a C_{2v} symmetry H_2O molecule.

All Cl^- ionisation bands are shifted up by 0.2 eV, so as the water inner bands ($3a_1$ and $1b_2$); the shift is larger (0.4 eV) for the outer water $1b_1$ band. It is noticeable also that within the together representation of water, the result photo-electron spectra is closer, qualitatively to experimentally observed result (see Figure (3.1b)). Effectively, the standard deviation of water BEs are almost equal in the case of the separated representation which results in analogous band shapes, while the $3a_1$ band is wider than the others.

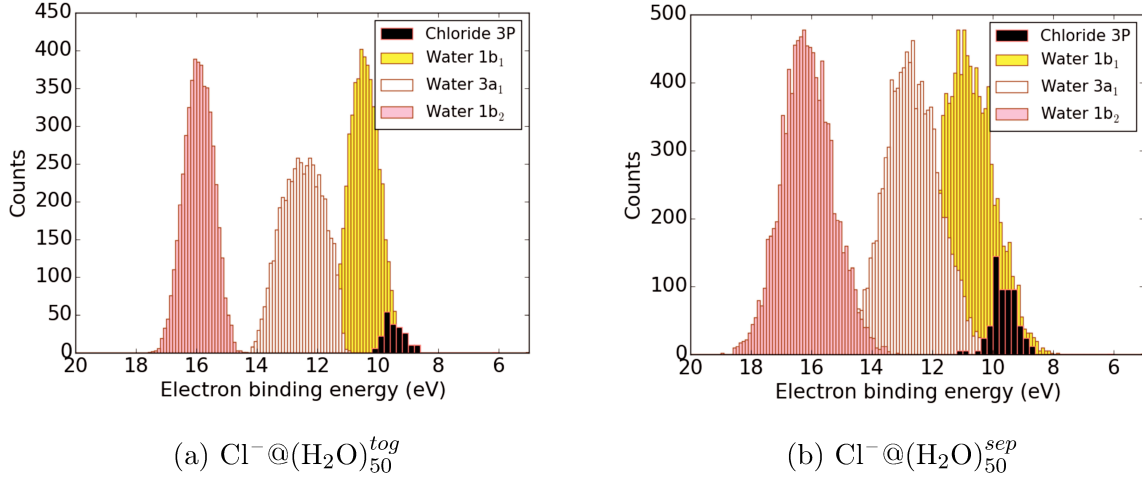


Figure 4.3: Electron binding energies spectra for the $[\text{Cl}^- @ (\text{H}_2\text{O})_{50}^{\text{tog}}]$ and $[\text{Cl}^- @ (\text{H}_2\text{O})_{200}^{\text{sep}}]$ systems from SAOP (SR-ZORA) calculated over 100 snapshots with triple-zeta basis sets.

After estimating the error done by using the separated representation, we keep using SAOP (SR-ZORA) calculation for the $\text{Cl}^- @ (\text{H}_2\text{O})_{200}$ system. The corresponding results compared to the $\text{Cl}^- @ (\text{H}_2\text{O})_{50}$, with the applied corrections, are illustrated in Figure (4.4), shows that the BE values show better agreement with experimental results [133, 131].

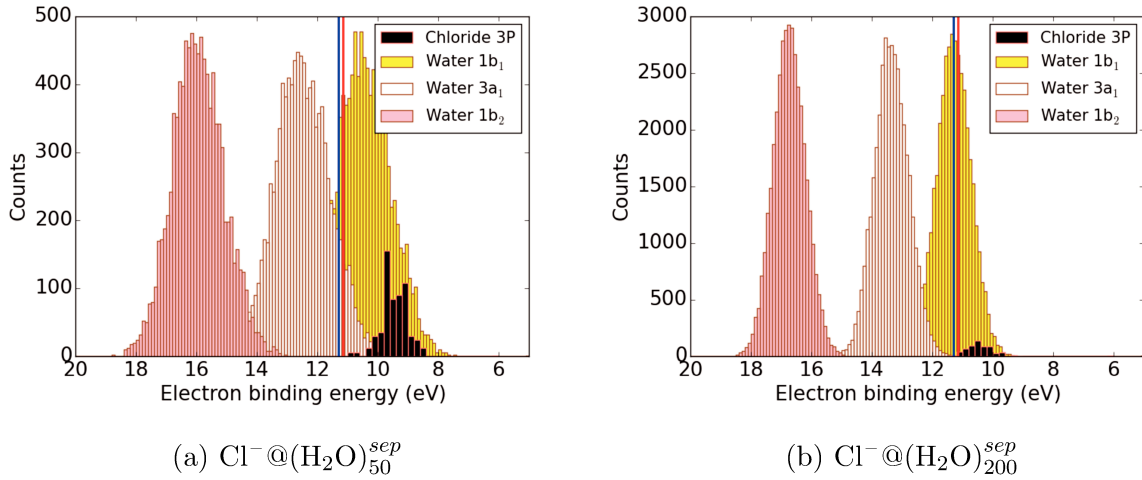


Figure 4.4: (Corrected) Electron binding energies spectra for the $[\text{Cl}^- @ (\text{H}_2\text{O})_{50}]$ and $[\text{Cl}^- @ (\text{H}_2\text{O})_{200}]$ from SAOP (SR-ZORA) calculated over 100 snapshots with triple-zeta basis, compared to experimental results, 11.31 eV from Kurahashi et al. [133] and 11.16 eV from Winter et al. [131].

Effectively, as the band energies values undergo shifts between 0.79 and 0.93 eV (see Table (4.3)) makes, after subtracting the correction value of 0.4 eV from the higher water BE (11.68 eV), we obtain 11.28 eV, a value which lies between the two experimental values of 11.16 eV [131] and 11.31 eV [133]. In addition the corrected $3a_1$ value of 13.42 eV ($13.60 - 0.18$) become closer to what is observed experimentally (13.50 [131] and 13.78 eV [133]). However, as shown by Markovich et al. we ideally would need larger water droplets to approach open the question the bulk values. An alternative to increasing the droplet size is to use an embedding

		binding energy/eV		
	orbital	Cl ⁻ @(H ₂ O) ₂₀₀		Difference/eV
		Ave	SD	to Cl ⁻ @(H ₂ O) ₅₀
chloride	3 <i>p</i>	10.63	0.38	0.99
	3 <i>s</i>	22.21	0.38	0.99
	2 <i>p</i>	190.18	0.37	0.99
	2 <i>s</i>	251.11	0.32	0.98
	1 <i>s</i>	2752.03	0.32	0.97
water	1 <i>b</i> ₁	11.68	0.55	0.85
	3 <i>a</i> ₁	13.60	0.56	0.93
	1 <i>b</i> ₂	16.92	0.54	0.79
	2 <i>a</i> ₁	29.34	0.54	0.81

Table 4.3: Average (Ave) BE (eV) of the hydrated chloride Cl⁻@(H₂O)₂₀₀ from SAOP (SR-ZORA) calculations over 100 snapshots with triple-zeta basis and effect of the water droplet size from [H₂O]₅₀ to [H₂O]₂₀₀.

potential computed for Cl⁻ in bulk water, as recently implemented in the embedding Quantum Espresso (eQE) package.

The results shows also that the chloride BEs are also tremendously shifted by (1 eV) making the droplets with 50 water molecules insufficient to capture the bulk SE(∞). This overestimation of chloride BE opens the question of the correctness of chloride-water interaction description within the FF [134] used in this study.

4.2 Force Field effects on the water and Halide bands

The molecular dynamics trajectories run by Réal et al. [134] (PCMT group) and used to generate conformations for results in chapter 3 showed that the first solvation shell consists of nine water molecules against six observed experimentally. This difference, of the anion apart from being due to the fact that a true sample contains conterions while the droplet MD do not, may results in considerable changes in water structures in the vicinity of the anion. molecular structure. Local structure is behind many species behaviours. For example small anions with a higher charge density have a tightly bond first solvation shell, form hydrogen-bonds to local water molecules and are structure makers. However larger but more diffuse anions without a strong solvation shell are structure breakers as in the case of large ions which tend to be gathered at the air-water interface [137, 138].

In the case of the chloride solution, it is possible to distinguish between two types of water molecules within the first solvation shell, those that hydrogen bond to the chloride, and those that remain local within the first solvation shell but which form hydrogen-bonds to other water molecules [137]. The used FF [134] (denoted FF16) shows artefacts in describing pairwise interactions between anions and surrounding water molecules and water-water interactions in the first solvation shell. To remediate this, an improvement to the FF [139] was proposed by adding a three-body correction that models more accurately water-water interactions in the first hydration shells of large-sized halide clusters and to improve in an averaged way the description of anion-water interactions.

To examine the performance of the new FF (denoted FF18) in comparison with the first one, we start by computing BE of chloride in small clusters of water containing up to seven water molecules and we compare the stabilisation energies (SEs) obtained with the two force fields to experimental values [135]. The small clusters come from classical MD trajectories prepared in our group by Réal et al. and the calculations are done at the level of supermolecule SAOP (SR-ZORA) averaged over 100 snapshots.

n	FF16			FF18			Markovich et al.		
	BE	SE	Δ_{SE}	BE	SE	Δ_{SE}	BE	SE	Δ_{SE}
0	3.38	0.00		3.38	0.00		3.61	0.00	
1	4.10	0.71	0.71	4.10	0.71	0.71	4.37	0.76	0.76
2	4.73	1.35	0.63	4.76	1.38	0.66	4.97	1.36	0.60
3	5.33	1.94	0.60	5.18	1.79	0.41	5.5	1.89	0.53
4	5.82	2.44	0.50	5.42	2.04	0.25	5.92	2.31	0.42
5	6.33	2.94	0.51	5.76	2.38	0.33	6.21	2.60	0.29
6	6.49	3.11	0.17	6.12	2.74	0.36	6.58	2.97	0.37
7	6.89	3.51	0.39	6.35	2.96	0.23	6.88	3.27	0.30

Table 4.4: Comparison of SAOP (SR-ZORA) BE, SE, and differential stabilisation energies $\Delta_{SE} = SE_n - SE_{n-1}$ for the clusters $[\text{Cl}(\text{H}_2\text{O})_n]^-$, $n=0,7$ water molecules between the first FF16 [139], the corrected FF18 [139] and experimental results [135].

As shown in Table (4.4) and illustrated in Figure (4.5), the theoretical values for 1 and 2 water molecules seem coherent with experiment, but for 3 or more waters, the droplets seems to somewhat over-stabilise the chloride for the FF16 and under-stabilise it for the FF18.

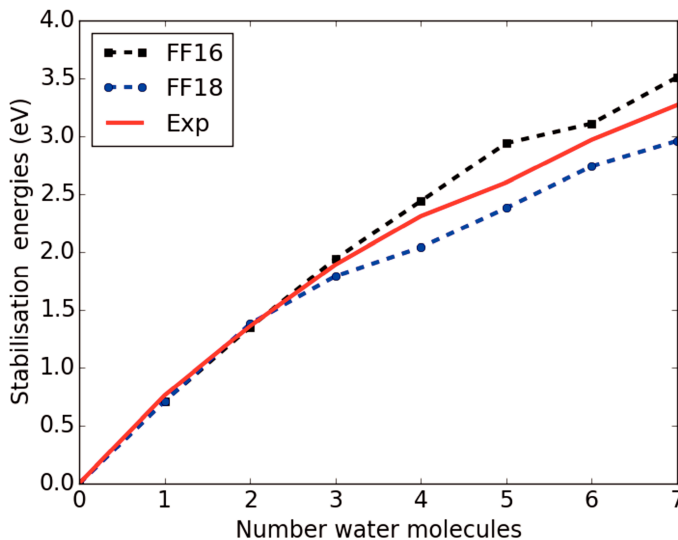


Figure 4.5: Comparison of stabilisation energies for Cl^- in water molecules obtained for the first FF16 [139], the corrected FF18 [139] and experimental results [135].

That said, if, for the FF16 the absolute value of the stabilisation is larger for theory, the rate of change of the stabilisations with increased droplet size varies more and that why we obtained overestimated values for chloride hydrated in 50 water molecules. In contrary, the FF18 tends to under-estimate the SEs. It is noticeable that an average of the two FFs yields

better agreement with experiment. As an example, for seven water molecules, we obtain a stabilisation energy of 3.51 eV with FF16 and 2.96 eV with FF18 while their average give 3.24 eV which is much closer to the experimental one of 3.27 eV. However using EOM-IP yields to increasing the BE value with an average of 0.3 eV for the $3p$ band which makes FF18 results agree in better way with experiment than FF16 ones.

