



HAL
open science

La méthode Equation of Motion Coupled Cluster pour la modélisation des états excités et propriétés des molécules contenant des éléments lourds

Loïc Halbert

► **To cite this version:**

Loïc Halbert. La méthode Equation of Motion Coupled Cluster pour la modélisation des états excités et propriétés des molécules contenant des éléments lourds. Chimie-Physique [physics.chem-ph]. Université de Lille, 2021. Français. NNT : 2021LILUR038 . tel-03848132

HAL Id: tel-03848132

<https://theses.hal.science/tel-03848132>

Submitted on 10 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE LILLE

École doctorale Ed 104 - Sciences de la Matière, du Rayonnement et de l'Environnement
Unité de recherche Laboratoire PhLAM

Thèse présentée par **Loïc HALBERT**

Soutenue le **10 novembre 2021**

En vue de l'obtention du grade de docteur de l'Université de Lille

Discipline **Physique des milieux dilués et optique fondamentale**

Spécialité **Chimie Physique Théorique**

La méthode Equation of Motion Coupled Cluster pour la modélisation des états excités et propriétés des molécules contenant des éléments lourds

Thèse dirigée par André SEVERO PEREIRA GOMES

Composition du jury

<i>Rapporteurs</i>	Rémi MAURICE	chargé de recherche au CNRS (HDR) / Univ. Rennes
	Thomas JAGAU	professeur / Katholieke Univer- siteit Leuven
<i>Examineurs</i>	Valérie VALLET	directrice de recherche présidente du jury CNRS / Univ
	Trond SAUE	directeur de recherche CNRS / Univ. Toulouse
	Emmanuel FROMAGER	professeur / Univ. Strasbourg
<i>Directeur de thèse</i>	André SEVERO PEREIRA GOMES	chargé de recherche au CNRS (HDR) / Univ. Lille

UNIVERSITÉ DE LILLE

École doctorale Ed 104 - Sciences de la Matière, du Rayonnement et de l'Environnement
Unité de recherche Laboratoire PhLAM

Thèse présentée par **Loïc HALBERT**

Soutenue le **10 novembre 2021**

En vue de l'obtention du grade de docteur de l'Université de Lille

Discipline **Physique des milieux dilués et optique fondamentale**

Spécialité **Chimie Physique Théorique**

La méthode Equation of Motion Coupled Cluster pour la modélisation des états excités et propriétés des molécules contenant des éléments lourds

Thèse dirigée par André SEVERO PEREIRA GOMES

Composition du jury

<i>Rapporteurs</i>	Rémi MAURICE	chargé de recherche au CNRS (HDR) / Univ. Rennes
	Thomas JAGAU	professeur / Katholieke Univer- siteit Leuven
<i>Examineurs</i>	Valérie VALLET	directrice de recherche présidente du jury CNRS / Univ
	Trond SAUE	directeur de recherche CNRS / Univ. Toulouse
	Emmanuel FROMAGER	professeur / Univ. Strasbourg
<i>Directeur de thèse</i>	André SEVERO PEREIRA GOMES	chargé de recherche au CNRS (HDR) / Univ. Lille

UNIVERSITÉ DE LILLE

Doctoral School Ed 104 - Sciences de la Matière, du Rayonnement et de l'Environnement

University Department Laboratoire PhLAM

Thesis defended by Loïc HALBERT

Defended on 10th November, 2021

In order to become Doctor from Université de Lille

Academic Field **Physics - diluted medium and fundamental optics**

Speciality **Theoretical chemical physics**

The Equation of Motion Coupled Cluster method for modeling excited states and properties of molecules containing heavy elements

Thesis supervised by André SEVERO PEREIRA GOMES

Committee members

<i>Referees</i>	Rémi MAURICE	chargé de recherche au CNRS (HDR) / Univ. Rennes
	Thomas JAGAU	professor / Katholieke Univer- siteit Leuven
<i>Examiners</i>	Valérie VALLET	directrice de recherche Committee President CNRS / Univ
	Trond SAUE	directeur de recherche CNRS / Univ. Toulouse
	Emmanuel FROMAGER	professor / Univ. Strasbourg
<i>Supervisor</i>	André SEVERO PEREIRA GOMES	chargé de recherche au CNRS (HDR) / Univ. Lille

Mots clés: chimie quantique, éléments lourds, structure électronique,
théorie à n corps

Keywords: quantum chemistry, heavy elements, electronic structure,
many-body theory

Cette thèse a été préparée au

Laboratoire PhLAM
CNRS UMR 8523
Université de Lille,
F-59655 Villeneuve d'Ascq Cedex
France



À Marie - À Mathieu

Ce qu'on n'a jamais su, il est possible, un jour, d'arriver à le savoir - ce qui est inenvisageable avec ce qu'on ne saura jamais - Sache ça.

Philippe Geluck

Si nous avons du jambon, nous pourrions faire des œufs au jambon si nous avions des œufs.

Frank Herbert

Si tu as une pomme, que j'ai une pomme, et que l'on échange nos pommes, nous aurons chacun une pomme. Mais si tu as une idée, que j'ai une idée et que l'on échange nos idées, nous aurons chacun deux idées.

George Bernard Shaw

LA MÉTHODE EQUATION OF MOTION COUPLED CLUSTER POUR LA MODÉLISATION DES ÉTATS EXCITÉS ET PROPRIÉTÉS DES MOLÉCULES CONTENANT DES ÉLÉMENTS LOURDS**Résumé**

Dans cette thèse, nous cherchons à obtenir certaines propriétés moléculaires pour des espèces contenant des éléments lourds ou présentant des intérêts atmosphériques. Pour cela, nous utilisons des techniques permettant de caractériser les électrons de cœur, avec les potentiels d'ionisation (IP) ou avec les énergies d'excitation (EE), offrant la possibilité par exemple d'interpréter respectivement les expériences *X-ray Photoelectron Spectroscopy* (XPS) et *X-ray Absorption Spectroscopy* (XAS). Nous cherchons également à caractériser les électrons de valence au travers de la polarisabilité qui est utilisée par exemple pour développer des champs de force. Quand nous travaillons avec des éléments lourds ou avec des électrons de cœur, il faut prendre en compte les effets relativistes. Nous avons donc employé l'hamiltonien de Dirac-Coulomb(-Gaunt). De plus, pour comparer nos résultats aux expériences, il nous faut des méthodes précises. Ainsi, nous travaillerons avec la méthode *Coupled-Cluster* (CC) et pour obtenir les (IP), les (EE) et également les affinités électroniques (EA), nous utiliserons *Equation of Motion Coupled-Cluster* (EOM-CC). Cependant, ces deux éléments (hamiltoniens à 4-composantes et méthodes post-Hartree-Fock) impliquent des coûts de calcul considérables, nécessitant les ressources de plateformes *High Performance Computing* (HPC).

Cette thèse se présentera donc selon les éléments décrits précédemment. Premièrement, nous étudierons la méthode *Core-Valence Separation* (CVS) qui nous permettra, à partir de EOM-CC, d'atteindre les propriétés des électrons de cœur (IP et EE). Comme ces électrons sont proches du noyau où les effets relativistes sont les plus importants, nous étudierons différents hamiltoniens, notamment *exact two-component molecular mean field Hamiltonian*. Deuxièmement, nous nous intéresserons aux approximations perturbatives (*Partitioned* et *Many Body Perturbation Theory 2d order* (MBPT(2)) à appliquer à la matrice EOM-CC pour limiter les coûts computationnels. Enfin, nous présenterons des travaux réalisés sur ExaCorr, une nouvelle implémentation de Coupled-Cluster relativiste pour les architectures hybrides et massivement parallèles. Nous terminerons en décrivant le formalisme et les équations de travail de la méthode *Linear Response Coupled-Cluster* (LR-CC), grâce à laquelle des polarisabilités moléculaires analytiques (dépendantes de la fréquence) peuvent être obtenues.

Mots clés : chimie quantique, éléments lourds, structure électronique, théorie à n corps

Laboratoire PhLAM

CNRS UMR 8523 – Université de Lille, – F-59655 Villeneuve d'Ascq Cedex – France

THE EQUATION OF MOTION COUPLED CLUSTER METHOD FOR MODELING EXCITED STATES AND PROPERTIES OF MOLECULES CONTAINING HEAVY ELEMENTS**Abstract**

In this thesis, we seek to obtain certain molecular properties for species containing heavy elements or presenting atmospheric interests. For this, we use techniques to characterize the core electrons, with ionization potentials (IP) or with excitation energies (EE), allowing for example to respectively interpret *X-ray Photoelectron Spectroscopy* (XPS) and *X-ray Absorption Spectroscopy* (XAS). We also seek to characterize valence electrons through the polarizability, which is used for example to develop force fields. When we work with heavy elements or with core electrons, we must take relativistic effects into account. We therefore used the Dirac-Coulomb(-Gaunt) Hamiltonian. Furthermore, to compare our results with experiments, we need precise methods. Thus, we will work with the *Coupled-Cluster* (CC) method, and will use the *Equation of Motion Coupled-Cluster* (EOM-CC) method to obtain the IPs, EEs and electron affinities (EA). However, these two elements (4-component Hamiltonians and post-Hartree-Fock methods) imply considerable computational costs, requiring the resources of *High Performance Computing* (HPC) platforms.

This thesis presents a study of the *Core-Valence Separation* (CVS) method, which will allow us to reach the properties of core electrons (IP and EE) with EOM-CC. We provide a detailed investigation of the performance of different Hamiltonians, in particular the *exact two-component molecular mean field* Hamiltonian. Second, we will focus on the perturbative approximations (*Partitioned* and *Many Body Perturbation Theory 2d order* (MBPT(2))) to be applied to the EOM-CC matrix to limit computational costs, including for core processes. Finally, we present the work carried out in ExaCorr, a new relativistic coupled cluster implementation for hybrid and massively parallel architectures. We will finish by outlining the formalism and working equations for the *Linear Response Coupled-Cluster* (LR-CC) method, through which analytical (frequency-dependent) molecular polarizabilities can be obtained.

Keywords: quantum chemistry, heavy elements, electronic structure, many-body theory

Remerciements

Je voudrais remercier l'ensemble de l'équipe PCMT pour son accueil. Je les remercie également pour le soutien qu'ils m'ont apporté quand il m'a fallu reprendre un Master II. Merci à Florent Réal pour les nombreuses conversations, ses conseils judicieux et son aide précieuse. Je voudrais adresser également mes remerciements à Valérie Vallet ; merci d'avoir toujours répondu à toutes mes questions aussi simplistes fussent-elles, et d'avoir toujours su être aussi disponible et pédagogue.