The next step is to compare the results for a 50-water droplet. Figure (4.6) plots data from Table (4.5) and collected from SAOP (SR-ZORA) calculations with the triple zeta basis set over 100 snapshots within the FDE approach with a treatment of the environment as a one subsystem, confirms that water bands maintain their positions with very negligible changes.

orbital		binding energy/eV				Difference/eV
		FF16		FF18		
		average	std	average	std	
chloride	$3p$	9.47	0.34	8.91	0.34	-0.56
	$3s$	21.05	0.34	20.49	0.34	-0.56
	$2p$	189.01	0.33	188.48	0.33	-0.53
	$2s$	249.95	0.33	249.42	0.33	-0.53
	$1s$	2750.87	0.32	2750.35	0.32	-0.52
water	$1b_1$	10.43	0.46	10.39	0.47	-0.04
	$3a_1$	12.49	0.67	12.46	0.68	-0.03
	$1b_2$	15.94	0.48	15.90	0.48	-0.03

Table 4.5: Comparison of the binding energies of $\text{Cl}^-@(\text{H}_2\text{O})_{50}^{\text{tog}}$ over 100 snapshots obtained from MDs performed of the two FFs at the level of SAOP (SR-ZORA).

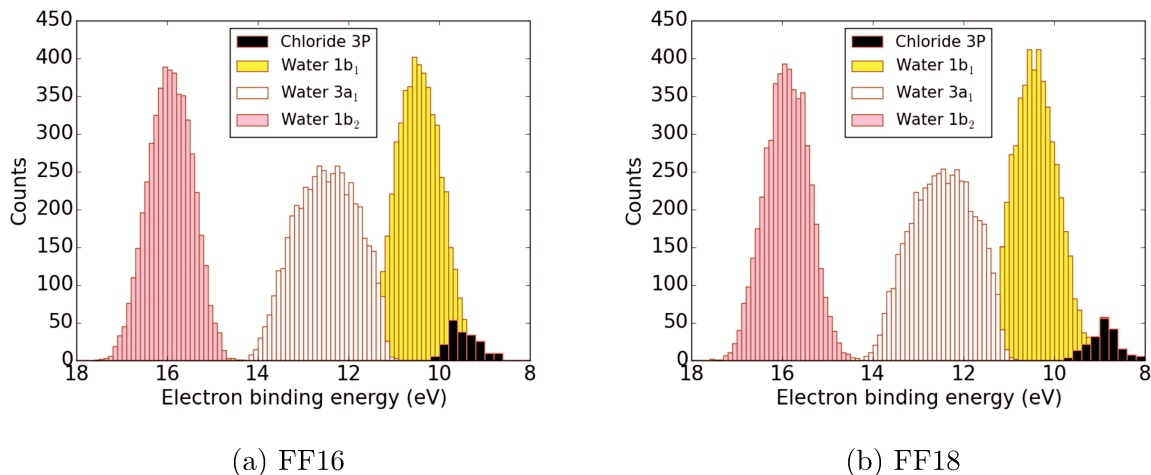


Figure 4.6: Calculated bands of $\text{Cl}^-@(\text{H}_2\text{O})_{50}^{\text{tog}}$ over 100 snapshots obtained from MDs performed of the two FFs at the level of SAOP (SR-ZORA) with a treatment of the water cluster as one fragment.

However using the FF18 results in a negative shift largest than 0.5 eV for chloride bands due to the difference in local structures reproduced by the two FFs. From our previous calculations based on the FF16, we can evaluate the changes that occur for the chloride p orbital energies, when we perform calculations at the level of SAOP (SR-ZORA), SAOP (DC) and CC-EOM-IP

over 100 snapshots of $\text{Cl}^-@(\text{H}_2\text{O})_{50}^{tog}$. These changes can be considered as corrections that we add to our FF18-SAOP (SR-ZORA) to presumed CC-EOM-IP values.

chloride		calculated binding energy/eV			experiment	
		SAOP (SR-ZORA)	SAOP (DC)	CC-EOM-IP	(a)	(b)
orbital	$P_{3/2}$	9.5(3)	9.4(4)	9.7(3)	9.5(2)	9.60(7)
	$P_{1/2}$		9.5(4)	9.9(3)	-	-
shift			-0.1	0.2		
			-0.0	0.4		

Table 4.6: Changes undergone by the chloride 3p orbital of $\text{Cl}^-@(\text{H}_2\text{O})_{50}$ for different level of computation over FF16 100 snapshots; (a) from Kurahashi et al. [133] and (b) from Winter et al. [131]

In order to predict CC-EOM-IP of $P_{3/2}$ and $P_{1/2}$ values, Table (4.6), shows that we have to add 0.2 eV and 0.4 to the SAOP (SR-ZORA) value respectively. Moreover the correction due to the use of separated representation of the environment, as mentioned before (see Table (4.2)), consist in subtracting 0.17 eV from the 3p orbital BE value. The FF18 $\text{Cl}^-@(\text{H}_2\text{O})_{200}^{sep}$ bands are then calculated over 100 snapshot at the SAOP (SR-ZORA) level (see Tab.(4.7)).

orbital		binding energy/eV			
		$\text{Cl}^-@(\text{H}_2\text{O})_{200}^{sep}$		experiment	
		Ave (corrected)	SD	(a)	(b)
chloride	$3p$	9.40 (9.43, 9.63)	0.36	9.5(2)	9.60(7)
	$3s$	20.98	0.36		
	$2p$	88.97	0.35		
	$2s$	249.91	0.35		
	$1s$	2750.84	0.35		
water	$1b_1$	11.54	0.49		
	$3a_1$	13.42	0.50		
	$1b_2$	16.75	0.47		
	$2a_1$	29.17	0.48		

Table 4.7: FF18 Average (Ave (corrected) subtracting 0.17 eV and adding 0.2 eV and 0.4 eV respectively for $P_{3/2}$ and $P_{1/2}$ EOM values) electron BE (eV) of the hydrated chloride from SAOP (SR-ZORA) calculations over 100 snapshots with triple-zeta basis; SD refers to standard deviation. (a) from Kurahashi et al. [133] and (b) from Winter et al. [131].

The results show that the predicted EOM-IP results agree in better way with experiment. For the $3p$ orbital, the applied correction leads to an average of the two estimated CC-EOM-IP $P_{3/2}$ and $P_{1/2}$ values of 9.53 eV. This value lies, as plotted in Figure (4.7) between the two provided experimental values. It is shown that we achieve a good agreement with experiment for water and also for chloride.

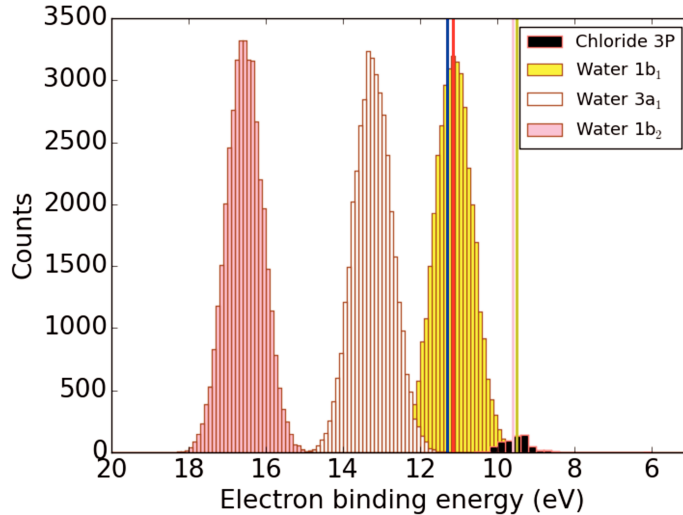


Figure 4.7: Corrected BEs of $\text{Cl}^- @ (\text{H}_2\text{O})_{200}^{\text{sep}}$ system from SAOP (SR-ZORA) calculations over FF18 100 snapshots with triple-zeta basis. Vertical lines refer to experimental values [133, 131].

4.3 Perspectives

As mentioned before, simulating halides in water bulk can represent the ultime step of probing the size effect on ionisation potentials. A subsystem DFT method capable of running ab-initio molecular dynamics simulations in periodic box (Quantum Espresso) to yield accurate models for liquid water [140] can be coupled to our protocole. Moreover, given that large anions are structure breaker and tend to be gathered in an air-water interface, it could be possible to generate MD snapshots in the corresponding conditions and apply the same computational protocole. Such investigation may bring precious information about the anion location in the surface or bulk, especially because the used photoelectron spectroscopy apparatus samples both bulk and interfacial anions [131], but without detection of any PE peak differences.

A third axis can also be followed. It concerns the effect of solvation on the core (or semi-valence) BEs. The particularity of these ionisations is that we need to take, in addition of the electronic correlation, the electronic relaxation which is very important [141, 142].

Chapter 5

FDE NMR properties calculation of platinum-thallium bonded complexes.

After having shown the good performance of the FDE in studying hydrated halides ionisation potentials, we would like to explore whether FDE performs well for a new category of molecular properties. In the following, we investigate the NMR properties of the $(\text{NC})_5\text{-Pt-Tl}$ complex, which shows the largest spin-spin coupling constant reported in the literature between two different nuclei (Pt-Tl). In addition, the electronic structure of this complex needs special attention due to the involvement of large relativistic effects as well as the different possibilities of photoreduction.

5.1 $[(\text{NC})-\text{Pt}-\text{Tl}]$ complexes in the literature

Thallium, element 81, discovered in May 1861 by Sir William Crookes, is a white soft metal and is a member of the boron-group (group 13). [143] Its main use for is most likely in organic synthesis where thallium compounds are used as oxidising agents, intermediates or catalysts. Most heavy atoms, such as W, Mo, Pt, Tl, Pb and Hg can form metal-metal bonds in their coordination compounds. [144, 145, 146] This is due to unoccupied p and d orbitals of the metal ions, geometrically suitable to form molecular orbitals.

Most metal-metal bonded complexes need bridging ligands to support the bond, otherwise it is not stable. Metal-metal interactions and in particular platinum-thallium, which can involve bond without bridging ligands (unsupported), interactions interest researchers. [147]

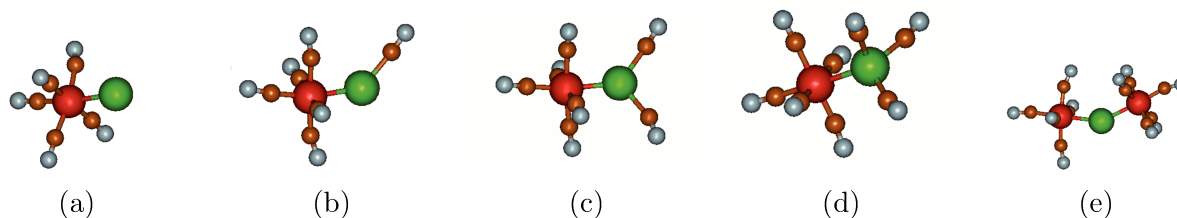


Figure 5.1: (a-d): compounds I-IV with the formula $[(\text{NC})_5\text{Pt-Tl}(\text{CN})_n]^{n-}$, $n = 0-3$ and (e): compound V with the formula $[(\text{NC})_5\text{Pt-Tl-Pt}(\text{CN})_5]^{3-}$. Pt (red), Tl (green), C (brown) and N (gray).

The Tl-Pt is also interesting as it can form several metal-metal bond as illustrated by

complexes I-V in Figure (5.1), which form the a family of platinum-thallium-cyanide metal-metal strongly unsupported bonded complexes with the formula $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{CN})_n]^{(n)-}$ ($n = 0-3$ for compound I, II, III and IV, respectively) and $[(\text{NC})_5\text{Pt}-\text{Tl}-\text{Pt}(\text{CN})_5]^{3-}$ (for compound V). Owing to the two stable and NMR-sensitive isotopes of thallium, ^{205}Tl and platinum ^{195}Pt , Tl-Pt, this complexes had their structures characterised by analysing the corresponding NMR spectra of $\text{Tl}(\text{ClO}_4)_3$, $\text{K}_2\text{Pt}(\text{CN})_4$ and NaCN synthesised in aqueous solution. [148, 149] is found and systematically studied in the literature. [148, 149]

For the complex I, after identifying the spin-spin coupling between the ^{205}Tl and ^{195}Pt nuclei from the spectrum shown in Figure (5.2a), the assignment of the ^{205}Tl (5.2b) and ^{13}C NMR spectra leads to the identification of two groups of cyanides (CN) ligands with the intensity ratio 1:4, spin-spin coupled to Pt and Tl. The chemical shift value and coupling constants of the ^{13}C most intensive signal are indicative of the C^C cyanides, leading to assign it to the four carbons of the $\text{Pt}(\text{CN})_4$ and identifying the final compound I structure, the $(\text{CN}^-)_5-\text{Pt}^{2+}-\text{Tl}^{3+}$ system. The corresponding experimental results are shown in Table (5.1).

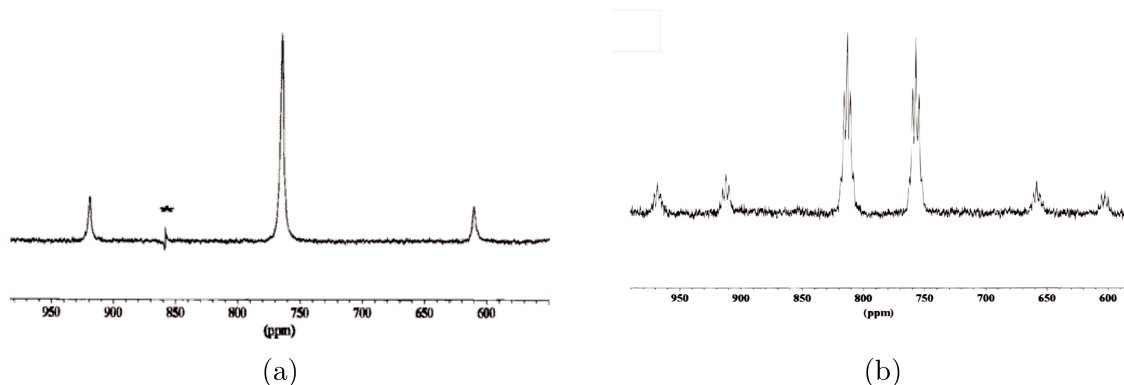


Figure 5.2: ^{205}Tl NMR spectrum of complex I. Aqueous solution containing 50 mM Tl, 50 mM $\text{Pt}(\text{CN})_4$, and (a): 100 mM KCN, (b): 100 mM Na^{13}CN . taken from Ref. [148]

This compound as their homologues $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{CN})_{n-1}]^{(n-1)-}$ have the ability to undergo a photo-induced redox reactions which make them very important for applications in the field of solar energy conversion into electric energy at the price of few modifications in order to adapt them to the requirements of the solar cell.

	δ_{Tl}	δ_{Pt}	δ_{CA}	δ_{CC}	$^1J_{\text{Tl-Pt}}$	$^2J_{\text{Tl-CA}}$	$^2J_{\text{Tl-CC}}$	$^1J_{\text{Pt-CA}}$	$^1J_{\text{Tl-CC}}$
$[(\text{NC})_5\text{Pt}-\text{Tl}]$	786	473	93.4	90.3	71060	12746	592	909	820

Table 5.1: (unsigned) NMR parameters for Pt-Tl-CN compound. Chemical shift (δ , in ppm) and spin-spin coupling (nJ , in Hz and n is the number of involved bonds). The chemical shifts were referred (in ppm) toward TlClO_4 , Na_2PtCl_6 , and (c) water-soluble sodium salt of TMS, for ^{205}Tl , ^{195}Pt , and ^{13}C NMR spectra, respectively. [148]

The Tl-Pt compound exhibits a strong spin-spin coupling between the thallium and platinum nuclei in aqueous solution. It surpasses not only the previously known values for Tl-Pt bond¹, but also the largest reported between two different nuclei. Relativistic effects have to

¹There are seven known compounds containing a direct Tl-Pt linkage. [150, 151, 152, 153, 154] Values of $^1J_{\text{Tl-Pt}}$ are not available for all of them. [148]

be taken into account when spin-spin coupling constants involve heavy elements. [155] In the case of thallium and due to the presence of s valence orbitals, the relativistic contraction of the 6s shell leads to an increase of 6s electron density on the nucleus which results in large values of J_{Tl-X} as it is the case of ${}^2J_{Tl-C^A}$ with 12746 Hz.

Due to the diamagnetic nature of the complex in solution (198 electrons), the bonding situation in the system can be interpreted as a donation of electrons between thallium(III) and platinum(II), leading to thallium(I) and platinum(IV) as final products. [148] Moreover, it is believed that this heavy metal-metal bond is strong and stable both in solution and in solid, for which the quantum mechanical calculations are complicated, but clearly needed to understand the different details of the bonding situation between the two metals, especially when the detailed mechanisms which determine the experimental data are not understood. [156]

At the level of theoretical calculations, Autschbach and Le Guennic [157, 158] have demonstrated that, it is necessary, at a scalar relativistic level, to take into account both the explicit water molecules (The first solvation water shell is represented by adding water molecules (maximum (5H₂O) to surround the Tl atom) and additional implicit (COSMO) solvent effects in order to reproduce the Pt–Tl spin-spin coupling constants and the ${}^{205}\text{Tl}$ and ${}^{195}\text{Pt}$ chemical shifts. It was observed also that, there is no obvious either with correlation between $J_{\text{Pt-Tl}}$ and the Pt–Tl bond lengths or with the inverse of the HOMO-LUMO gap. Moreover, they showed that, as the experimentally observed trends are largely caused by solvent coordination of the complexes, the calculated spin-spin coupling constants are improved systematically upon introducing more realistic computational models for the treatment of the solvent. This study showed, relatively, good agreement with experiment (due to errors cancelation), but the static solvent coordination is not very significant statistically for a dynamic complex in a solvent. Hence, the results are not very representative.

Recently, Ducati et al. [159] have performed a more realistic study that models the concerned system, in addition of the complex, by more water molecules (20), complemented by an implicit model (COSMO) to take into account long-range effects. Using PBE-based Car-Parrinello molecular dynamics [160] (CPMD) to generate a set of snapshot geometries (256) and PBE0 relativistic hybrid KS NMR calculations, they have confirmed that the J-coupling constants of such system, mainly for Tl–Pt one, to have a strong dependence on the coordination of water at the Tl site. Moreover, the computational model gave reasonable agreement with the experimental data (within over 10% deviation from experiment for the $J_{\text{Pt-Tl}}$ value). They suggest, beyond the impracticability of using relativistic correlated wave function level of theory, that it might deem necessary to use hybrid functionals in the MD simulations to better describe the particular three-centre-four-electron bonding $C^A\text{–Pt–Tl}$ moiety, which is expected to be sensitive to the presence of water since that such improvement may minimise the KS delocalisation error. [161] However, this study is restricted to spin-spin couplings without any mention of shieldings or chemical shift and their sensitivity to solvent molecules.

5.2 FDE for the $(\text{NC})_5\text{–Pt–Tl}$ complex

Taking in consideration all the impracticalities to treat the system with correlated wave-function methods (there isn't yet a code that can do NMR properties with coupled-cluster method and 4-components relativity level), as well as the high cost of describing electronic structure for the NMR properties for such systems even with DFT, we want to investigate

this system within the FDE approach trying to obtain simple and more economical theoretical models. We first investigate the complex I, which we consider to be the most interesting one, since it shows the largest Tl–Pt spin-spin coupling constant reported in the literature for two different nucleus (so the largest in the family) and is the most water-coordination-dependent compound which makes the solvation effects the most important in the series.

First, the strategy consists in minimising the size of the part that can be treated as subsystem of interest (we denote it from here as subsystem I) and secondly that yields acceptable results in comparison with supermolecular calculations. The 256 geometries [159] will be used in the present work and the PBE0 is adopted in different ADF calculations. We use also the same basis set consist in *jcpl* for Tl and Pt and *TZ2P* for other atoms.

5.2.1 Chemical model and relativistic Hamiltonian

As mentioned before, Ducati et al. [159] in addition to considering 20 explicit water molecules, used COSMO in order to capture long range effects, within a scalar-relativistic-ZORA Hamiltonian. They also checked, for one snapshot, that the effect of COSMO is equivalent to that of explicitly treating 125 water molecules. Thus, before establishing our computational protocol, and given that, for implementations questions, we cannot use FDE and COSMO together, we start by going beyond 20 water molecules. For this purpose, we use the simulation box containing 64 water molecules to select the convenient chemical model by comparing the effects of the water molecules number, for one snapshot, between systems with 35 and 64 water molecules. In other words, we considered two structures as supermolecule, One using all molecules within the simulation box, and another in which we kept only 35 water molecules closest to the {Tl–Pt complex as illustrated in Figure (5.3).

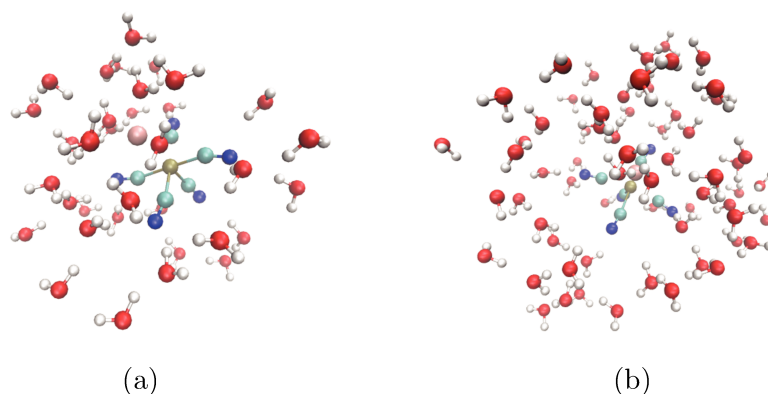


Figure 5.3: Perspective view of the chemical models, (a) : $(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{31}$ and (b) : $(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{64}$

Ducati et al. [159] so far restricted themselves a SR-ZORA Hamiltonian for the J-coupling computations. While they found this to be enough to capture the relativistic effects, that will not necessarily be true to other properties. Since we are also interested in NMR shieldings, we will investigate the effect of spin-orbit coupling on these properties. Tabs.(5.2) and (5.3) summarise supermolecular calculations, the comparison between results of SR and SO ZORA Hamiltonians in the cases of 35 and 64 water molecules (one structured snapshot only) for NMR shieldings and J-couplings.

	SR-ZORA		SO-ZORA	
	(H ₂ O) ₃₅	(H ₂ O) ₆₄	(H ₂ O) ₃₅	(H ₂ O) ₆₄
Tl	7869	7829	12682	12719
Pt	2143	2216	5631	5705
C ^A	70	67	117	111
C ^C	49	48	83	83

Table 5.2: Size and SO coupling effects on calculated shieldings (ppm) for 1 snapshot for the chemical models (CN)₅Pt–Tl (H₂O)_n (n=35,64).

As a result, whether or not use spin-orbit coupling going from a 35 water molecules representation to a 64 one, the calculated shieldings are found to be insensitive in the contrast of the spin-spin couplings that undergo considerable shifts (more than 7 KHz) which demonstrates the J-coupling-dependance to long-range solvation effects. For the shieldings, the weak sensitivity to the number of explicit water molecules can be related to the fact that after a certain number of surrounding molecules, the atomic sites may be totally shielded and therefore adding more molecules does change anything. The spin-orbit coupling significantly impacts shielding, while the effects on J-couplings is smaller in relative smaller terms.

		SR-ZORA		SO-ZORA	
P	R	35	64	35	64
Tl	Pt	48482	55979	47387	55297
	C ^A	15995	16575	14997	15523
	C ^C	-712	-698	-723	-712
Pt	C ^A	1260	1216	1288	1240
	C ^C	923	920	917	915

Table 5.3: Size and SO coupling effects on calculated spin-spin couplings (Hz) between perturbing (P) and responding (R) element for 1 snapshot for the chemical models (CN)₅Pt–Tl (H₂O)_n (n=35,64).

In the light of results obtained from this comparison and to optimise the computational time, we will include SOC in all further coming calculations, and we will select a chemical hydration model encompassing 64 water molecules.