Je devine à quel point il est peu évident d'accompagner un thésard qui assume 500 heures de cours par an et qui a, en plus, la charge de quelques responsabilités administratives, un grand merci à André Gomes...

Un salut amical à M. Abdel Hadi Kassiba ; j'ai mis 25 ans entre ma première année universitaire et la rédaction de cette thèse, mais le travail est là.

A mon âge, j'ai tellement de monde à remercier encore... Pour ceux qui me sont proches. Merci à Dom, mon meilleur ami, un formidable allié comme j'en souhaite à tous.

J'ai une pensée particulière pour ma famille, mes parents François et Dominique et mes sœurs Sabine et Laurence. Merci également à ma belle-famille Nat, Seb, Brigitte, merci pour tout. Une pensée également pour José qui m'a appris que même si « il n'y a que Kate Moss qui est éternelle », l'utilisation du verbe « impose » nous oblige à utiliser le subjonctif et quoi qu'il en soit, « *Frozen Core* impose que les amplitudes soient nulles », même si je vous assure qu'elles sont nulles !

Enfin, merci Mathieu, mon fils, de me faire comprendre la relativité des choses. Mon seul regret reste de ne pas t'avoir fait comprendre l'importance de l'approche diagrammatique, mais il semblerait qu'à 9 ans on ait d'autres préoccupations...

Mais par-dessus tout, et si je ne pouvais remercier qu'une seule personne ; je voudrais remercier Marie, ma femme, pour son aide, ses conseils toujours avisés et son indéfectible soutien.

Sommaire

Résumé	xv
Remerciements	xvii
Sommaire	xix
Liste des tableaux	xxiii
Table des figures	xxv
Introduction	1
1 Méthodes	5
1.1 Méthode HARTREE-FOCK	5
1.1.1 Hamiltonien non-relativiste	5
1.1.2 Fonction d'onde	7
1.1.3 Énergie HARTREE-FOCK	8
1.1.4 Expansion en fonction de base, Basis Set	9
1.1.5 Size-Consistent Size-Extensive	11
1.2 La Méthode full Configuration Interaction fCI	11
1.2.1 Énergie de corrélation	12
1.2.2 Méthodes multi-références (MR)	13
1.3 Théorie perturbative	15
1.3.1 RAYLEIGH-SCHRÖDINGER <i>Perturbation Theory</i>	15
1.3.2 MÖLLER-PLESSET <i>Perturbation Theory</i>	16
1.4 Méthode Coupled-Cluster	17
1.4.1 Énergie et Amplitudes	17
1.4.2 Approche diagrammatique	20
1.5 Equation of Motion Coupled-Cluster	22
1.5.1 Description générale	23

1.5.2	Énergie d'excitation, Potentiel d'ionisation, Affinité électronique	24
1.5.3	Matrice EOM-CCSD	25
1.5.4	EOM-CC et l'algorithme de Davidson	26
1.6	Relativité restreinte	28
1.6.1	Solution pour un système mono-électronique	29
1.6.2	Solution pour un système poly-électronique	31
1.6.3	Hamiltonien de BREIT - Électrodynamique Quantique	32
1.6.4	Approximation à 2-composantes de l'hamiltonien du Dirac-Coulomb : ${}^2\text{DC} - {}^2\text{DC}^{\text{M}}$	33
2	Les différentes problématiques	35
	PROBLÉMATIQUE 1 : OBTENIR LES IP ET LES EE DE CŒUR	35
2.1	<i>X-ray Photoelectron Spectroscopy (XPS)</i>	36
2.1.1	Présentation d'une expérience XPS	36
2.1.2	Expérimentalement	37
2.2	Chimie quantique et potentiels d'ionisation des électrons de cœur	40
2.2.1	Lien entre les calculs et les expériences XPS	40
2.2.2	Classes de méthode	41
2.2.3	Vers le chapitre 3	44
	PROBLÉMATIQUE 2 : QUELLES APPROXIMATIONS RÉALISER?	45
2.3	Approximations	45
2.3.1	Approximations dans l'origine des intégrales	45
2.3.2	Approximations perturbatives des termes de la matrice \mathcal{H}_{eff}	47
2.3.3	Vers le chapitre 4	49
	PROBLÉMATIQUE 3 : EXPLOITER DE FAÇON EFFICACE DES PLATEFORMES HPC POUR DES CALCULS CORRÉLÉS	50
2.4	High-Performance Computing (HPC)	50
2.4.1	Hardware	51
2.4.2	Software	53
2.4.3	Vers le chapitre 5	55
3	EOM-(IP; EE; EA) - CVS	57
3.1	Utilisation de CVS avec EOM-CC	58
3.2	Options pour CVS : <i>Frozen Core (FC)</i> et <i>No-Double (ND)</i>	60
3.3	Influence de l'hamiltonien	64
3.4	Conclusion	67
4	{P}-EOM-CCSD-{MBPT(2)}	69
4.1	Les différentes approximations	70
4.2	Détails computationnels et présentation des résultats	72

4.3	Potentiels d'ionisation	73
4.3.1	Triiodure I_3^-	73
4.3.2	Les ions hypohalogénites XO^- , $X \in \{Cl; Br; I; At; Ts\}$. . .	74
4.4	Énergie d'excitation et affinité électronique	86
4.5	Conclusion	89
4.6	Remarques sur les coûts computationnels	91
5	ExaCorr	93
5.1	Une nouvelle bibliothèque : ExaTensor	94
5.2	Exemple d'expression informatique pour les calculs tensoriels .	95
5.3	Résultats sur Summit	96
5.3.1	La symétrie	96
5.3.2	Complexe d'uranyl : le trinitrate d'uranyl $[UO_2(NO_3)_3]^-$.	97
5.4	CC2	99
5.4.1	Coût computationnel	100
5.4.2	Résultats préliminaires	101
5.4.3	Travaux en cours de réalisation et perspectives vers ExaTensor	104
5.5	LR-CCSD - α	107
5.5.1	Polarisabilité dans le cadre de CC	108
5.5.2	Expressions programmables	111
5.6	Conclusion	114
	Conclusion	115
	Bibliographie	117
A	Équations	143
A.1	Équations CC2	143
A.2	Méthodes approximées	144
A.2.1	EOM-MBPT(2)	144
A.2.2	Partitioned-EOM	145
B	Annexes du Chapitre 3	147
B.1	Courbes d'énergie potentielle	147
B.1.1	Courbes	147
B.1.2	Tables : Position des minimums - Énergies de transition .	148
B.2	Géométrie des molécules	150
C	Article 1 : Relativistic EOM-CCSD and CVS	151
D	Article 2 : ExaCorr	169

E Diagrammes	191
Table des matières	205

Liste des tableaux

2.1	Correspondance entre la notation par couche électronique (<i>edge</i>) et la notation par orbitale.	39
2.2	Coût computationnel de quelques méthodes permettant d'obtenir les énergies d'excitation (<i>EE</i>).	49
2.3	Exemple de CPU, GPU et de FPGA.	53
2.4	Taxonomie de Flynn.	54
3.1	Représentation des différentes limites pour les approximations utilisées pour EOM-IP-CVS.	63
3.2	I^- Énergies de corrélation pour (MP2) et (CCSD) et différence ΔE_c par rapport à 4DC (en eV).	66
4.1	Potentiel d'ionisation de I_3^- en (eV).	74
4.2	Énergie de transition Δ_T entre $X^{1/2}$ - $X^{3/2}$	83
4.3	HCl - Énergie E (en eV) - Δ_f la différence entre IP- full-IP - Δ_{IP} celle entre IP_{App} -IP (en eV).	84
4.4	HBr - Énergie E (en eV) et %SI pour $3p^{-1}$ et $3d^{-1}$ - Δ_f la différence entre IP- full-IP - Δ_{IP} celle entre IP_{App} -IP (en eV).	85
4.5	CH_2I_2 - Affinité électronique en (eV) - Δ_{EA} différence entre EA_{App} -EA - Δ_{exp} la différence avec l'expérience [134].	86
4.6	CH_2IBr - Affinité électronique (en eV) - Δ_{EA} différence entre EA_{App} -EA.	87
4.7	I_3^- - Énergies d'excitation (eV)- Δ_{EE} différence entre EE_{App} -EE (en eV).	88
4.8	CH_2I_2 - Énergies d'excitation (eV)- Δ_{EE} différence entre EE_{App} -EE (en eV).	89
5.1	Durée des calculs en secondes pour des dimères $AuAr_n^+$ avec ExaCorr et RELCCSD release 2019.	97
5.2	Durée des calculs pour obtenir les amplitudes $T + \Lambda$ et Totale pour le complexe trinitrate d'uranyl $[UO_2(NO_3)_3]^-$ - X2C-AMFI.	98

5.3	H ₂ O : calcul non-relativiste, cc-pv2 ζ , comparaison avec le code MRCC et Dalton (en E_h).	102
5.4	Étude CC2 pour une chaîne linéaire d'atomes de néon.	104
5.5	Étude CC2 pour une chaîne linéaire d'atomes d'argon.	105
B.1	ClO ⁻ : Minimums des <i>Courbes d'énergie potentielle</i>	148
B.2	ClO ⁻ : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).	148
B.3	BrO ⁻ : Minimums des <i>Courbes d'énergie potentielle</i>	148
B.4	BrO ⁻ : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).	148
B.5	IO ⁻ : Minimums des <i>Courbes d'énergie potentielle</i>	149
B.6	IO ⁻ : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).	149
B.7	AtO ⁻ : Minimums des <i>Courbes d'énergie potentielle</i>	149
B.8	AtO ⁻ : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).	149
B.9	XO ⁻ - Distances internucléaires en Å.	150
B.10	CH ₂ I ₂ & CH ₂ IBr - Coordonnées cartésiennes en Å.	150