5.2.2 Computational protocol

The determination of the computational protocol mainly depends on the answer of how the FDE reproduce the NMR parameters in comparison with supermolecule (SM) calculations. To envision the beginnings of an answer, we proceed to perform our tests on the target system of 35 water molecules to speed up the calculations. In the tests, the whole system is divided into two parts, a subsystem I, containing the compound I and a number of water molecules going from 0 to 30 selected on the criterion of their proximity to the Tl-site (NMR parameters, mainly $J_{\text{Tl-Pt}}$ converge faster when we add water molecules to Tl-site), and a subsystem II which contains the rest of water molecules that varies between 35 and 5. The subsystem II is treated in the tests as one unity (we denote it as together representation). Moreover, we do not relax the embedding potential with the Freeze-and-Thaw (FnT) procedure. This choice is motivated by the fact that in the case of a neutral subsystem, resorting to the FnT is not

that necessary as when charged subsystems are present and therefore the polarisation of one subsystem to the other is important. [162] The results of these tests are shown in Figs.(5.4) and (5.5). For the shieldings it appears that in general after including 4 water molecules in the subsystem I, the obtained values shows are very close to the SM values. The difference between the SM and $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{31}$ shieldings values is of 92 *ppm* for Tl, 222 *ppm* for Pt, 12 *ppm* for C^A and 3 *ppm* for C^C .

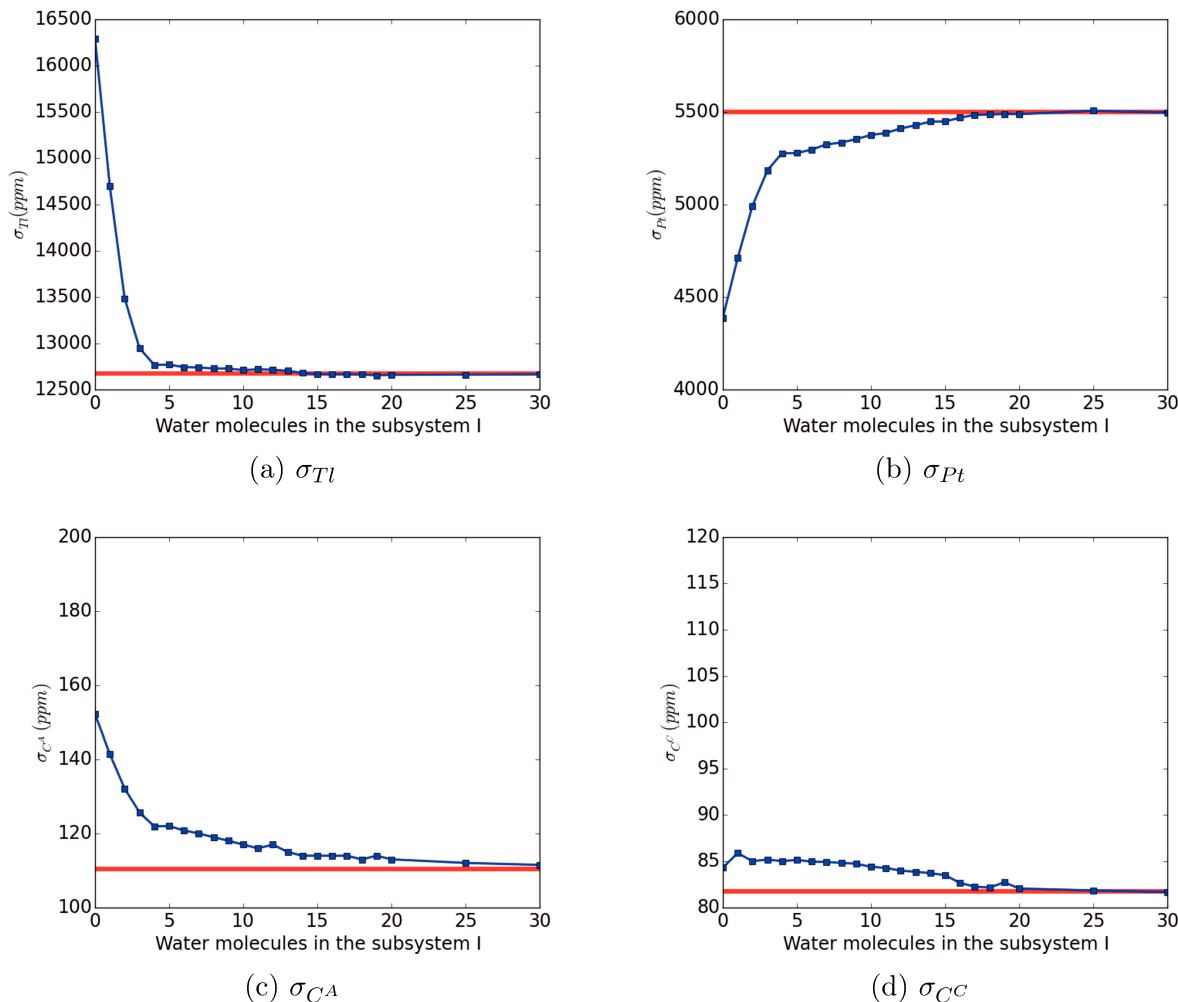


Figure 5.4: Convergence of FDE (blue square lines) with respect to the supermolecule results (horizontal red lines) for the shieldings (in *ppm*) for models $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_n]@(\text{H}_2\text{O})_{n-1}$ with increasing n , the number of explicitly treated water molecules.

By adding more water molecules in subsystem I, the J -coupling values converge to the SM values already with 15 explicit water molecules. The difference between the SM and $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{15}]@(\text{H}_2\text{O})_{20}$ is of 2218 *Hz* for the $J_{\text{Tl}-\text{Pt}}$, 35 *Hz* for the $J_{\text{Tl}-C^A}$, 14 *Hz* for the $J_{\text{Tl}-C^C}$, 17 *Hz* for the $J_{\text{Pt}-C^A}$ and 12 *Hz* for the $J_{\text{Pt}-C^C}$. This lead to choose a model with 4 explicit water molecules for the shielding parameters and 15 for the J -couplings. Figure (5.6) which illustrates the computing time for FDE calculation on 8 processors, shows that performing FDE calculation with 4 water molecules in the subsystem costs around $14/104 = 13\%$ of the SM computation time. This enables us to run 7.43 calculations in lieu of 1 and thus, performing 256 snapshots FDE calculations rather 35 SM ones at the same computational cost.

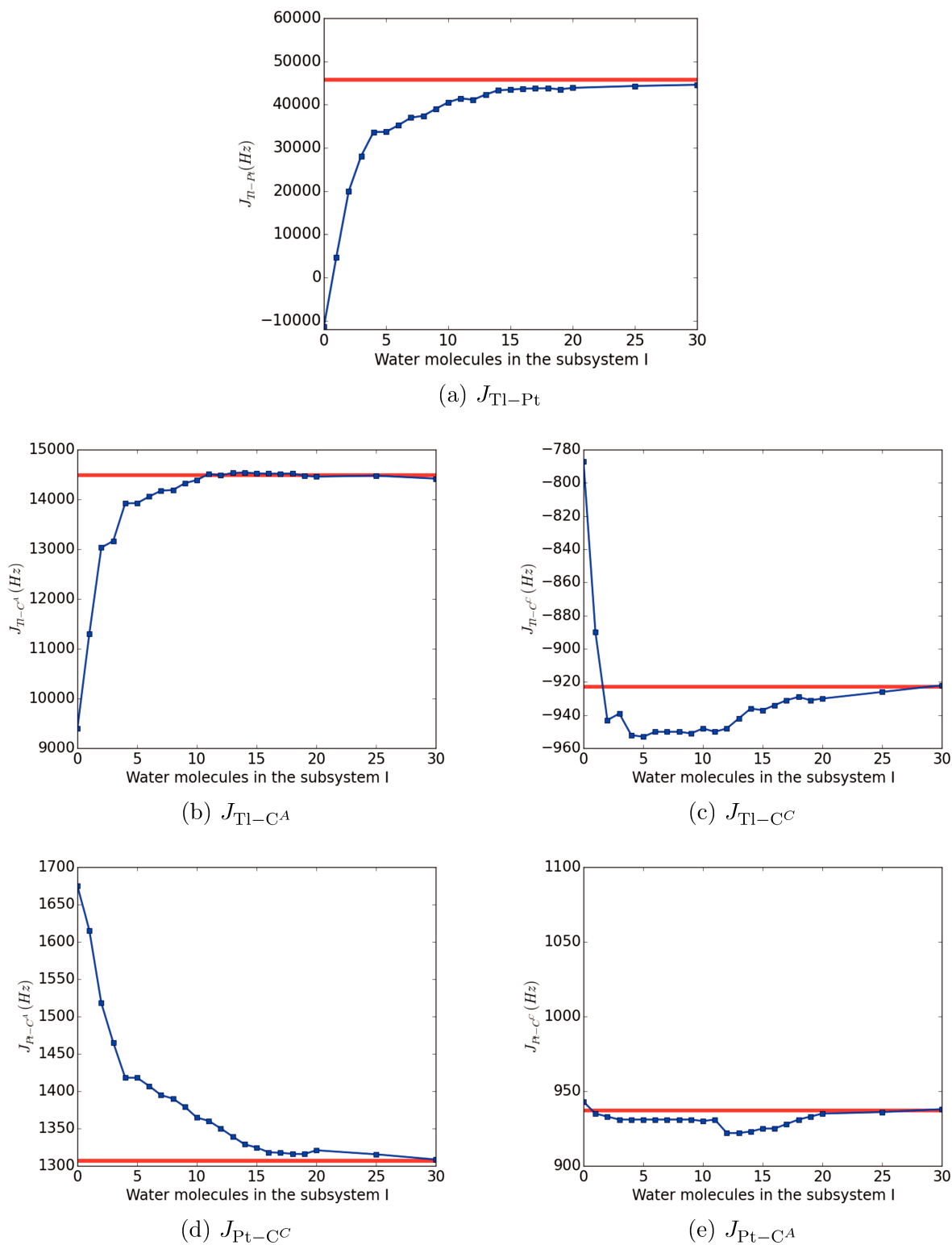


Figure 5.5: Convergence of FDE (blue square lines) with respect to the supermolecule results (horizontal red lines) for the J -couplings (in Hz) for the shieldings (in ppm) for models $[(CN)_5Pt-Tl(H_2O)_n]@(H_2O)_{n-1}$ with increasing n , the number of explicitly treated water molecules.

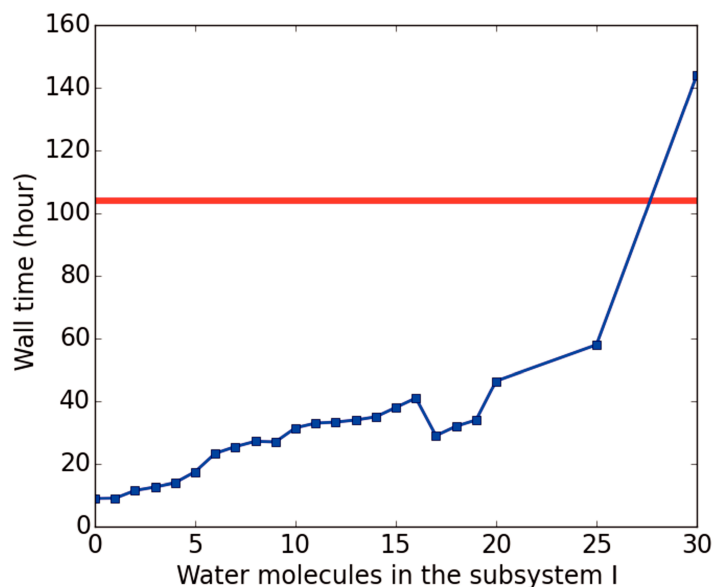


Figure 5.6: Comparison of the computing time between supermolecule and FDE calculations for models $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_n]@(\text{H}_2\text{O})_{n-1}$ with increasing n , the number of explicitly treated water molecules).

In the light of this comparison, we choose to use for our computational protocol two models for the active subsystems (subsystems I), as are illustrated in Figure (5.7)

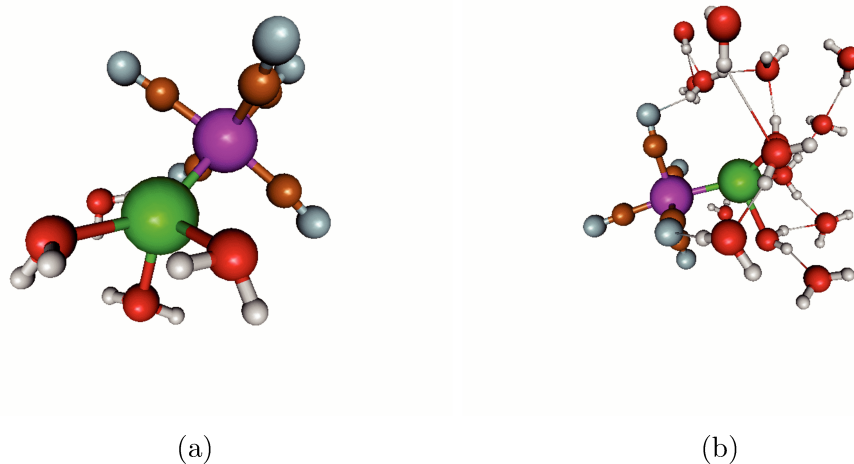


Figure 5.7: The active systems composed, in addition of $[(\text{NC})_5\text{Pt}-\text{Tl}$, (a) $(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$; (b) $(\text{H}_2\text{O})_{15}]@(\text{H}_2\text{O})_{49}$.

At the end of this part, additional tests on the effect of separated fragmentation of the environment (subsystem II), yield to small changes in comparison with the together representation with a very important saving of computing time, with less than 4 hours for $[(\text{CN})_5\text{Pt}-\text{Tl}]@(\text{H}_2\text{O})_{35}^{\text{sep}}$ in lieu of more than 9 hours for $[(\text{CN})_5\text{Pt}-\text{Tl}]@(\text{H}_2\text{O})_{35}^{\text{tog}}$.

Finally, to recall, in the next, calculations will be done within the FDE approach and a separated representation the water environment.

5.3 Statistics' significance

The simulation of liquids require to include the effects of the temperature. One calculation value is never enough to obtain relevant values for a dynamic system. However hundreds or thousands of calculations without checking the statistics could be a big waste of time and computing resources as the selected snapshots might be correlated for a given property. As checked for the halides ionisation potential study, we have examined our snapshots results with two statistical criteria, the autocorrelation functions and the evolution of the averages. Figs.(5.8) and (5.9) show that for 256 snapshots, the statistics stabilise the reported averages for $J_{\text{Tl-Pt}}$ and σ_{Tl} . It is also demonstrated that the snapshots are non correlated. Similar trends are observed for all investigated properties which makes the calculations to be statistically significant.

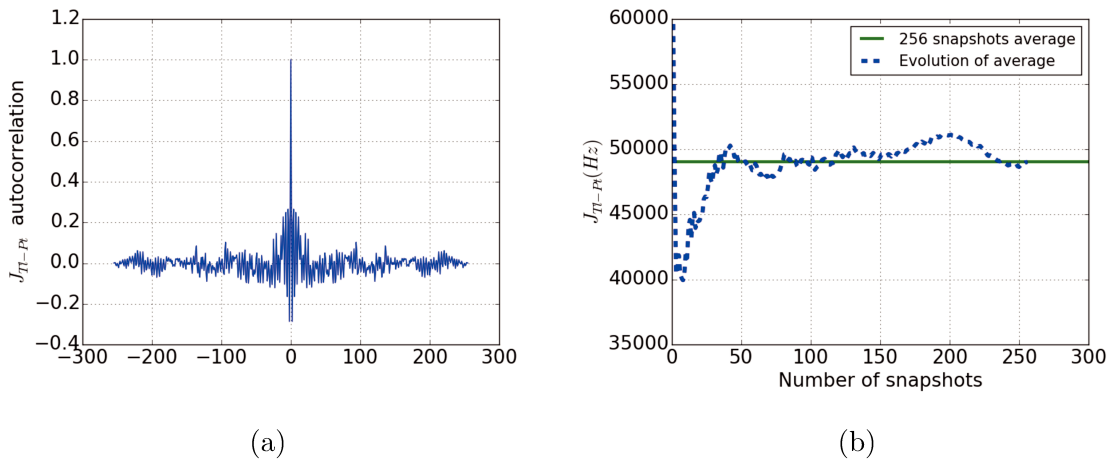


Figure 5.8: $J_{\text{Tl-Pt}}$ (a) : autocorrelation function and (b) : average evolution, obtained from 256 $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ calculations.

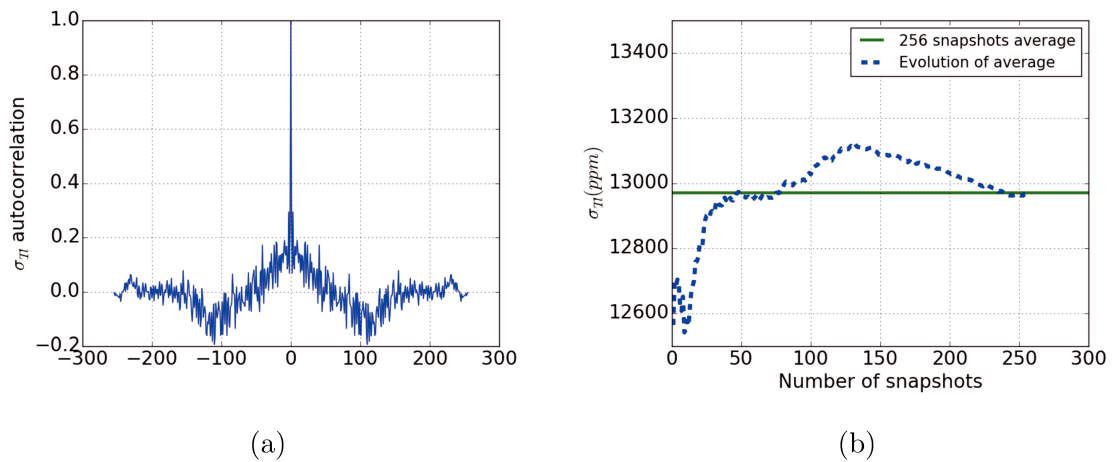


Figure 5.9: σ_{Tl} (a) : autocorrelation function and (b) : average evolution, obtained from 256 $[(\text{CN})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ calculations.

5.4 Final results

The final results of our computations are presented here. We compare the average J -coupling values mainly the $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ and $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{15}]@(\text{H}_2\text{O})_{49}$ models to Ducati et al. [159] theoretical calculations and to available experimental results. [148] However, the chemical shifts are calculated relative to the available reference data.

5.4.1 Spin-spin couplings

Table (5.4) summarise all the J -coupling constants, obtained from averaging over 256 snapshots. The $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{15}]@(\text{H}_2\text{O})_{49}$ model improves the $J_{\text{Tl}-\text{Pt}}$ by more than 12000 Hz in comparison with $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ relative to the previous theoretical values that used the same dynamics. This trend is observed also for $J_{\text{Tl}-\text{C}^{\text{C}}}$ and $J_{\text{Pt}-\text{C}^{\text{A}}}$. The $J_{\text{Tl}-\text{C}^{\text{A}}}$ also converge to the previous theoretical calculations but overestimates by about 2500 Hz the experimental results. This may confirm that there might be issues with the PBE-based dynamics and its common delocalisation errors with DFT-GGA functionals. The remaining $J_{\text{Pt}-\text{C}^{\text{C}}}$ values do not undergo large shifts.

$J(Hz)$	$[(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{31}$	$[(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$	$[(\text{H}_2\text{O})_{15}]@(\text{H}_2\text{O})_{49}$	Ref [159]	Expt [148]
Tl-Pt	32632 ± 17100	37003 ± 17884	49032 ± 19173	63350	71060
Tl-C ^A	14224 ± 2521	14678 ± 2424	15356 ± 2533	16718	12746
Tl-C ^C	-857 ± 583	-880 ± 486	-870 ± 509	-713	592
Pt-C ^A	1493 ± 122	1474 ± 117	1372 ± 173	1089	909
Pt-C ^C	923 ± 147	922 ± 147	932 ± 149	867	820

Table 5.4: Effects of different models on J -coupling constants obtained from averaging over 256 snapshots.

On the other hand, the reported $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{31}$ model values come to estimate the effects of adding around 30 water molecules on the J -coupling constants. The comparison to the $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{60}$ model results in a difference around 4400 Hz for the $J_{\text{Tl}-\text{Pt}}$. That makes, in the assumption of considering that adding water molecules keeps the same effect even for outer shells, the $J_{\text{Tl}-\text{Pt}}$ at 125 water molecules and within the FDE approach, to reach something around $49032 + 2 * 4600 = 58000 Hz$. This places our results very close to the result of the study by Ducati et al. [159], in which the effect, as mentioned before, of COSMO combined to 20 explicit water molecules is found, for one geometry, to be equivalent to 125 water molecules. Moreover, the comparison to experiment depends mainly on the used dynamics and tests with hybrid functionals based dynamics are really needed to explore further the reasons behind the difference to experiment.

The last, but may be the most important remark is related to the signs of these constants. While the experimental results are unsigned, theoretical investigations provide signed values. The sign of a J -coupling constant is related to the order of the linking bond explained in subsection (1.4.2). As expected, the bonds involving platinum are positive which means in general the presence of odd bond, here of order one. However the bonds involving thallium with carbons are expected to be negative. This is true in the case of C^C, but not for C^A for which the situation is more complicated. In fact the Tl-Pt-C^A bond can be seen as a 3-center 4-electron

(3c-4e) bond model, derived from concepts developed for electron-deficient bonding². [163, 164, 165, 166] More detailed investigations are highly recommended to understand what happens locally in order to explain the variation of bond orders.

5.4.2 Shieldings and chemical shifts

The calculated shieldings are presented in Table (5.5). It is shown that the main changes occur when we adopt the $[(\text{NC})_5\text{Pt}-\text{Tl}(\text{H}_2\text{O})_{15}]\text{@}(\text{H}_2\text{O})_{49}$ model except for the C^C that maintain relatively a constant value.

$\sigma(\text{ppm})$	$[(\text{H}_2\text{O})_4]\text{@}(\text{H}_2\text{O})_{31}$	$[(\text{H}_2\text{O})_4]\text{@}(\text{H}_2\text{O})_{60}$	$[(\text{H}_2\text{O})_{15}]\text{@}(\text{H}_2\text{O})_{49}$
Tl	13127 ± 456	13123 ± 465	12972 ± 507
Pt	5061 ± 356	5113 ± 350	5331 ± 348
C^A	128 ± 8	125 ± 8	117 ± 9
C^C	89 ± 8	89 ± 8	87 ± 9

Table 5.5: Effects of different models on the calculated atomic shieldings (σ) obtained from averaging over 256 snapshots.

Due to the difficulties encountered in determining the absolute shieldings values experimentally, the comparison to the experiment requires the calculation of chemical shifts which implies to know the used reference data. However, the reference values are not provided in the experimental paper [148], but we know that the carbons chemical shifts are measured relatively to the TMS carbon 13 shieldings which are measured directly for pure liquid in Ref. [167]. This value of 183.94 ppm is used here to calculate the chemical shifts of the C^A and C^C as :

$$\delta_C = 183.94 - \sigma_C \quad (5.1)$$

In addition to the previous models, carbons chemical shifts are also computed for the $[(\text{NC})_5\text{Pt}-\text{Tl}]$ model with out any water molecules.

$\delta(\text{ppm})$	$[(\text{NC})_5\text{Pt}-\text{Tl}]$	$[(\text{H}_2\text{O})_4]\text{@}(\text{H}_2\text{O})_{31}$	$[(\text{H}_2\text{O})_4]\text{@}(\text{H}_2\text{O})_{60}$	$[(\text{H}_2\text{O})_{15}]\text{@}(\text{H}_2\text{O})_{49}$	Expt [148]
C^A	6 ± 6	56 ± 8	59 ± 8	67 ± 9	93.4
C^C	111 ± 13	95 ± 8	95 ± 8	97 ± 9	90.3

Table 5.6: Effects of different models on the chemical shifts (δ) obtained from averaging over 256 snapshots compared to experimental results [148].