Table des figures

1.1	Représentation des coordonnées d'espace pour un système de 2 noyaux et 2 électrons.	6
1.2	Représentation diagrammatique des opérateurs d'excitation t_1 et t_2	21
1.3	Exemple de représentation diagrammatique des opérateurs $\{1-; 2-\}$ électrons.	21
1.4	Représentation diagrammatique d'une des contractions de l'équation (eq.1.47).	22
1.5	Résultat de la contraction correspondant à l'équation (eq.1.47).	22
1.6	Représentation de Feynman de la « self-energy » (a.) et de la « polarisation du vide » (b.).	33
2.1	Spectre XPS du Chlore (niveaux 1s et 2p) durant l'absorption de HCl sur la glace (eV) [106].	36
2.2	Exemple d'énergie accessible par la technologie synchrotron.	38
2.3	Spectres XPS expérimentaux du PET et du xénon.	39
2.4	Les intérêts de la <i>X-ray Photoelectron Spectroscopy (XPS)</i> dans un calcul HF.	41
2.5	Représentation des deux étapes nécessaires aux méthodes Δ HF, Δ DFT et Δ MP2.	42
2.6	Les différentes approximations de la matrice <i>Equation of Motion (EOM)</i>	48
2.7	Représentation de l'architecture hardware de Summit.	52
3.1	Spectre complet EOM-CC IP et cumul du nombre de solutions pour I^-	59
3.2	Énergies des ionisations de cœur de l'ion iodure I^- à partir de la diagonalisation complète de \tilde{H}	61
3.3	Représentation de la matrice <i>EOM</i> pour les potentiels d'ionisation (IP) pour les différentes options (CVS), (FC) et (ND).	62
3.4	Effet des différentes options appliquées à CVS pour I^-	64
3.5	Influence de l'hamiltonien sur les (IP) de cœur de l'ion iodure I^-	66

4.1	XO^- $X \in \{Cl; Br; I; At; Ts\}$. ΔE Différence entre les méthodes approximées IP_{App} (P-IP, MBPT(2) et P-MBPT(2)) et EOMCC-IP en (eV).	75
4.2	PEC pour AtO^- , IP, P-IP, MBPT(2)-IP et P-MBPT(2)-IP.	78
4.3	Différences en énergies et en positions pour les PEC - ΔE (eV) en fonction de Δr (Å) pour les 3 IP_{App} et les 4 composés, classées par chaque IP_{App}	79
4.4	Différences en énergies et en positions pour les PEC - ΔE (eV) en fonction de Δr (Å) pour les 3 IP_{App} et les 4 composés, classées par composé.	80
4.5	Erreur dans les énergies $\Delta E = \Delta IP_{App} - \Delta IP$ (eV) pour les transitions : $X^{1/2}-X^{3/2}$, $A^{3/2}-X^{3/2}$ et $A^{1/2}-X^{3/2}$, pour les différents XO^- et selon les différentes IP_{App}	82
5.1	Structure du complexe d'uranyl $[UO_2(NO_3)_3]^-$ déduite de la structure rayon-X du cristal $RbUO_2(NO_3)_3$ [14].	97
5.2	Abaque CC2.	102
5.3	Exemple de diagrammes présents dans la construction du tenseur ξ .111	
B.1	Courbes d'énergies potentielles - XO^- $X \in \{Cl; Br; I; At\}$ pour toutes les méthodes IP.	147

Introduction

Cette thèse a trait à des ions et à des molécules en phase gazeuse présentant des intérêts pour la chimie atmosphérique ou présentant des isotopes radioactifs pouvant être rejetés en cas d'accident nucléaire.

Pour comprendre le comportement d'un système chimique dans son environnement, il faut avant tout déterminer sa structure électronique. Cela nous donne accès à son comportement dynamique quand il interagira avec des gouttelettes d'eau, des cristaux de glace ou encore avec des agrégats de poussières minérales dues à l'érosion ou organiques comme les suies.

D'un point de vue expérimental, pour déterminer ces propriétés électroniques, les méthodes spectroscopiques n'ont fait que progresser ces dernières décennies et cela dans tous les domaines, que ce soit pour les énergies accessibles ou pour la précision des mesures obtenues. Nous pensons, en écrivant cela, à la *X-ray Photoelectron Spectroscopy (XPS)* et aux énergies accessibles grâce aux *synchrotrons*, mais également aux expériences réalisées en *X-ray Absorption Spectroscopy (XAS)*. Il appartient donc à la chimie quantique de pouvoir aider quand les expériences sont difficilement réalisables, dans la résolution des spectres obtenus ou encore dans la prédiction d'effets très fins.

Pour les éléments lourds ($Z > 40$ [164]) ou pour les électrons de cœur, il est désormais admis qu'il faut résoudre l'équation de DIRAC, prenant en compte les vitesses relativistes de ces électrons ; dans ce sens nous avons utilisé l'hamiltonien DIRAC-COULOMB(-GAUNT) intégré dans le code *Dirac*¹. À cela il convient parfois d'ajouter certaines corrections issues de l'électrodynamique quantique (*Quantum Electrodynamics (QED)*).

1. Dans cette thèse le terme « *Dirac* » se référera au programme P.A.M. D.I.R.A.C. : Program for Atomic and Molecular Direct Iterative Relativistic All-electron Calculations [165, 1].

La méthode *Coupled-Cluster* (CC) a été choisie pour récupérer une partie de l'énergie de corrélation. Obtenir les énergies d'excitation (EE), les affinités électroniques (EA) et les potentiels d'ionisation (IP) en utilisant la méthode *Equation of Motion Coupled-Cluster* (EOM-CC), qui dérive de CC et cela dans un cadre relativiste, est donc naturellement apparu comme intéressant à étudier.

Cependant, calculer ces énergies entraîne de nombreux défis. Premièrement, la sélection des orbitales moléculaires (OM) d'intérêts, occupées et virtuelles, de façon à récupérer un maximum de la corrélation électronique tout en décrivant correctement les processus physiques, présente quelques difficultés. Deuxièmement, la taille des systèmes est limitée par le coût computationnel et cela à deux égards : le nombre de fonctions de Base (Basis Set) qui détermine le nombre d'intégrales biélectroniques, et le nombre de déterminants qui entrent en jeu dans l'obtention des valeurs propres de la matrice *Equation of Motion Coupled-Cluster* (EOM-CC). Troisièmement, ces méthodes de chimie quantique doivent être implémentées tout en tenant compte de l'utilisation efficace de plateformes de parallélisation des calculs, indispensable pour l'étude de systèmes de grande envergure.

Ainsi, cette thèse s'axe sur l'étude d'espèces isolées, principalement des composés halogénés en phase gazeuse. Après le premier chapitre théorique, le deuxième chapitre sera consacré à donner quelques éléments généraux sur les problématiques que sont : les méthodes pour obtenir IP et EE de cœur ; les approximations que nous pouvons apporter à nos méthodes de chimie quantique pour réduire les coûts de calcul computationnel ; et les défis que présentent les *High-Performance Computing* (HPC). Dans le troisième chapitre, nous répondrons à la première problématique en présentant la technique *Core-Valence Separation* (CVS) appliquée à EOM-CC. Dans le chapitre suivant, ce sont les approximations *Partition* et *Many-Body Perturbation Theory 2d order* (MBPT(2)) que nous étudierons. Enfin, le dernier chapitre traitera d'un nouveau module, ExaCorr, désormais intégré au code *Dirac*, dont le rôle est de paralléliser les calculs de produits tensoriels. Nous y exposerons notamment nos travaux en cours sur la méthode CC2 (CC2). Toujours dans le cadre des propriétés de la matière, pour terminer nous présenterons nos avancées pour l'implémentation de la polarisabilité (α), qui nous est apparue comme une candidate idéale à

implémenter dans le formalisme *Linear Response Coupled-Cluster (LR-CC)*.

Méthodes

Nous présenterons succinctement dans ce chapitre les formalismes qui ont été utilisés pour la réalisation des différents travaux : il n’y a ni résultats expérimentaux, ni résultats calculatoires. Pour cette partie, il conviendra de se reporter aux ouvrages de Szabo et Ostlund [193] concernant le développement de la méthode *HARTREE-FOCK (HF)* et *Configuration Interaction (CI)* et de Crawford et Schaefer III [43] pour son introduction à la méthode *Coupled-Cluster (CC)*. Concernant les parties plus spécialisées, le lecteur averti se reportera aux ouvrages de Dyllal et Faegri [56] pour les effets relativistes et Shavitt et Bartlett [178] pour l’étude diagrammatique de la méthode *Coupled-Cluster (CC)*. Enfin, puisque les théories perturbatives servent dans cette thèse, nous présenterons la méthode de *RAYLEIGH-SCHRÖDINGER* et la théorie de *MÖLLER-PLESSET (MP2)*.

1.1 Méthode *HARTREE-FOCK*

1.1.1 Hamiltonien non-relativiste

Considérons une molécule de M noyaux, de positions \mathbf{R}_α et charges Z_α (dans l’approximation de *BORN-OPPENHEIMER* [22]¹) et de N électrons, de posi-

1. Les noyaux, de positions \mathbf{R}_α , seront considérés comme fixes et les fonctions d’ondes électroniques et nucléaires seront découplées.

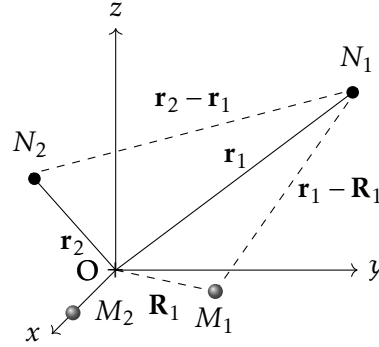


FIGURE 1.1 – Représentation des coordonnées d'espace pour un système de 2 noyaux M_1 et M_2 distants de \mathbf{R}_1 et 2 électrons N_1 et N_2 de positions \mathbf{r}_1 et \mathbf{r}_2 .

tions \mathbf{r}_i (fig.1.1). Sans effet relativiste, l'hamiltonien électronique s'écrit² :

$$H(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_i^N \left(-\frac{1}{2} \nabla_{\mathbf{r}_i}^2 - \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{r}_i - \mathbf{R}_{\alpha}|} \right) + \frac{1}{2} \sum_i^N \sum_{i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.1)$$

Dans l'ordre, les termes de l'équation 1.1 correspondent à l'opérateur énergie cinétique, l'interaction coulombienne noyau-électron et l'interaction coulombienne électron-électron. Généralement, cet hamiltonien est décomposé en deux : une partie mono-électronique correspondant aux deux premiers termes $h(i)$ et une partie bi-électronique correspondant à l'interaction électron-électron $g(i, j)$:

$$H(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_i^N h(i) + \frac{1}{2} \sum_i^N \sum_{i \neq j}^N g(i, j). \quad (1.2)$$

Pour les systèmes polyatomiques, l'énergie due à l'interaction coulombienne noyau-noyau est calculée a fortiori ou a posteriori sans influence sur l'hamiltonien électronique.

L'énergie E est calculée à partir de l'équation de SCHRÖDINGER indépendante du temps :

$$H(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = E \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N). \quad (1.3)$$

2. En unité atomique : $\hbar = 1$, $e = 1$, $m_e = 1$.

$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ est la fonction d'onde; les termes \mathbf{x}_i sont les degrés de liberté, i.e. les *spin-orbitales* : $\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$ avec $\mathbf{r}_i \in \mathbb{R}^3$ et σ_i est le spin, généralement noté α ou β . Cette fonction d'onde est antisymétrique par rapport à l'échange de particules (inversion de coordonnées \mathbf{x}_i) respectant le caractère fermionique des électrons.