First as shown in Table (5.6), it is revealed that the inclusion of water molecules is necessary. The C^A chemical shift depends tremendously on its coordination to water which confirms previous observations. [159] The inclusion of water molecules in the active subsystem improves considerably the results but still places it far from the experiment value. Hence, as mentioned before, testing hybrid functionals based dynamics appears to be decisive before trying to computing NMR parameters with large water environment. For the C^C , the obtained chemical

²Electron deficiency is a term describing atoms or molecules having fewer than the number of electrons required for maximum stability. At the atomic level, main group atoms having less than 8 electrons or transition metal atoms having less than 18 electrons are described as electron-deficient. At the molecular level, molecules which have an incompletely filled set of bonding molecular orbitals are considered to be electron-deficient.

shifts agree much better with experiment within the FDE models and the $[(\text{H}_2\text{O})_4]@(\text{H}_2\text{O})_{31}$ model can be considered to be large enough in terms of hydration to reach agreement with experiment. The comparison of chemical shifts for thallium and platinum could bring further perspectives and orient the next investigations, but at the price of computing the shieldings for the used references. This necessitates to perform new dynamics since the used references (TlClO_4 and Na_2PtCl_6) values are measured in liquid water.

Conclusions and perspectives

The current challenges of theoretical modelling target developments and improvements of the efficiency of theoretical approaches to treat increasingly large systems at an optimal level of accuracy. In this context, the Frozen Density Embedding (FDE) approach provides a very powerful tool for the quantum chemical treatment of large systems. It is based on a partitioning of the electron density into the density of an active subsystem and a frozen environment. In the calculation of the density of the active subsystem, the effect of the frozen environment is represented by an effective embedding potential, that contains the electrostatic potential of the environment, an exchange-correlation component and a kinetic-energy component. In contrast to most other embedding schemes used in theoretical chemistry, the FDE scheme provides a formulation that is in principle exact.

This thesis focusses on discussing the performance of FDE for the calculation and prediction of molecular properties for heavy-element based complexes in the presence of a water solvent. The first chapter of this thesis introduces the methods of theoretical chemistry, in particular the two quantum chemical families of approaches, namely correlated wave function theory (WFT) and density functional theory (DFT). Relativistic effects and relativistic Hamiltonians are also presented as relativistic quantum chemistry is the framework to be used for chemically and physically relevant simulations of heavy element containing molecules.

The second chapter presents the various strategies designed to model large-scale molecular systems, all within the family of quantum embedding methods. This lead us to introduce the FDE scheme, which allows to partition a molecular systems into interacting subsystems and to choose the most suitable electronic structure approach to treat each of these. In the FDE framework, the whole system can be described with several accuracy of DFT (DFT-in-DFT) or by a merge of more computationally demanding WFT for the subsystem of interest coupled to a DFT embedding potential.

In the next three chapter, we have analysed the performance of the FDE scheme for the simulations to account for hydration effects on two classes of properties, namely electron binding energies (Chapters 3 and 4) and magnetic ($J - J$ coupling and NMR shieldings) in Chapter 5, all done in the context of relativistic quantum modelling as we target mostly heavy elements.

We first discussed how FDE can account for the large hydration shift on valence electron binding energies of halides for the whole series starting with fluoride up to astatide, taking as a chemical model water droplets encompassing 50 water molecules, in a first chemical model. In the study published in Physical Review Letters, we have demonstrated that with the combination of relativistic EOM-CC for the active subsystem and DFT for the environment, afforded by FDE, one can rival with quite sophisticated theoretical approaches based on periodic quasi-particle calculations which are the current state-of-the-art for condensed matter simulations.

We have also explored, in a follow-up study, the sensitivity of the electron binding energies to structural changes around the halides and among these water, and how these energies (that of the halide and of water) evolve with the size of the water droplet. Our results have proven that long-range effects contribute to the computed photo-electron spectra. It remains to complete simulations at the relativistic EOM-CC level for larger droplets, to reach full quantitative agreement with experimental data, with a competitive computational cost.

The last chapter explores the performance of FDE for the description of solvent effects on magnetic properties (indirect spin-spin couplings and NMR shielding tensors) for a complex $\text{PtTl}(\text{CN})_5$ containing a metal-metal bond between the heavy centers (Pt, Tl), this time purely at relativistic DFT level. This complex was chosen as it so far exhibits the largest reported spin-spin coupling constant between two metals. FDE allows us to capture semi-quantitatively hydration effects beyond the first hydration shell thereby reducing the computational cost with respect to supermolecular calculations. For spin-spin couplings, we have shown that much like prior theoretical results, we require an extensive first hydration shell around the complex, but nevertheless arrive at a semi-quantitative agreement with experiment. For NMR shieldings on the other hand, FDE allows us to significantly reduce the amount of water molecules explicitly added to the active subsystem to the first hydration shell around the Tl atom. Ideally longer-range solvation effects might be introduced with an implicit solvation model such as COSMO. This might open up the perspective to employing FDE with more accurate electronic structure methods for this property for this class of compounds.

Finally, the established computational and workflow protocols could be transferred to the computation of different surrounded species properties. Particularly, they are already used, with small adaptations, for the calculation of inner shells ionisation potentials for the chloride adsorbed on ice surface in the master thesis of R. Opoku. Moreover, investigating molecular properties for heavy element species in surfaces and especially in solid environments is very interesting outlook, since additional complications are necessarily present. Systems for which a first challenge may consists in dividing the whole electron density across chemical bonds without losing a good representation of the target system electronic structure.

Bibliography

- [1] Kenny B Lipkowitz and Donald B Boyd. *Reviews in Computational Chemistry 7*. Wiley Online Library, 1996.
- [2] Frank Jensen. *Introduction to computational chemistry*. John Wiley & Sons, 2017.
- [3] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.
- [4] Tjalling Koopmans. Über die zuordnung von wellenfunktionen und eigenwerten zu den einzelnen elektronen eines atoms. *Physica*, 1(1-6):104–113, 1934.
- [5] Christopher J Cramer. *Essentials of computational chemistry: theories and models*. John Wiley & Sons, 2013.
- [6] Multi-configurational self-consistent field, <https://quantique.u-strasbg.fr/istpc/doku.php?id=istpc2017:start>.
- [7] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2002.
- [8] Rodney J Bartlett. Coupled-cluster theory and its equation-of-motion extensions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2(1):126–138, 2012.
- [9] Anna I Krylov. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker’s guide to fock space. *Annu. Rev. Phys. Chem.*, 59:433–462, 2008.
- [10] John F Stanton and Rodney J Bartlett. The equation of motion coupled-cluster method. a systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.*, 98(9):7029–7039, 1993.
- [11] Donald C Comeau and Rodney J Bartlett. The equation-of-motion coupled-cluster method. applications to open-and closed-shell reference states. *Chem. Phys. Lett.*, 207(4-6):414–423, 1993.
- [12] Marcel Nooijen and Rodney J Bartlett. Equation of motion coupled cluster method for electron attachment. *J. Chem. Phys.*, 102(9):3629–3647, 1995.
- [13] John F Stanton and Jürgen Gauss. Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.*, 101(10):8938–8944, 1994.

- [14] M Wladyslawski and M Nooijen. The photoelectron spectrum of the $\text{no} \sim 3$ radical revisited: A theoretical investigation of potential energy surfaces and conical intersections. In *ACS Symposium Series*, volume 828, pages 65–92. Washington, DC; American Chemical Society; 1999, 2002.
- [15] Roy McWeeny. *Methods of molecular quantum mechanics*. Academic press, 1992.
- [16] Sergey V Levchenko. *Equation-of-motion Coupled-cluster Model with Single and Double Substitutions: Theory and Applications*. University of Southern California, 2005.
- [17] Armin Shayeghi. *Bindungsverhältnisse und optisches Verhalten kleinster Gold-Silber-Legierungsbausteine*. PhD thesis, Technische Universität, 2016.
- [18] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864, 1964.
- [19] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133, 1965.
- [20] JF Janak. Proof that $\epsilon_n = \epsilon$ in density-functional theory. *Phys. Rev. B*, 18(12):7165, 1978.
- [21] R Van Meer, OV Gritsenko, and EJ Baerends. Physical meaning of virtual kohn–sham orbitals and orbital energies: an ideal basis for the description of molecular excitations. *J. Chem. Theory Comput.*, 10(10):4432–4441, 2014.
- [22] Ulf von Barth and Lars Hedin. A local exchange-correlation potential for the spin polarized case. i. *J. Phys. C: Solid State Phys.*, 5(13):1629, 1972.
- [23] David M Ceperley and BJ Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45(7):566, 1980.
- [24] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58(8):1200–1211, 1980.
- [25] Walter Kohn. Nobel lecture: Electronic structure of matterwave functions and density functionals. *Rev. Mod. Phys.*, 71(5):1253, 1999.
- [26] C Richard A Catlow, Berend Smit, and RA van Santen. *Computer modelling of microporous materials*. Elsevier, 2004.
- [27] PR Antoniewicz and Leonard Kleinman. Kohn-sham exchange potential exact to first order in $\rho(\mathbf{k})/\rho_0$. *Phys. Rev. B*, 31(10):6779, 1985.
- [28] Axel D Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098, 1988.
- [29] John P Perdew, Kieron Burke, and Yue Wang. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B*, 54(23):16533, 1996.

- [30] Carlo Adamo and Vincenzo Barone. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The m pw and m pw1pw models. *J. Chem. Phys.*, 108(2):664–675, 1998.
- [31] Axel D Becke. Density functional calculations of molecular bond energies. *J. Chem. Phys.*, 84(8):4524–4529, 1986.
- [32] John P Perdew. Jp perdew, k. burke, and m. ernzerhof, phys. rev. lett. 77, 3865 (1996). *Phys. Rev. Lett.*, 77:3865, 1996.
- [33] John P Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33(12):8822, 1986.
- [34] John P Perdew, John A Chevary, Sy H Vosko, Koblar A Jackson, Mark R Pederson, Dig J Singh, and Carlos Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B*, 46(11):6671, 1992.
- [35] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785, 1988.
- [36] Junwei Lucas Bao, Laura Gagliardi, and Donald G Truhlar. Self-interaction error in density functional theory: An appraisal. *J. Phys. Chem. Lett.*, 9(9):2353–2358, 2018.
- [37] AD Becke. Ad becke, j. chem. phys. 98, 1372 (1993). *J. Chem. Phys.*, 98:1372, 1993.
- [38] Axel D Becke. A new mixing of hartree–fock and local density-functional theories. *J. Chem. Phys.*, 98(2):1372–1377, 1993.
- [39] Axel D Becke. Density-functional thermochemistry. iii. the role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, 1993.
- [40] PJ Stephens, FJ Devlin, CFN Chabalowski, and Michael J Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98(45):11623–11627, 1994.
- [41] Sergio Filipe Sousa, Pedro Alexandrino Fernandes, and Maria Joao Ramos. General performance of density functionals. *J. Phys. Chem. A*, 111(42):10439–10452, 2007.
- [42] OV Gritsenko, PRT Schipper, and EJ Baerends. Approximation of the exchange–correlation kohn–sham potential with a statistical average of different orbital model potentials. *Chem. Phys. Lett.*, 302(3-4):199–207, 1999.
- [43] Oleg Gritsenko, Robert van Leeuwen, Erik van Lenthe, and Evert Jan Baerends. Self-consistent approximation to the kohn–sham exchange potential. *Phys. Rev. A*, 51(3):1944, 1995.
- [44] Oleg V Gritsenko. Ov gritsenko, r. van leeuwen, and ej baerends, int. j. quantum chem. 61, 231 (1997). *Int. J. Quantum Chem.*, 61:231, 1997.
- [45] R Shankar, K Senthilkumar, and P Kolandaivel. Calculation of ionization potential and chemical hardness: A comparative study of different methods. *Int. J. Quantum Chem.*, 109(4):764–771, 2009.