1.1.2 Fonction d'onde

Dans la méthode *HARTREE-FOCK* (*HF*), la fonction d'onde Φ_0 est une approximation de la fonction d'onde de l'état fondamental Ψ_0 construite comme un unique déterminant de *SLATER* :

$$\Phi_0(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \quad (1.4)$$

Les χ_i sont les fonctions *spin-orbitales*, orthonormales, représentant les *orbitales moléculaires* (*OM*) occupées, accessibles aux N électrons. Elles sont décomposables en un produit d'une fonction d'onde spatiale et d'une fonction spin :

$$\chi_i(\mathbf{x}) = \phi_i(\mathbf{r}) \sigma_{\alpha, \beta}. \quad (1.5)$$

Cette fonction d'onde Φ_0 est normalisée par le préfacteur $(N!)^{-1/2}$. En utilisant la notation *braket* de *DIRAC*, le déterminant de *SLATER* s'écrit sous la forme simplifiée :

$$|\Phi_0\rangle = |\chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) \dots \chi_N(\mathbf{x}_N)\rangle = |\dots \chi_i \chi_j \dots \chi_N\rangle = |\dots i j \dots N\rangle. \quad (1.6)$$

Cette forme pour la fonction d'onde est antisymétrique. En effet, l'ordre des coordonnées *spin-orbitales* est toujours le même : $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$; l'inversion de deux colonnes change le signe du déterminant de *SLATER* simulant correctement

le comportement de la fonction d'onde :

$$|\dots i j \dots N\rangle = -P(ij)|\dots i j \dots N\rangle = -|\dots j i \dots N\rangle. \quad (1.7)$$

1.1.3 Énergie HARTREE-FOCK

L'énergie *HF* est obtenue comme la valeur moyenne de l'hamiltonien, utilisant comme fonction d'onde le déterminant de SLATER, c'est à dire : $E_{\text{HF}} = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle$. En développant cette expression en fonction des opérateurs mono- et bi- électroniques de l'hamiltonien, nous obtenons l'expression suivante [193] :

$$E_{\text{HF}} = \sum_i^N \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{i,j}^N \langle ij || ij \rangle, \quad (1.8)$$

ici :

$$\langle i | \hat{h} | i \rangle = h_{ii} = \int \chi_i^*(\mathbf{x}) h(i) \chi_i(\mathbf{x}) d\mathbf{x}, \quad (1.9)$$

et $\langle ij || ij \rangle = \langle ij | ij \rangle - \langle ij | ji \rangle$ est l'intégrale de répulsion électronique dans sa formulation antisymétrique, avec :

$$\langle ij | kl \rangle = \iint \frac{\chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) \chi_k(\mathbf{x}_1) \chi_l(\mathbf{x}_2)}{|\mathbf{r}_2 - \mathbf{r}_1|} d\mathbf{x}_1 d\mathbf{x}_2, \quad (1.10)$$

Nous avons l'expression de l'énergie *HF*, (eq.1.8), mais pour obtenir les orbitales qui correspondront à l'énergie la plus basse, i.e. l'état *fondamental*, nous allons utiliser un processus d'optimisation. Computationnellement, *HF* peut être réécrite comme un problème aux valeurs propres :

$$f(\mathbf{x}) \chi_i(\mathbf{x}) = \varepsilon_i \chi_i(\mathbf{x}), \quad (1.11)$$

où ε_i est l'énergie associée à chaque *spin-orbitale*. L'opérateur de Fock prend la forme de l'opérateur mono-électronique $h(\mathbf{x})$ associé à un potentiel v_{HF} :

$$f(\mathbf{x}) = h(\mathbf{x}) + v_{\text{HF}}(\mathbf{x}), \quad (1.12)$$

qui peut être développé sous la forme :

$$f_i(\mathbf{x}) = h_i(\mathbf{x}) + \sum_{j=1}^N |J_j(\mathbf{x}) - K_j(\mathbf{x})| \quad (1.13)$$

Dans cette expression, nous avons l'opérateur coulombien J et l'intégrale d'échange K (qui n'a pas d'équivalent classique) :

$$J_j \chi_i(\mathbf{x}_i) = \left(\int \chi_j^*(\mathbf{x}_j) \chi_j(\mathbf{x}_j) r_{ij}^{-1} d\mathbf{x}_j \right) \chi_i(\mathbf{x}_i) \quad (1.14)$$

$$K_j \chi_i(\mathbf{x}_i) = \left(\int \chi_j^*(\mathbf{x}_j) \chi_i(\mathbf{x}_j) r_{ij}^{-1} d\mathbf{x}_j \right) \chi_j(\mathbf{x}_i) \quad (1.15)$$

Résoudre $f(\mathbf{x})\chi_i(\mathbf{x}) = \varepsilon_i\chi_i(\mathbf{x})$ est en fait un problème aux pseudo-valeurs propres car l'opérateur de Fock, au travers de $v_{\text{HF}}(\mathbf{x})$, dépend lui-même des *spin-orbitales* ; l'optimisation est donc obtenue par un processus itératif. Le potentiel $v_{\text{HF}}(\mathbf{x})$ de *HF* est un potentiel moyen dû à la présence de tous les autres électrons. La procédure de résolution de l'équation de HARTREE-FOCK est de ce fait appelée *Self-Consistent-Field (SCF)*.

Enfin, par la présence du champ moyen v_{HF} , les positions des électrons ayant des spins différents ne sont pas corrélées (il manque la corrélation d'un point de vue coulombien). De ce fait, historiquement décrit par Löwdin [124], la différence entre l'énergie de l'état *fondamental* exacte E_0 et l'énergie HARTREE-FOCK est l'*énergie de corrélation* :

$$E_{\text{corr}} = E_0 - E_{\text{HF}}. \quad (1.16)$$

1.1.4 Expansion en fonction de base, Basis Set

Pour pouvoir résoudre les intégrales mono- et bi- électroniques, respectivement les équations (eq.1.9) et (eq.1.10), il faut connaître les fonctions χ_i . Ces *orbitales atomiques (OA)*, mono-électroniques sont représentées par des *fonctions de base (Base)*, qui facilitent le calcul. C'est le point de départ de tous les calculs moléculaires *Ab initio*. N'oublions pas justement que les *OM* ψ_i et l'énergie sont

obtenues par une combinaison linéaire des OA (*Linear Combination of Atomic Orbitals (LCAO)*) après optimisation des fonctions *spin-orbitales* :

Ainsi pour les OM, après optimisation :

$$\psi_i = \sum_j^{N_{AO}} c_{ij} \phi_i = \sum_j^{N_{AO}} c_{ij} \sum_k^N d_{jk} G_k, \quad (1.17)$$

où c_{ij} sont les coefficients optimisés et $d_{jk} G_k$ des fonctions mathématiques représentant l'OA.

Dans cette thèse, les fonctions de base utilisées sont de la forme *Cartesian Gaussian-Type Orbital (GTO)* pour les calculs CC :

$$G_{lmn}(\mathbf{r}, a, \mathbf{R}_A) = x_A^l y_A^m z_A^n \exp(-ar_A^2), \quad \begin{cases} a > 0, \mathbf{r}_A = \mathbf{r} - \mathbf{R}_A \\ l \geq 0, m \geq 0, n \geq 0 \end{cases} . \quad (1.18)$$

L'utilisation des GTO facilite les calculs. En effet, le produit de deux gaussiennes sur deux centres différents peut se réduire à une seule gaussienne [193] et les intégrations sont obtenues « facilement » [133, 40].

Le nombre de gaussiennes pour décrire chaque atome est plus ou moins grand et chaque *Base* a ses propres caractéristiques [96]; c'est à l'utilisateur de décider. Nous noterons les *fonctions de polarisation* qui permettent notamment une meilleure description de la densité électronique proche du noyau et les *fonctions diffuses* qui concernent la densité à longue distance. Szabo [193, p. 180] écrit à ce propos :

There are probably as many basis sets defined for polyatomic calculations as there are quantum chemists. The choice of a basis set is not nearly the black art, however, that it may first appear.

Ainsi, selon l'utilisation, nous pourrions faire l'étude avec des bases minimales (STO-3G) pour un débogage ou une étude avec des bases relativistes *core-correlated* (dyall-acv3 ζ) pour obtenir les potentiels d'ionisation des électrons des OM de cœur.

1.1.5 Size-Consistent Size-Extensive

Deux autres notions sont importantes en chimie quantique, ce sont les caractères :

- *Size-Consistent*, à savoir si la courbe énergétique d'un système qui se dissocie tend vers l'énergie correspondante à la somme de celle des systèmes isolés;
- *Size-Extensive*, notion identique à celle de la thermodynamique, à savoir si l'énergie d'un système composé de N sous-systèmes n'interagissant pas est identique à la somme des énergies de ces N sous-systèmes isolés.

La méthode *HF* peut être *Size-Consistent* et *Size-Extensive* si le processus de dissociation est correctement représenté³.

1.2 La Méthode full Configuration Interaction fCI

La méthode *fCI* est l'expression de la fonction d'onde « exacte » dans le sens où tous les déterminants de SLATER (celui représentant l'état *fondamental* mais également tous les déterminants excités) pouvant être construits à partir d'un nombre donné d'*OM*, participent à la fonction d'onde :

$$|\Psi_{CI}\rangle = c_0|\Phi_0\rangle + \sum_i^{\text{occ}} \sum_a^{\text{vir}} c_i^a |\Phi_i^a\rangle + \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \sum_{\substack{i<j \\ j<k}}^{\text{occ}} \sum_{\substack{a<b \\ b<c}}^{\text{vir}} c_{ijk}^{abb} |\Phi_{ijk}^{abc}\rangle + \dots \quad (1.19)$$

Les coefficients c sont obtenus par diagonalisation jusqu'à la convergence de l'énergie *fCI*⁴. Elle correspond à la résolution exacte de l'équation de SCHRÖDINGER, la seule approximation se faisant alors sur le caractère fini de la *Base* choisie. L'énergie obtenue E_{fCI} sert désormais à définir l'énergie de corrélation E_{cor} : $E_{cor} = E_{fCI} - E_{HF}$. Cependant, cette méthode est coûteuse car le nombre de dé-

3. Nous différencierons alors *Unrestricted Hartree-Fock* de *Restrictive Hartree-Fock*, cette dernière n'étant pas utilisable lors de la rupture d'une liaison covalente par exemple.