- [46] David J Tozer and Frank De Proft. Computation of the hardness and the problem of negative electron affinities in density functional theory. *J. Phys. Chem. A*, 109(39):8923–8929, 2005.
- [47] RG Pearson. J. absolute electronegativity and hardness: applications to organic chemistry. *Org. Chem.*, 54:1423–1430, 1989.
- [48] Ralph G Pearson. Absolute electronegativity and hardness: application to inorganic chemistry. *Inorg. Chem.*, 27(4):734–740, 1988.
- [49] Edward M Purcell, H Co Torrey, and Robert V Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 69(1-2):37, 1946.
- [50] John S Rigden. Quantum states and precession: The two discoveries of nmr. *Rev. Mod. Phys.*, 58(2):433, 1986.
- [51] Juha Vaara. Theory and computation of nuclear magnetic resonance parameters. *PCCP*, 9(40):5399–5418, 2007.
- [52] CP Slichter. Haerberlen u.high resolution nmr in solids: Selective averaging, supplement 1 to advances in magnetic resonance, 1976.
- [53] Melinda J Duer. *Solid state NMR spectroscopy: principles and applications*. John Wiley & Sons, 2008.
- [54] Kenneth L Williamson. Substituent effects on nuclear magnetic resonance coupling constants and chemical shifts in a saturated system: Hexachlorobicyclo [2.2. 1] heptenes. *J. Am. Chem. Soc.*, 85(5):516–519, 1963.
- [55] Nathan Janes and Eric Oldfield. Prediction of silicon-29 nuclear magnetic resonance chemical shifts using a group electronegativity approach: applications to silicate and aluminosilicate structures. *J. Am. Chem. Soc.*, 107(24):6769–6775, 1985.
- [56] Dieter Cremer and Jürgen Gräfenstein. Calculation and analysis of nmr spin–spin coupling constants. *PCCP*, 9(22):2791–2816, 2007.
- [57] nmr interactions, [//www.weizmann.ac.il/chemphys/assaf_tal/sites/chemphys.assaf_tal/files/uploads/nmr_primer_2015/lecture_ii_-_nmr_interactions.pdf](http://www.weizmann.ac.il/chemphys/assaf_tal/sites/chemphys.assaf_tal/files/uploads/nmr_primer_2015/lecture_ii_-_nmr_interactions.pdf).
- [58] Małgorzata Olejniczak, Radovan Bast, and André Severo Pereira Gomes. On the calculation of second-order magnetic properties using subsystem approaches in a relativistic framework. *PCCP*, 19(12):8400–8415, 2017.
- [59] Trygve Helgaker, Michał Jaszuński, and Kenneth Ruud. Ab initio methods for the calculation of nmr shielding and indirect spin- spin coupling constants. *Chem. Rev.*, 99(1):293–352, 1999.
- [60] Trond Saue and HJ Aa Jensen. Linear response at the 4-component relativistic level: Application to the frequency-dependent dipole polarizabilities of the coinage metal dimers. *J. Chem. Phys.*, 118(2):522–536, 2003.
- [61] Kenneth G Dyall and Knut Fægri Jr. *Introduction to relativistic quantum chemistry*. Oxford University Press, 2007.

- [62] David Young. *Computational chemistry: a practical guide for applying techniques to real world problems*. John Wiley & Sons, 2004.
- [63] M Seth, KG Dyall, R Shepard, and A Wagner. The calculation of ff spectra of lanthanide and actinide ions by the mdcf-ci method. *J. Phys. B: At. Mol. Opt. Phys.*, 34(12):2383, 2001.
- [64] Ivan Infante, André Severo Pereira Gomes, and Lucas Visscher. On the performance of the intermediate hamiltonian fock-space coupled-cluster method on linear triatomic molecules: The electronic spectra of np o 2+, np o 2 2+, and pu o 2 2+. *J. Chem. Phys.*, 125(7):074301, 2006.
- [65] Ivan Infante, Ephraim Eliav, Marius J Vilkas, Yasuyuki Ishikawa, Uzi Kaldor, and Lucas Visscher. A fock space coupled cluster study on the electronic structure of the uo 2, uo 2+, u 4+, and u 5+ species. *J. Chem. Phys.*, 127(12):124308, 2007.
- [66] Jochen Autschbach. Perspective: relativistic effects. *J. Chem. Phys.*, 136(15):150902, 2012.
- [67] Wenjian Liu. Ideas of relativistic quantum chemistry. *Mol. Phys.*, 108(13):1679–1706, 2010.
- [68] Trond Saue. Relativistic hamiltonians for chemistry: a primer. *ChemPhysChem*, 12(17):3077–3094, 2011.
- [69] Leslie L Foldy and Siegfried A Wouthuysen. On the dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys. Rev.*, 78(1):29, 1950.
- [70] Andre Severo Pereira Gomes. *On the calculation of molecular properties of heavy element systems with ab initio approaches: from gas-phase to complex systems*. PhD thesis, Universite de Lille, 2016.
- [71] Erik van Lenthe, Evert-Jan Baerends, and Jaap G Snijders. Relativistic total energy using regular approximations. *J. Chem. Phys.*, 101(11):9783–9792, 1994.
- [72] Ch Chang, M Pelissier, and Ph Durand. Regular two-component pauli-like effective hamiltonians in dirac theory. *Phys. Scr.*, 34(5):394, 1986.
- [73] Christoph van Wüllen. Molecular density functional calculations in the regular relativistic approximation: Method, application to coinage metal diatomics, hydrides, fluorides and chlorides, and comparison with first-order relativistic calculations. *J. Chem. Phys.*, 109(2):392–399, 1998.
- [74] Kenneth G Dyall. An exact separation of the spin-free and spin-dependent terms of the dirac-coulomb-breit hamiltonian. *J. Chem. Phys.*, 100(3):2118–2127, 1994.
- [75] Jetze Sikkema, Lucas Visscher, Trond Saue, and Miroslav Iliaš. The molecular mean-field approach for correlated relativistic calculations. *J. Chem. Phys.*, 131(12):124116, 2009.
- [76] Takahito Nakajima and Kimihiko Hirao. The douglas-kroll-hess approach. *Chem. Rev.*, 112(1):385–402, 2011.

- [77] Paweł Tecmer, André Severo Pereira Gomes, Stefan Knecht, and Lucas Visscher. Communication: Relativistic fock-space coupled cluster study of small building blocks of larger uranium complexes, 2014.
- [78] Trond Saue and Lucas Visscher. Relativistic all-electron approaches to the study of f element chemistry. *Computational methods in lanthanide and actinide chemistry*, pages 55–87, 2015.
- [79] Johan Åqvist and Arieh Warshel. Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem. Rev.*, 93(7):2523–2544, 1993.
- [80] Laura E Ratcliff, Stephan Mohr, Georg Huhs, Thierry Deutsch, Michel Masella, and Luigi Genovese. Challenges in large scale quantum mechanical calculations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 7(1):e1290, 2017.
- [81] Jensen F. *Introduction to Computational Chemistry*. John Wiley & Sons, second edition, 2007.
- [82] Benedetta Mennucci and Jacopo Tomasi. Continuum solvation models: A new approach to the problem of solutes charge distribution and cavity boundaries. *J. Chem. Phys.*, 106(12):5151–5158, 1997.
- [83] A. Klamt and G. Schramm. Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2*, pages 799–805, 1993.
- [84] Anders Osted, Jacob Kongsted, Kurt V Mikkelsen, Per-Olof Åstrand, and Ove Christiansen. Statistical mechanically averaged molecular properties of liquid water calculated using the combined coupled cluster/molecular dynamics method. *J. Chem. Phys.*, 124(12):124503, 2006.
- [85] Benedetta Mennucci and Roberto Cammi. *Continuum solvation models in chemical physics: from theory to applications*. John Wiley & Sons, 2008.
- [86] Andreas Klamt and GJGJ Schüürmann. Cosmo: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2*, (5):799–805, 1993.
- [87] Andreas Klamt and Volker Jonas. Treatment of the outlying charge in continuum solvation models. *J. Chem. Phys.*, 105(22):9972–9981, 1996.
- [88] Ansgar Schäfer, Andreas Klamt, Diana Sattel, John CW Lohrenz, and Frank Eckert. Cosmo implementation in turbomole: Extension of an efficient quantum chemical code towards liquid systems. *PCCP*, 2(10):2187–2193, 2000.
- [89] Andreas Klamt, Volker Jonas, Thorsten Bürger, and John CW Lohrenz. Refinement and parametrization of cosmo-rs. *J. Phys. Chem. A*, 102(26):5074–5085, 1998.
- [90] Gregory Voth. *Computational approaches for studying enzyme mechanism*, volume 577. Academic Press, 2016.
- [91] Arieh Warshel and Michael Levitt. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103(2):227–249, 1976.

- [92] George A Kaminski and William L Jorgensen. A quantum mechanical and molecular mechanical method based on cm1a charges: applications to solvent effects on organic equilibria and reactions. *J. Phys. Chem. B*, 102(10):1787–1796, 1998.
- [93] Loup Verlet. Computer” experiments” on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, 159(1):98, 1967.
- [94] Gerrit Groenhof. Introduction to qm/mm simulations. In *Biomolecular Simulations*, pages 43–66. Springer, 2013.
- [95] Xavier Assfeld and Jean-Louis Rivail. Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method. *Chem. Phys. Lett.*, 263(1-2):100–106, 1996.
- [96] Jiali Gao, Patricia Amara, Cristobal Alhambra, and Martin J Field. A generalized hybrid orbital (gho) method for the treatment of boundary atoms in combined qm/mm calculations. *J. Phys. Chem. A*, 102(24):4714–4721, 1998.
- [97] Alejandro Rodríguez, Carolina Oliva, Miguel Gonzalez, Marc van der Kamp, and Adrian J Mulholland. Comparison of different quantum mechanical/molecular mechanics boundary treatments in the reaction of the hepatitis c virus ns3 protease with the ns5a/5b substrate. *J. Phys. Chem. B*, 111(44):12909–12915, 2007.
- [98] R Mark Nicoll, Sally A Hindle, Grant MacKenzie, Ian H Hillier, and Neil A Burton. Quantum mechanical/molecular mechanical methods and the study of kinetic isotope effects: modelling the covalent junction region and application to the enzyme xylose isomerase. *Theor. Chem. Acc.*, 106(1-2):105–112, 2001.
- [99] Hans Martin Senn and Walter Thiel. Qm/mm methods for biomolecular systems. *Angew. Chem. Int. Ed.*, 48(7):1198–1229, 2009.
- [100] Tomasz Adam Wesolowski and Arieh Warshel. Frozen density functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.*, 97(30):8050–8053, 1993.
- [101] Martin Fuchs and Matthias Scheffler. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory. *Comput. Phys. Commun.*, 119(1):67–98, 1999.
- [102] Tomasz Adam Wesolowski. Density functional theory with approximate kinetic energy functionals applied to hydrogen bonds. *J. Chem. Phys.*, 106(20):8516–8526, 1997.
- [103] Tomasz Adam Wesolowski, Henry Chermette, and Jacques Weber. Accuracy of approximate kinetic energy functionals in the model of kohn–sham equations with constrained electron density: The fh nch complex as a test case. *J. Chem. Phys.*, 105(20):9182–9190, 1996.
- [104] Tomasz Adam Wesolowski and Jacques Weber. Kohn-sham equations with constrained electron density: The effect of various kinetic energy functional parametrizations on the ground-state molecular properties. *Int. J. Quantum Chem.*, 61(2):303–311, 1997.
- [105] Pietro Cortona. Self-consistently determined properties of solids without band-structure calculations. *Phys. Rev. B*, 44(16):8454, 1991.

- [106] Peter Elliott, Kieron Burke, Morrel H Cohen, and Adam Wasserman. Partition density-functional theory. *Phys. Rev. A*, 82(2):024501, 2010.
- [107] Morrel H Cohen and Adam Wasserman. On the foundations of chemical reactivity theory. *J. Phys. Chem. A*, 111(11):2229–2242, 2007.
- [108] Morrel H Cohen and Adam Wasserman. On hardness and electronegativity equalization in chemical reactivity theory. *J. Stat. Phys.*, 125(5-6):1121–1139, 2006.
- [109] Morrel H Cohen, Adam Wasserman, and Kieron Burke. Partition theory: A very simple illustration. *J. Phys. Chem. A*, 111(49):12447–12453, 2007.
- [110] Morrel H Cohen, Adam Wasserman, Roberto Car, and Kieron Burke. Charge transfer in partition theory. *J. Phys. Chem. A*, 113(10):2183–2192, 2009.
- [111] Chen Huang, Michele Pavone, and Emily A Carter. Quantum mechanical embedding theory based on a unique embedding potential. *J. Chem. Phys.*, 134(15):154110, 2011.
- [112] Octavio Roncero, Alexandre Zanchet, Pablo Villarreal, and Alfredo Aguado. A density-division embedding potential inversion technique. *J. Chem. Phys.*, 131(23):234110, 2009.
- [113] André Severo Pereira Gomes and Christoph R Jacob. Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. *Annu. Rep. R. Soc. Chem. Sect. C. Phys. Chem*, 108:222–277, 2012.
- [114] Patrick Huang and Emily A Carter. Advances in correlated electronic structure methods for solids, surfaces, and nanostructures. *Annu. Rev. Phys. Chem.*, 59:261–290, 2008.
- [115] Niranjana Govind, Yan Alexander Wang, and Emily A Carter. Electronic-structure calculations by first-principles density-based embedding of explicitly correlated systems. *J. Chem. Phys.*, 110(16):7677–7688, 1999.
- [116] Christoph R Jacob and Lucas Visscher. Calculation of nuclear magnetic resonance shieldings using frozen-density embedding. *J. Chem. Phys.*, 125(19):194104, 2006.
- [117] Andreas W Götz, Jochen Autschbach, and Lucas Visscher. Calculation of nuclear spin-spin coupling constants using frozen density embedding. *J. Chem. Phys.*, 140(10):104107, 2014.
- [118] Tomasz A Wesolowski, Sapana Shedge, and Xiuwen Zhou. Frozen-density embedding strategy for multilevel simulations of electronic structure. *Chem. Rev.*, 115(12):5891–5928, 2015.
- [119] Jason D Goodpaster, Taylor A Barnes, Frederick R Manby, and Thomas F Miller III. Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.*, 140(18):18A507, 2014.
- [120] Yizhak Marcus. Effect of ions on the structure of water: structure making and breaking. *Chem. Rev.*, 109(3):1346–1370, 2009.
- [121] CJ Fecko, JD Eaves, JJ Loparo, A Tokmakoff, and PL Geissler. Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water. *Science*, 301(5640):1698–1702, 2003.

- [122] John B Asbury, Tobias Steinell, Kyungwon Kwak, SA Corcelli, CP Lawrence, JL Skinner, and MD Fayer. Dynamics of water probed with vibrational echo correlation spectroscopy. *J. Chem. Phys.*, 121(24):12431–12446, 2004.
- [123] Nils Huse, Satoshi Ashihara, Erik TJ Nibbering, and Thomas Elsaesser. Ultrafast vibrational relaxation of o–h bending and librational excitations in liquid h₂o. *Chem. Phys. Lett.*, 404(4-6):389–393, 2005.
- [124] ML Cowan, Barry D Bruner, N Huse, JR Dwyer, B Chugh, ETJ Nibbering, T Elsaesser, and RJD Miller. Ultrafast memory loss and energy redistribution in the hydrogen bond network of liquid h₂o. *Nature*, 434(7030):199, 2005.
- [125] Sander Woutersen and Huib J Bakker. Resonant intermolecular transfer of vibrational energy in liquid water. *Nature*, 402(6761):507, 1999.
- [126] Ralph Jimenez, Graham R Fleming, PV Kumar, and Mark Maroncelli. Femtosecond solvation dynamics of water. *Nature*, 369(6480):471, 1994.
- [127] Nenad Trinajstić. *Chemical graph theory*. Routledge, 2018.
- [128] William Jorgensen. *The organic chemist’s book of orbitals*. Elsevier, 2012.
- [129] Bernd Winter and Manfred Faubel. Photoemission from liquid aqueous solutions. *Chem. Rev.*, 106(4):1176–1211, 2006.
- [130] MS Banna, BH McQuaide, R Malutzki, and V Schmidt. The photoelectron spectrum of water in the 30 to 140 ev photon energy range. *J. Chem. Phys.*, 84(9):4739–4744, 1986.
- [131] B Winter, R Weber, W Widdra, M Dittmar, M Faubel, and IV Hertel. Full valence band photoemission from liquid water using euv synchrotron radiation. *J. Phys. Chem. A*, 108(14):2625–2632, 2004.
- [132] Greg Hura, Daniela Russo, Robert M Glaeser, Teresa Head-Gordon, Matthias Krack, and Michele Parrinello. Water structure as a function of temperature from x-ray scattering experiments and ab initio molecular dynamics. *PCCP*, 5(10):1981–1991, 2003.
- [133] Naoya Kurahashi, Shutaro Karashima, Ying Tang, Takuya Horio, Bumaliya Abulimiti, Yoshi-Ichi Suzuki, Yoshihiro Ogi, Masaki Oura, and Toshinori Suzuki. Photoelectron spectroscopy of aqueous solutions: Streaming potentials of na_x (x= cl, br, and i) solutions and electron binding energies of liquid water and x-. *J. Chem. Phys.*, 140(17):174506, 2014.
- [134] Florent Réal, André Severo Pereira Gomes, Yansel Omar Guerrero Martínez, Tahra Ayed, Nicolas Galland, Michel Masella, and Valérie Vallet. Structural, dynamical, and transport properties of the hydrated halides: How do at- bulk properties compare with those of the other halides, from f- to i-? *J. Chem. Phys.*, 144(12):124513, 2016.
- [135] Gil Markovich, Stuart Pollack, Rina Giniger, and Ori Cheshnovsky. Photoelectron spectroscopy of cl-, br-, and i- solvated in water clusters. *J. Chem. Phys.*, 101(11):9344–9353, 1994.
- [136] Gil Markovich, Stuart Pollack, Rina Giniger, and Ori Cheshnovsky. The solvation of iodine anions in water clusters: Pes studies. *Z. Phys. A*, 26(1):98–100, 1993.

- [137] Pavel Jungwirth and Douglas J Tobias. Specific ion effects at the air/water interface. *Chem. Rev.*, 106(4):1259–1281, 2006.
- [138] Marcel D Baer, I-Feng William Kuo, Hendrik Bluhm, and Sutapa Ghosal. Interfacial behavior of perchlorate versus chloride ions in aqueous solutions. *J. Phys. Chem. B*, 113(48):15843–15850, 2009.
- [139] Florent Réal, Valérie Vallet, and Michel Masella. Improving the description of solvent pairwise interactions using local solute/solvent three-body functions. the case of halides and carboxylates in aqueous environment. *J. Comput. Chem.*, 40(11):1209–1218, 2019.
- [140] Alessandro Genova, Davide Ceresoli, and Michele Pavanello. Avoiding fractional electrons in subsystem dft based ab-initio molecular dynamics yields accurate models for liquid water and solvated oh radical. *J. Chem. Phys.*, 144(23):234105, 2016.
- [141] AF Orchard and G Thornton. A study of final state structure in the x-pe spectra of the rare-earth oxides part ii: Ionisation of a 4f electron. *J. Electron. Spectrosc. Relat. Phenom.*, 10(1):1–14, 1977.
- [142] AF Orchard and G Thornton. A study of final state structure in the x-pe spectra of the rare earth oxides, part iii: Ionisation of a 4d electron. *J. Electron. Spectrosc. Relat. Phenom.*, 13(1):27–38, 1978.
- [143] Anthony Gordon Lee. *chemistry of thallium*. Elsevier Pub. Co., 1971.
- [144] Alessandro Ceriotti, Paolo Chini, Roberto Della Pergola, and Giuliano Longoni. Carbonylnickelates. 3. synthesis and chemical characterization of the $[ni_{12}(co)_{21}h_{4-n}]_{n(n=2, 3, 4)}$ clusters. *Inorg. Chem.*, 22(11):1595–1598, 1983.
- [145] Colin J Schaverien, John C Dewan, and Richard R Schrock. Multiple metal-carbon bonds. 43. well-characterized, highly active, lewis acid free olefin metathesis catalysts. *J. Am. Chem. Soc.*, 108(10):2771–2773, 1986.
- [146] Dennis J Underwood, Roald Hoffmann, Kazuyuki Tatsumi, Akira Nakamura, and Yasuhiro Yamamoto. Triangular platinum and nickel clusters: the” tinker-toy” construction of chains with high nuclearity. *J. Am. Chem. Soc.*, 107(21):5968–5980, 1985.
- [147] Vincent J Catalano, Byron L Bennett, Stamatis Muratidis, and Bruce C Noll. Unsupported pt (0)- tl (i) bonds in the simple $[pt(pph_2py)_3tl]^+$ complexes. *J. Am. Chem. Soc.*, 123(1):173–174, 2001.
- [148] Mikhail Maliarik, Julius Glaser, Imre Tóth, Mateus Webba da Silva, and László Zékány. A new class of oligonuclear platinum-thallium compounds with a direct metal- metal bond-3. unusual equilibria in aqueous solution. *Eur. J. Inorg. Chem.*, 1998(5):565–570, 1998.
- [149] Katja E Berg, Julius Glaser, Michael C Read, and Imre Toth. Nonbuttressed metal-metal-bonded complexes of platinum and thallium in aqueous solution: Characterization of $[(nc)_5pt-tl(cn)]$ -by multinuclear nmr. *J. Am. Chem. Soc.*, 117(28):7550–7551, 1995.
- [150] Jeffrey K Nagle, Alan L Balch, and Marilyn M Olmstead. $Tl_2pt(cn)_4$: a non-columnar, luminescent form of pt (cn)₄-containing platinum-thallium bonds. *J. Am. Chem. Soc.*, 110(1):319–321, 1988.

- [151] Alan L Balch and Steven P Rowley. Solubilizing the thallium-platinum unit of $\text{tl}_2\text{pt}(\text{cn})_4$. preparation and use of a new crown ether/phosphine hybrid ligand for linking main-group and transition-metal ions. *J. Am. Chem. Soc.*, 112(16):6139–6140, 1990.
- [152] Oliver Renn, Bernhard Lippert, and Ilpo Mutikainen. A mixed pt_2tl complex of 1-methylthymine: structural evidence for the stereoactivity of the tli electron lone pair and unexpected intracomplex nucleobase stacking. *Inorg. Chim. Acta*, 208(2):219–223, 1993.
- [153] Rafael Uson, Juan Fornies, Milagros Tomas, Raquel Garde, and Pablo J Alonso. Synthesis and structure of $(\text{nbu}_4)_2[\text{tl}\{\text{Pt}(\text{C}_6\text{F}_5)_4\}_2]$, the first paramagnetic compound containing thallium (ii). *J. Am. Chem. Soc.*, 117(6):1837–1838, 1995.
- [154] Leijun Hao, Jagadese J Vittal, and Richard J Puddephatt. Platinum cluster cryptates: trapping of mercury (0) and thallium (i). *Inorg. Chem.*, 35(2):269–270, 1996.
- [155] Cynthia J Jameson. Spin-spin coupling. In *Multinuclear NMR*, pages 89–131. Springer, 1987.
- [156] Guibin Ma. *On Thallium (III) and binuclear platinum-thallium complexes with N-donor ligands in solution and in solid*. PhD thesis, Kemi, 2001.
- [157] Jochen Autschbach and Boris Le Guennic. Solvent effects on ^{195}pt and ^{205}tl nmr chemical shifts of the complexes $[(\text{nc})_5\text{pt}\text{tl}(\text{cn})_n]^{n-(n=0-3)}$, and $[(\text{nc})_5\text{pt}\text{tl}\text{pt}(\text{cn})_5]^{3-}$ studied by relativistic density functional theory. *ChemPlusChem*, 10(10):2581–2589, 2004.
- [158] Boris Le Guennic, Kazuko Matsumoto, and Jochen Autschbach. Nmr properties of platinum thallium bonded complexes: analysis of relativistic density functional theory results. *Magn. Reson. Chem.*, 42(S1):S99–S116, 2004.
- [159] Lucas C Ducati, Alex Marchenko, and Jochen Autschbach. Nmr j-coupling constants of tl-pt bonded metal complexes in aqueous solution: Ab initio molecular dynamics and localized orbital analysis. *Inorg. Chem.*, 55(22):12011–12023, 2016.
- [160] Richard Car and Mark Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55(22):2471, 1985.
- [161] Jochen Autschbach and Monika Srebro. Delocalization error and functional tuning in kohn–sham calculations of molecular properties. *Acc. Chem. Res.*, 47(8):2592–2602, 2014.
- [162] Andre Severo Pereira Gomes, Christoph R Jacob, and Lucas Visscher. Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory. *PCCP*, 10(35):5353–5362, 2008.
- [163] Norman Neill Greenwood and Alan Earnshaw. *Chemistry of the Elements*. Elsevier, 2012.
- [164] Frank Weinhold and Clark R Landis. *Valency and bonding: a natural bond orbital donor-acceptor perspective*. Cambridge University Press, 2005.
- [165] RE Rundle. Electron deficient compounds1. *J. Am. Chem. Soc.*, 69(6):1327–1331, 1947.
- [166] RE Rundle. Electron deficient compounds. ii. relative energies of “half-bonds”. *J. Chem. Phys.*, 17(8):671–675, 1949.
- [167] Karol Jackowski and Włodzimierz Makulski. ^{13}c shielding scale for mas nmr spectroscopy. *Magn. Reson. Chem.*, 49(9):600–602, 2011.