4. Dans le cas d'orbitales moléculaires non dégénérées, si plusieurs solutions sont trouvées, l'une d'elles correspondra à l'état fondamental et les autres correspondront alors à des états excités.

terminants progresse selon une loi binomiale $\binom{2K}{N}$, avec $2K$ *spin-orbitales* et N électrons. Pour pallier ce problème, nous pouvons tronquer cette expression après la deuxième excitation : *Configuration Interaction Singles Doubles (CISD)*, les calculs deviennent réalisables pour des systèmes plus importants ; mais cette méthode n'est plus *Size-Consistent* [180].

1.2.1 Énergie de corrélation

Il est intéressant ici de s'attarder sur l'énergie de corrélation. La corrélation électronique est caractérisée dans la littérature par deux limites : la corrélation statique et la corrélation dynamique [198].

Corrélation statique

La corrélation statique a trait aux états dégénérés ou quasi-dégénérés pour lesquels la représentation de la fonction d'onde par un seul déterminant de SLATER s'avère insuffisante. Les exemples typiques sont : la dissociation de la molécule de dihydrogène H_2 par la méthode *Restricted-HF* qui inclut la présence des configurations en $\{H^+; H^-\}$ (caractère à « longue distance » de la corrélation statique) ; la quasi-dégénérescence des *OM* 2s et 2p du Béryllium (caractère à « courte distance » de la corrélation statique [198]).

L'utilisation de méthodes post-*HF* multi-déterminantales, comme CISD ou plus généralement *Multi-Reference Configuration Interaction (MRCI)*, permet de retrouver une partie de cette énergie.

Corrélation dynamique

La corrélation dynamique concerne la corrélation de 2 électrons de par leurs positions. Par le principe d'exclusion de PAULI, deux électrons d'un même système ne peuvent avoir les mêmes degrés de liberté. Une partie de cette corrélation est déjà prise en compte dans la méthode *HF* par le fait que deux électrons de même spin ne peuvent se retrouver au même endroit : « trou de FERMI ». Cependant, la méthode *HF* en raison du champ moyen, ne reproduit pas

correctement la répulsion électronique entre deux électrons de *spin* différents, i.e. problème du « trou de COULOMB ».

Pour cette énergie, il faut considérer deux types différents de corrélation :

- Corrélation radiale : elle a pour but de corriger l'éloignement des électrons. Si nous ajoutons des fonctions plus diffuses, et par conséquent des *OM* virtuelles pouvant s'étendre plus loin, alors l'éloignement radial des électrons est mieux représenté ;
- Corrélation angulaire : elle a pour but de corriger l'écartement entre les électrons. Si nous ajoutons des fonctions de symétries différentes, nous améliorons la distribution spatiale, i.e. nous obtenons une meilleure flexibilité angulaire. C'est un problème qui apparaît notamment dans le cas des *OM* non-liantes. Les fonctions de polarisation qui possèdent un moment angulaire supérieur à ceux des *OA* occupées lors d'un calcul atomique *SCF* permettent le changement de la forme des orbitales atomiques nécessaire à la description des molécules [4, 177].

Il faut noter le travail de Dunning [51] sur les *correlation-consistent basis set* permettant de récupérer une partie de cette énergie par une couverture systématique de l'espace radial et angulaire sans introduire de biais entre ces corrélations.

1.2.2 Méthodes multi-références (MR)

full Configuration Interaction (FCI) est suffisamment flexible pour corriger les problèmes de corrélation, mais comme nous l'évoquions, cette méthode est si coûteuse qu'elle ne peut être envisagée. Ainsi, pour corriger les problèmes de corrélation statique, de *CISD* ou de toutes les autres méthodes basées sur un unique déterminant de *SLATER*, nous utilisons une fonction de référence multi-déterminantale, c'est-à-dire basée sur plusieurs déterminants de *SLATER*. De la sorte, il sera possible d'atteindre d'autres excitations. Ce sont les méthodes multi-références [116], qui combinent l'avantage de *CI* en évitant un coût computationnel trop important.

Cependant, avant de commencer à appliquer une méthode multi-référence, il nous faut choisir les déterminants de *SLATER*. Pour cela, nous découpons l'espace des orbitales en trois parties, un espace inactif formé par des orbitales pleines, un

espace actif sur lequel nous allons travailler (qui contient des orbitales occupées et virtuelles) et un espace d'orbitales virtuelles inoccupées et qui le resteront.

Nous pouvons citer notamment la méthode *Multi-Reference Configuration Interaction (MRCI)*[180]⁵. C'est un *CISD* sur plusieurs fonctions de référence R différentes :

$$|\Phi_{\text{MRCI}}\rangle = \sum_R c(R)|\Phi(R)\rangle + \sum_R \sum_{ia} c_i^a(R)|\Phi_i^a(R)\rangle + \sum_R \sum_{ijab} c_{ij}^{ab}(R)|\Phi_{ij}^{ab}(R)\rangle, \quad (1.20)$$

$\{i, j\}$ sont des *OM* occupées, $\{a, b\}$ les virtuelles et $|\Phi_i^a(R)\rangle$ et $|\Phi_{ij}^{ab}(R)\rangle$ représentent donc l'ensemble des simples- doubles- excitations associées à R .

Elle est écrite également sous la forme :

$$|\Phi_{\text{MRCI}}\rangle = \sum_I c_I |\Phi_I\rangle + \sum_S \sum_a c_S^a |\Phi_S^a\rangle + \sum_P \sum_{ab} c_P^{ab} |\Phi_P^{ab}\rangle, \quad (1.21)$$

où a et b sont des orbitales virtuelles, et I , S et P sont les références « internes », i.e. la référence, avec N , $N-1$ et $N-2$ électrons respectivement.

Nous pouvons citer en deuxième point les méthodes *Multi-Configurational Self-Consistent-Field (MCSCF)*, pour lesquelles plusieurs déterminants de SLATER sont considérés comme référence. Par exemple (*CASSCF*) [182] dont l'idée est de faire un *FCI* mais uniquement sur l'espace actif. À cette méthode s'ajoutent des variations selon la sélection des configurations qui constituent la référence, comme (*RASSCF*), où un sous-ensemble des configurations (*CASSCF*) est pris en compte.

La principale difficulté est le choix de l'espace actif qui dépend surtout du processus que nous souhaitons décrire : rupture de liaisons chimiques, excitations, etc. Cela nécessite un travail de réglage entre le choix des orbitales qui entrent en jeu, et le nombre d'orbitales de l'espace actif.

5. Il est à noter que cette méthode récupère également une partie de la corrélation dynamique.

1.3 Théorie perturbative

1.3.1 RAYLEIGH-SCHRÖDINGER *Perturbation Theory*

La théorie perturbative commence par l'équation de RAYLEIGH-SCHRÖDINGER :

$$\hat{H}^{(\lambda)} = \hat{H}^{(0)} + \lambda \hat{V}, \quad (1.22)$$

où $\hat{H}^{(0)}$ et $\hat{H}^{(\lambda)}$ sont les hamiltoniens développés respectivement à l'ordre (0) et (λ). \hat{V} est le potentiel considéré comme une perturbation. Enfin, λ est un coefficient variant sur $[0;1]$, si $\lambda = 0$ le système n'est pas perturbé et il l'est complètement si $\lambda = 1$. La solution recherchée est de la forme :

$$\hat{H} |\Psi_n\rangle = E_n |\Psi_n\rangle, \quad (1.23)$$

En appliquant un développement en série de TAYLOR en fonction de λ :

$$|\Psi_n\rangle = |\Psi_n^{(0)}\rangle + \lambda |\Psi_n^{(1)}\rangle + \lambda^2 |\Psi_n^{(2)}\rangle + \dots \quad (1.24)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots, \quad (1.25)$$

En remplaçant dans 1.23, nous pouvons écrire en fonction des ordres de λ :

$$\left(\hat{H}^{(0)} - E_n^{(0)} \right) |\Psi_n^{(0)}\rangle = 0 \quad (1.26)$$

$$\left(\hat{H}^{(0)} - E_n^{(0)} \right) |\Psi_n^{(1)}\rangle = \left(E_n^{(1)} - \hat{V} \right) |\Psi_n^{(0)}\rangle \quad (1.27)$$

$$\left(\hat{H}^{(0)} - E_n^{(0)} \right) |\Psi_n^{(2)}\rangle = \left(E_n^{(1)} - \hat{V} \right) |\Psi_n^{(1)}\rangle + E_n^{(2)} |\Psi_n^{(0)}\rangle, \quad (1.28)$$

Par la suite, en considérant uniquement l'état fondamental, *i.e.* $n=0$, et en considérant : $\hat{H}^{(0)} |\Phi_n\rangle = E_n^{(0)} |\Phi_n\rangle$ connu,

à l'ordre 0 :

$$|\Psi_0^{(0)}\rangle = |\Phi_0\rangle \quad (1.29)$$

à l'ordre 1, en utilisant $\langle \Phi_0 | \Psi_0^{(n)} \rangle = 0, \forall n > 0$, une projection sur $\langle \Phi_0 |$ puis

sur $\langle \Phi_n |$:

$$E_0^{(1)} = \langle \Phi_0 | \hat{V} | \Phi_0 \rangle \quad (1.30)$$

$$|\Psi_0^{(1)}\rangle = - \sum_{n \neq 0} \frac{\langle \Phi_n | \hat{V} | \Phi_0 \rangle}{E_n^{(0)} - E_0^{(0)}} |\Phi_n\rangle \quad (1.31)$$

à l'ordre 2 :

$$E_0^{(2)} = - \sum_{n \neq 0} \frac{|\langle \Phi_0 | \hat{V} | \Phi_n \rangle|^2}{E_n^{(0)} - E_0^{(0)}} \quad (1.32)$$

$$|\Psi_0^{(2)}\rangle = \sum_{n \neq 0} \frac{-\langle \Phi_n | \hat{V} | \Psi_0^{(1)} \rangle + E_0^{(1)} \langle \Phi_n | \Psi_0^{(1)} \rangle}{E_n^{(0)} - E_0^{(0)}} |\Phi_n\rangle \quad (1.33)$$

1.3.2 MÖLLER-PLESSET *Perturbation Theory*

La méthode *MÖLLER-PLESSET Perturbation Theory (MP2)* [135, 110] est un cas particulier de la théorie de *RAYLEIGH-SCHRÖDINGER* appliquée à *HF*.

$\hat{H}^{(0)}$ est défini comme :

$$\hat{H}^{(0)} = \hat{F} = \sum_i^N f(\mathbf{x}_i), \quad (1.34)$$

et l'opérateur perturbation \hat{V} tel que :

$$V(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{2} \sum_i^N \sum_{i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_i^N v_{\text{HF}}(\mathbf{x}_i) \quad (1.35)$$

$|\Phi_0\rangle$ est défini comme le déterminant *HF* et $|\Phi_n\rangle$ comme le déterminant simple, double, ..., excitation ⁶($|\Phi_i^a\rangle, |\Phi_{ij}^{ab}\rangle, \dots$). L'opérateur de Fock nous donne à l'ordre

6. $\{i,j,\dots\}$ représentant les orbitales occupées et $\{a,b,\dots\}$ les orbitales virtuelles, avec : $\hat{a}_a^\dagger \hat{a}_i |\Phi_0\rangle = |\Phi_i^a\rangle$.

(0) :

$$E_0^{(0)} = \sum_i^{\text{occ}} \varepsilon_i \quad (1.36)$$

Nous obtenons également :

$$E_0^{(0)} + E_0^{(1)} = E_{HF} \quad (1.37)$$

Par la suite, deux éléments importants en chimie quantique s'appliquent ici : la règle de SLATER qui a pour conséquence que deux déterminants (ϕ_1 et ϕ_2) différant de plus de deux OA sont tels que $\langle \phi_1 | \hat{H} | \phi_2 \rangle = 0$ et le fait qu'à convergence HF l'orthogonalité $\langle \Phi_0 | \Phi_i^a \rangle = 0$ soit assurée. Ainsi, de l'expression 1.32 il ne reste que le terme $\langle \Phi_0 | \hat{V} | \Phi_{ij}^{ab} \rangle = \langle ij || ab \rangle$, finalement⁷ :

$$E_{Corr}^{MP2} = -\frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{vir}} \frac{|\langle ij || ab \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \quad (1.38)$$

1.4 Méthode Coupled-Cluster

1.4.1 Énergie et Amplitudes

Dans le but d'obtenir une méthode plus performante que celles issues de théories perturbatives, et moins coûteuse que FCI, la méthode CC a été développée. La fonction d'onde Coupled-Cluster (CC) $|\Psi^{CC}\rangle$ est définie par l'application de *ansatz* exponentiel $e^{\hat{T}}$ sur la fonction d'onde de référence⁸. Ici, cela correspond à la fonction d'onde *ground-state* HF.

$$|\Psi^{CC}\rangle = e^{\hat{T}} |\text{HF}\rangle = e^{\hat{T}} |\Phi_0\rangle \quad (1.39)$$

7. Le facteur 1/4 venant de la sommation complète sur toutes les OM est omis si on sélectionne les OM telles que : $i < j$, $a < b$, car il faut prendre en compte le fait que : $\langle ij || ab \rangle = -\langle ji || ab \rangle = -\langle ij || ba \rangle$.

8. La paramétrisation n'est pas linéaire comme dans CI.

et pour l'énergie :

$$\left\langle \Phi_0 \left| e^{-\hat{T}} \hat{H} e^{\hat{T}} \right| \Phi_0 \right\rangle = E \quad (1.40)$$

Pour *Coupled-Cluster Singles Doubles (CCSD)* : $\hat{T} = \hat{T}_1 + \hat{T}_2$, avec :

$$\hat{T}_1 = \sum_{ai} t_i^a \{ \hat{a}_a^\dagger \hat{a}_i \} \quad \hat{T}_2 = \frac{1}{4} \sum_{abij} t_{ij}^{ab} \{ \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i \}, \quad (1.41)$$

les coefficients t_i^a et t_{ij}^{ab} sont appelés *Amplitudes*.

Écrit dans le formalisme de la seconde quantification, l'hamiltonien électronique⁹ est :

$$\hat{H} = \sum_{pq} \langle p | \hat{h} | q \rangle \hat{a}_p^\dagger \hat{a}_q + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \quad (1.42)$$

et réécrit dans sa forme *Normal-Ordered* et en utilisant le théorème de Wick [43, p 43]¹¹

$$\hat{H}_N = \hat{H} - \langle \Phi_0 | \hat{H} | \Phi_0 \rangle :$$

$$\hat{H}_N = \sum_{pq} \langle p | f | q \rangle \{ \hat{a}_p^\dagger \hat{a}_q \} + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \}, \quad (1.43)$$

que nous écrivons de manière plus concise sous la forme :

$$\hat{H}_N = \sum_{pq} f_q^p \{ \hat{a}_p^\dagger \hat{a}_q \} + \frac{1}{4} \sum_{pqrs} V_{rs}^{pq} \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \}. \quad (1.44)$$

L'équation 1.40 ne se résout pas aisément sous cette forme ; nous voyons en effet que cette équation n'est pas variationnelle. Il existe une technique plus évi-

9. Comme précédemment : approximations Born-Oppenheimer, homogénéité des unités, etc.

10. p, q, r, s sont, sans distinction, les indices correspondants aux opérateurs de création ou d'annihilation.

11. Avec la relation sur des *{hole ; particule}-lines* quelconques : $\hat{a}_p^\dagger \hat{a}_q = \{ \hat{a}_p^\dagger \hat{a}_q \} + \{ \hat{a}_p^\dagger \hat{a}_q \}$.

dente, par l'utilisation de la projection suivante : $\left(\langle \Phi_i^a | \oplus \langle \Phi_{ij}^{ab} | \right) e^{-\hat{T}} \hat{H}_N e^{\hat{T}} |\Phi_0\rangle = 0$ et en utilisant le théorème de BAKER-CAMPBELL-HAUSDORFF [2], nous pouvons résoudre ce système d'équations non-linéaires :

$$\left(\langle \Phi_i^a | \oplus \langle \Phi_{ij}^{ab} | \right) \left(\hat{H}_N \left(\frac{1}{2} \hat{T}_2^2 + \frac{1}{2} \hat{T}_2 \hat{T}_1^2 + \frac{1}{24} \hat{T}_1^4 + \hat{T}_2 \hat{T}_1 + \frac{1}{6} \hat{T}_1^3 + \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 + \hat{T}_1 + 1 \right) \right)_C |\Phi_0\rangle = 0, \quad (1.45)$$

L'indice C indique la connectivité entre les opérateurs \hat{H} et $\{\hat{T}_1; \hat{T}_2\}$. Notons que l'hamiltonien *Similarity-Transformed* est défini comme : $\hat{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}} = (\hat{H} e^{\hat{T}})_C$.

Ainsi, nous pouvons obtenir par itération jusqu'à un certain critère en énergie :

$$E_{CCSD} = \sum_a \sum_i t_i^a f_a^i + \frac{1}{4} \sum_{ab} \sum_{ij} (t_{ij}^{ab} + 2t_i^a t_j^b) V_{ab}^{ij} \quad (1.46)$$

Cette méthode permet de retrouver une partie de l'énergie de corrélation $E_{\text{cor CCSD}} = E_{CCSD} - E_{HF}$. Elle est *Size-Consistent* et *Size-Extensive* (cf. 1.1.5) [139, 128, 80]. En effet, la présence de termes appelés « disconnected », les produits d'opérateurs \hat{T} (comme $\hat{T}_2 \hat{T}_1$, $\hat{T}_2 \hat{T}_1^2$, etc.), permettent de retrouver un caractère *Size-Consistent* (contrairement au développement linéaire de *CISD*).

La méthode *Coupled-Cluster Singles Doubles (CCSD)* est basée sur un unique déterminant de SLATER et ne peut donc pas reproduire correctement les problèmes *multireference*, le « critère t_1 » permet de se rendre compte de l'importance du caractère multi-déterminantal de la fonction d'onde [114]. Nous notons la parution dernière de l'article de Bartlett [15] qui utilise la matrice densité *CC* pour affiner ce critère.

La méthode *CCSD* nécessite un nombre d'opérations en $\mathcal{O}(N^6)$ ¹³ (où N est le nombre d'*orbitales moléculaires (OM)*), ce qui limite son utilisation. A titre de

12. Cette équation n'est pas tronquée et cela quel que soit l'ordre du développement de l'exponentielle.

13. Comme *CISD*.

comparaison, le coût de HF est en $\mathcal{O}(N^4)$; pour les méthodes perturbées $MP2$ est en $\mathcal{O}(N^5)$, celui de $(MP4)$ est en $\mathcal{O}(N^7)$ mais la convergence des résultats n'est pas assurée vis-à-vis de l'ordre des perturbations [144]; pour $CISD$ le coût est également en $\mathcal{O}(N^6)$.

Enfin, il convient de noter que la méthode $CCSD(T)$ $\mathcal{O}(N^7)$, où les triples excitations sont incluses de façon perturbative (cf. 1.3), est *l'étalon-or* de la chimie quantique par les intérêts qu'elle présente tant par sa précision que par son coût computationnel [207, 157].

1.4.2 Approche diagrammatique

Il ne convient pas ici de reprendre l'ensemble des développements qui permettent de passer des équations algébriques aux équations programmables. Cependant, la plus grande partie de cette thèse est basée sur l'approche diagrammatique; il nous a donc semblé important de montrer un exemple.

Reprenons l'équation (eq.1.45). Il faut alors écrire l'ensemble des combinaisons possibles afin d'obtenir cette équation sous sa forme de produits tensoriels. Nous pouvons pour cela utiliser le théorème de Wick.

Imaginons ne prendre que le terme : $\langle \Phi_{ij}^{ab} | \hat{H}_N \hat{T}_2 | \Phi_0 \rangle \leftarrow 0$ ¹⁴, et limitons-nous à l'opérateur 2-électrons : $\hat{H}_N : \hat{V}_N = \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \} = \frac{1}{4} \sum_{pqrs} V_{rs}^{pq} \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \} :$

$$\langle \Phi_0 | \{ \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \} \frac{1}{4} \sum_{pqrs} V_{rs}^{pq} \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \} \frac{1}{4} \sum_{cdkl} t_{kl}^{cd} \{ \hat{a}_c^\dagger \hat{a}_d^\dagger \hat{a}_l \hat{a}_k \} | \Phi_0 \rangle \leftarrow 0 \quad (1.47)$$

Pour obtenir les produits tensoriels autorisés par la seconde quantification et par le théorème de Wick, nous pouvons commencer par les contractions suivantes :

$$\begin{array}{c} \text{Diagram showing contractions between } \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \text{ and } \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \text{ with lines connecting } i \text{ to } p, j \text{ to } q, b \text{ to } s, \text{ and } a \text{ to } r. \\ \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \hat{a}_c^\dagger \hat{a}_d^\dagger \hat{a}_l \hat{a}_k = -\delta_{is} \delta_{jr} \delta_{bc} \delta_{ad} \delta_{pl} \delta_{qk}, \end{array} \quad (1.48)$$


14. La flèche \leftarrow signifiant que l'expression *contribue* à 0 sans être elle-même nulle.

ce qui nous donne :

$$\langle \Phi_0 | -\frac{1}{16} \sum_{kl} V_{ji}^{lk} t_{kl}^{ba} | \Phi_0 \rangle \leftarrow 0 \quad (1.49)$$

Cependant, il faut écrire toutes les combinaisons possibles.


Une autre approche a alors été envisagée : *les expressions diagrammatiques*. Chaque opérateur est représenté par une *brique élémentaire*. Les opérateurs correspondant à t_1 et t_2 sont présentés sur la figure (fig.1.2).



$$\hat{\mathbf{T}}_1 : \sum_{ai} t_i^a \{ \hat{a}_a^\dagger \hat{a}_i \} \quad \hat{\mathbf{T}}_2 : \frac{1}{4} \sum_{abij} t_{ij}^{ab} \{ \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i \}$$

FIGURE 1.2 – Représentation diagrammatique des opérateurs d'excitation t_1 et t_2 .

Nous remarquerons que les *OM* occupées, nommées *hole-lines*, et les *OM* virtuelles, nommées *particule-lines* sont représentées respectivement par des flèches descendantes et ascendantes. Les opérateurs $\{1-;2-\}$ électrons sont présentés sur la figure (fig.1.3)¹⁵.



$$\mathbf{F} : f_j^i \{ \hat{a}_i^\dagger \hat{a}_j \} \quad \mathbf{V} : V_{kl}^{ij} \{ \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k \}$$

FIGURE 1.3 – Exemple de représentation diagrammatique des opérateurs $\{1-;2-\}$ électrons.

Pour reprendre l'équation (eq.1.47), nous cherchons donc l'expression tensorielle qui correspond donc au passage présenté sur la figure (fig.1.4), ce qui se rapporte à la figure (fig.1.5).

15. Un seul des opérateurs 1-électron sur les 4 existants et un seul des opérateurs 2-électrons sur les 9 existants sont présentés.

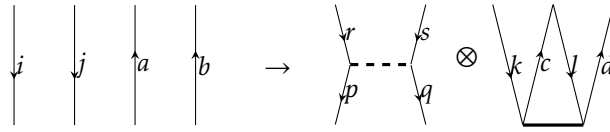


FIGURE 1.4 – Représentation diagrammatique d’une des contractions de l’équation (eq.1.47).

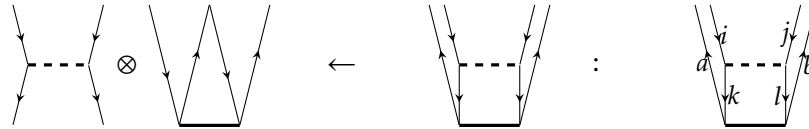


FIGURE 1.5 – Résultat de la contraction correspondant à l’équation (eq.1.47).

Nous devons dessiner toutes les combinaisons ; mais à partir d’un certain nombre de règles, nous pouvons déduire de ce diagramme l’expression :

$$\langle \Phi_{ij}^{ab} | \hat{V}_N \hat{T}_2 | \Phi_0 \rangle \leftarrow \frac{1}{2} \sum_{kl} V_{ij}^{kl} t_{kl}^{ab} \quad (1.50)$$

L’ensemble des équations et des diagrammes correspondants est donné dans le livre de Shavitt et Bartlett [178].

1.5 Equation of Motion Coupled-Cluster

Nous allons utiliser la méthode *Equation of Motion Coupled-Cluster (EOM-CC)* dans le but d’obtenir *les potentiels d’ionisation (IP)*, *les énergies d’excitation (EE)* et *les affinités électroniques (EA)*. Cette méthode qui s’applique sur une fonction de référence CC serait l’équivalent d’un FCI si la base des déterminants excités était complète. Rappelons que CC ne fournit pas, en soit, une façon d’obtenir les états excités autrement qu’en ayant pour origine un déterminant excité HF.

1.5.1 Description générale

Soit une fonction d'onde CC, $|CC\rangle$, et une fonction d'onde quelconque $|\Phi_x\rangle$ dont nous voulons l'énergie, il est toujours possible d'écrire :

$$\hat{H}|CC\rangle = E_{CC}|CC\rangle \quad (1.51)$$

$$\hat{H}|\Phi_x\rangle = E_x|\Phi_x\rangle \quad (1.52)$$

Nous avons besoin de $\omega_x = E_x - E_{CC}$, nous utilisons alors l'hamiltonien *normal-ordered* :

$$\hat{H}_N|CC\rangle = \Delta E_{CC}|CC\rangle \quad (1.53)$$

$$\hat{H}_N|\Phi_x\rangle = \Delta E_x|\Phi_x\rangle, \quad (1.54)$$

où : $\Delta E_{CC} = E_{CC} - E_0$ et $\Delta E_x = E_x - E_0$, ainsi $\omega_x = \Delta E_x - \Delta E_{CC}$. Puis, avec : $|CC\rangle = e^{\hat{T}}|HF\rangle$. S'il est possible d'écrire : $|\Phi_x\rangle = \hat{\mathcal{R}}_x|CC\rangle$, nous obtenons :

$$\hat{H}_N \hat{\mathcal{R}}_k e^{\hat{T}}|HF\rangle = \Delta E_x \hat{\mathcal{R}}_k e^{\hat{T}}|HF\rangle, \quad (1.55)$$

Comme nous avons préalablement défini l'hamiltonien *Normal-Ordered Similarity-Transformed* \mathcal{H} (en multipliant par e^{-T}), avec l'équation (eq.1.53) et $\hat{\mathcal{R}}$ dans l'équation (eq.1.55)¹⁶, nous pouvons écrire :

$$\mathcal{H}\hat{\mathcal{R}}|HF\rangle = \Delta E_x \hat{\mathcal{R}}|HF\rangle \quad (1.56)$$

$$\hat{\mathcal{R}}(\mathcal{H}|HF\rangle) = \hat{\mathcal{R}}(\Delta E_{CC}|HF\rangle) \quad (1.57)$$

En soustrayant ces 2 équations, nous obtenons :

16. Si $\hat{\mathcal{R}}$ et $e^{\hat{T}}$ commutent.

$$[\mathcal{H}, \hat{\mathcal{R}}] |HF\rangle = \omega_x \hat{\mathcal{R}} |HF\rangle \quad (1.58)$$

$$(\mathcal{H} \hat{\mathcal{R}} |HF\rangle)_C = \omega_x \hat{\mathcal{R}} |HF\rangle \quad (1.59)$$

L'équation (eq.1.59) devient alors un problème aux valeurs propres.

Enfin, il faut également s'intéresser à la partie gauche de cette équation.

$$(\langle HF | \hat{\mathcal{L}} \mathcal{H})_C = \omega_x \langle HF | \hat{\mathcal{L}}, \quad (1.60)$$

L_k et R_k obéissent à la condition de biorthogonalité : $L_k \cdot R_l = \delta_{kl}$

1.5.2 Énergie d'excitation, Potentiel d'ionisation, Affinité électronique

Pour les énergies d'excitation, en nous limitant aux *Simples* et *Doubles* excitations nous pouvons alors définir $\hat{\mathcal{R}}$ et $\hat{\mathcal{L}}$ sous la forme \hat{R}_{EE} et \hat{L}_{EE} ¹⁷ :

$$\hat{R}_{EE} = r_0 + \sum_{ai} r_i^a \{ \hat{a}_a^\dagger \hat{a}_i \} + \frac{1}{4} \sum_{abij} r_{ij}^{ab} \{ \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i \} \quad (1.61)$$

$$\hat{L}_{EE} = l_0 + \sum_{ai} l_a^i \{ \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \} + \frac{1}{4} \sum_{abij} l_{ab}^{ij} \{ \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_b \hat{a}_a \}, \quad (1.62)$$

\hat{R}_{EE} et \hat{L}_{EE} sont respectivement les opérateurs d'excitation et de désexcitation.

Dans cette thèse, nous travaillerons avec \hat{R} , avec *CCSD* ; nous avons alors la représentation matricielle suivante de \mathcal{H} (eq.1.59) :

$$\begin{array}{c} |\Phi_k^c\rangle \quad |\Phi_{kl}^{cd}\rangle \\ \left\langle \Phi_i^a \right| \quad \left\langle \Phi_{ij}^{ab} \right| \end{array} \left(\begin{array}{cc} \mathcal{H}_{SS} & \mathcal{H}_{SD} \\ \mathcal{H}_{DS} & \mathcal{H}_{DD} \end{array} \right) \quad (1.63)$$

Les différents termes de cette matrice représentent les tenseurs de passage

17. r_0 et l_0 sont des constantes [178, p 433].

du *bra* au *ket*. Ils sont séparés en blocs *S* pour *Single* et *D* pour *Double*.

Les définitions sont semblables pour les *potentiels d'ionisation (IP)* et les *affinités électroniques (EA)*.

$$\hat{R}_{IP} = \sum_i r_i \{\hat{a}_i\} + \frac{1}{2} \sum_{aij} r_{ij}^a \{\hat{a}_a^\dagger \hat{a}_j \hat{a}_i\} \quad (1.64)$$

$$\hat{R}_{EA} = \sum_a r^a \{\hat{a}_a^\dagger\} + \frac{1}{2} \sum_{abi} r_j^{ab} \{\hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i\} \quad (1.65)$$

Pour trouver les différents termes de cette matrice, il est plus avantageux d'utiliser l'approche diagrammatique [67, 187]. Attention, le coût est en $\mathcal{O}(N^6)$ pour *EE* et en $\mathcal{O}(N^5)$ pour *IP* et *EA*.

Les avantages de cette méthode sont nombreux. Nous citerons Landau [112] :

EOM-IP-CC describes open-shell doublet states as “ionized” states derived from a well-behaved closed-shell reference wave function. By avoiding open-shell reference, the method is free from spin encontamination and is less susceptible to symmetry breaking problems.

Nous notons également la revue de Krylov [109], dans laquelle d'autres exemples, comme l'effet Jan-Teller et les systèmes à transfert de charge sont montrés, la méthode *spin-flip* est également décrite. *EOM-CC* est *Size-Intensive* et *Size-Consistent* [139, 128, 80]

1.5.3 Matrice EOM-CCSD

En omettant la symétrie, la taille de la matrice *EOM-CC* peut être calculée en utilisant la loi binomiale. Avec '*N*' électrons, '*n*' l'ordre de l'excitation, '*2K*' le nombre de *spin-orbitales*, le nombre de déterminants est de :

$$\binom{N}{n} \binom{2K-N}{n}$$

La dimension de la matrice est la suivante :

$$\begin{array}{c} |\Phi_k^c\rangle \quad |\Phi_{kl}^{cd}\rangle \\ \langle \Phi_i^a | \quad \left(\begin{array}{cc} o.v * o.v & o.v * o^2.v^2 \\ o^2.v^2 * o.v & o^2.v^2 * o^2.v^2 \end{array} \right) \\ \langle \Phi_{ij}^{ab} | \end{array}$$

Prenons l'exemple de HCl avec la base dyall-av3 ζ ; il y a $o = 18$ orbitales occupées et $v = 208$ orbitales virtuelles, ainsi cette matrice a pour taille $((3744) + (14017536))^2$. Il est donc inutile de la construire, comme il est impossible de l'inverser en un temps raisonnable.

1.5.4 EOM-CC et l'algorithme de Davidson

Nous utilisons donc une procédure itérative : dans *Dirac*, nous avons recours à l'algorithme de Davidson [45]. C'est un algorithme itératif qui s'arrête selon un critère de convergence basé sur la norme des résidus¹⁸. La matrice *EOM* étant non-hermitienne par son côté *CC* et complexe par son côté relativiste, c'est en fait une généralisation de l'algorithme de Davidson créé par Hirao *et al.* [86] qui est en fait implémentée¹⁹.

De façon générale, nous décrivons la résolution, pour les équations *droite* et *gauche* sous la forme :

$$\sigma^R = \mathcal{H}\mathbf{b} \quad (1.66)$$

$$\sigma^L = \mathbf{b}^\dagger \mathcal{H}, \quad (1.67)$$

où, les σ , sont obtenus par projection de \mathcal{H} sur un vecteur d'essai \mathbf{b} et, à convergence, nous obtenons le vecteur-propre solution associé à une valeur propre.

Nous allons développer un peu plus ces propos en reprenant les grandes idées de cet algorithme pour le cas de matrices non-symétriques. D'abord, si nous voulons les vecteurs propres de \mathcal{H} de rang (N), nous commençons par définir que \mathbf{b} n'est pas qu'un vecteur d'essai mais en fait l'un des éléments d'une matrice $\mathbf{B}^{(m)} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ de taille $N \times m$. Nous n'allons pas travailler sur la matrice \mathcal{H}

18. i.e. basé sur la minimisation de l'erreur de la valeur propre par rapport à un vecteur propre d'essai.

19. Nous garderons dans cette thèse, comme il est d'usage, le nom Davidson pour décrire cette méthode itérative.

mais sur une réduction de cette matrice : $\tilde{\mathcal{H}} = \bar{\mathbf{B}}^{(m)H} \mathcal{H} \mathbf{B}^{(m)}$, où $\bar{\mathbf{B}}^{(m)H} \mathbf{B}^{(m)} = \mathbf{1}$ (bi-orthogonaux). Cette « petite » matrice $\tilde{\mathcal{H}}$ peut être diagonalisée classiquement sur des vecteurs propres $C^{(m)}$, également bi-orthogonaux, à droite $(\tilde{\mathcal{H}} - \lambda_k^{(m)}) C_k^{(m)} = 0$ et à gauche $(\tilde{\mathcal{H}}^H - \lambda_k^{(m)}) \bar{C}_k^{(m)} = 0$. Nous avons donc les vecteurs propres et les valeurs propres associés à cette matrice dans le sous-espace (m) . En remontant la chaîne des transformations, si nous avons : $\bar{C}_k^{(m)} \tilde{\mathcal{H}} C_k^{(m)} = \lambda_k^{(m)}$, nous pouvons écrire $\bar{C}_k^{(m)} \bar{\mathbf{B}}^{(m)H} \mathcal{H} \mathbf{B}^{(m)} C_k^{(m)} = \lambda_k^{(m)}$. Comme nous voulons $\mathcal{H} r_p = \omega_p r_p$, si $\omega_p = \lambda_k$, alors $r_p^{(m)} = \mathbf{B}^{(m)} C_k^{(m)}$ à droite et $\bar{r}_p^{(m)} = \bar{\mathbf{B}}^{(m)} \bar{C}_k^{(m)}$ à gauche. Ici, il convient de faire attention, les valeurs propres ne sont strictement identiques, d'un point de vue mathématique, que si les vecteurs propres sont identiques, ce qui n'est possible uniquement que si $(m) = (N)$. En fait, il faut considérer que c'est à la convergence de $\mathbf{B}^{(m)} C_k^{(m)}$ et de $\bar{\mathbf{B}}^{(m)} \bar{C}_k^{(m)}$ que nous aurons les solutions acceptables. En effet, nous commençons par un sous-espace relativement petit et le nous faisons progresser ; pour cela nous ajoutons des vecteurs b_{m+1} et \bar{b}_{m+1} respectivement à \mathbf{B} et $\bar{\mathbf{B}}$, ces nouveaux vecteurs sont créés selon un vecteur résiduel ξ qui mesure la différence entre $r_p^{(m)}$ et $\mathbf{B}^{(m+1)} C_k^{(m+1)}$ (pareillement pour la gauche avec $\bar{\xi}$). Les itérations s'arrêtent quand ce vecteur résiduel est suffisamment petit, selon un critère de convergence prédéfini. Il faut signaler que r_p et \bar{r}_p sont bi-orthogonaux ; il faut donc renormaliser les vecteurs propres après la convergence. Dans *Dirac*, la matrice \mathbf{B} commence par une taille $N \times m$, où N a la taille $t_1 + t_2$ et m est le nombre de solutions désirées. Elle se compose de 0, sauf aux lieux $\mathbf{B}(j, m) = 1$ où j correspond à la place du déterminant $|\phi_{i, \dots}^{a, \dots}\rangle$ qui nous intéresse pour la solution m .

Pour terminer sur l'algorithme de Davidson, intéressons-nous à la résolution de l'équation pour *IP* avec *EOM-CCSD* : $(\mathcal{H} \hat{\mathcal{R}} |HF\rangle)_C = \omega_k \hat{\mathcal{R}} |HF\rangle$. D'un point de vue mathématique, obtenir l'énergie E_{IP_k} , c'est obtenir la valeur propre correspondant au vecteur propre \mathbf{r}_k . Ce vecteur propre a pour base les déterminants *1 hole/ 0 particule* (1h/0p) $|\Phi_i\rangle$, i.e. simplement ionisé et *2 holes/1 particule* (2h/1p)

$|\Phi_{ij}^a\rangle$, i.e. simplement ionisé mais avec la relaxation induite par une excitation.

$$\mathcal{H}\mathbf{r}_{\mathbf{k}} = E_{IP_{\mathbf{k}}}\mathbf{r}_{\mathbf{k}}, \quad \mathbf{r}_{\mathbf{k}} : \begin{pmatrix} r_1 \\ \dots \\ r_i \\ \dots \\ r_{ij}^a \\ \dots \end{pmatrix}_{\mathbf{k}} \begin{pmatrix} |\Phi_1\rangle \\ \dots \\ |\Phi_i\rangle \\ \dots \\ |\Phi_{ij}^a\rangle \\ \dots \end{pmatrix} \quad (1.68)$$

Nous considérons à posteriori que l'énergie $E_{IP_{\mathbf{k}}}$ est l'énergie d'ionisation d'un électron de OM « i » si la valeur de r_i est prédominante. Nous définissons ensuite : $\%SI_k = 100(r_i)_k^2$. Si %SI est proche de 100, nous pouvons qualifier l'ionisation de mono-déterminante. S'il apparaît plusieurs déterminants simplement ionisés ayant des coefficients forts, nous aurons un état multi-déterminantal. Nous pouvons remarquer l'importance de la relaxation 2holes/1 particule dans le cas d'un état simple excitation présentant un caractère %SI plus faible.

1.6 Prise en compte de la relativité restreinte dans le comportement des électrons

Cette thèse est basée sur des hamiltoniens relativistes. Sans développer le formalisme mathématique, nous indiquerons dans cette partie quelques-unes des étapes clés; nous remarquerons dans ce sens la revue de Saue [162] et l'article de Liu [120].

Les vitesses des électrons de cœur sont relativistes. Pour les électrons 1s, nous avons : $v \sim Zc\alpha$, où Z est le numéro atomique, c la vitesse de la lumière et α la constante de structure fine ($\alpha \sim 137^{-1}$). Un calcul sur l'or nous montre qu'avec $Z=79$, nous obtenons une vitesse $v = 0,58c$. De la même manière, ces vitesses relativistes entraînent une contraction des longueurs. Un calcul par rapport au rayon de BOHR a_0 , nous donne un « rayon relativiste » a_{rel} , tel que :

$$a_{rel} = \frac{\hbar}{m\alpha} \sqrt{1 - \frac{v^2}{c^2}} < a_0, \quad (1.69)$$

ce rayon corrigé par la relativité devient pour l'or : $a_{rel} = 0,88a_0$.

Cette contraction des OA est décrite dans les articles de Baerends *et al.* [9] et de Schwerdtfeger [169]. Nous verrons également le couplage spin-orbite, décrit notamment dans les articles de Pyykkö *et al.* [156, 154].

Ces vitesses impliquent surtout, dans un premier temps, qu'il nous faut utiliser un autre modèle que l'hamiltonien de SCHRÖDINGER pour prendre en compte les effets relativistes.

1.6.1 Solution pour un système mono-électronique

DIRAC [48] s'intéresse à la conservation de l'équation de SCHRÖDINGER par transformation lorentzienne²⁰(LORENTZ-POINCARRÉ). Partant de l'énergie d'une particule libre en mouvement $E^2 = m^2c^4 + c^2p^2$, il arrive, en introduisant l'idée maîtresse que le spin (au travers des matrices de Pauli) est présent dans p^2 : $(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p}) = p^2 \mathbb{1}$, à l'équation suivante pour une particule libre :

$$\begin{bmatrix} mc^2 & c(\boldsymbol{\sigma} \cdot \mathbf{p}) \\ c(\boldsymbol{\sigma} \cdot \mathbf{p}) & -mc^2 \end{bmatrix} \begin{bmatrix} \psi^L \\ \psi^S \end{bmatrix} = E \begin{bmatrix} \psi^L \\ \psi^S \end{bmatrix} \quad (1.70)$$

où σ sont les matrices PAULI $\sigma_{x,y,z}$.

La résolution de ce système linéaire a impliqué de grands changements. Dans un premier temps, il y a apparition de solutions d'énergie positive et négative $E = \pm\sqrt{m^2c^4 + c^2p^2}$ (les valeurs propres), séparées par $\pm 2mc^2$. Cela implique la notion de l'antiparticule e^+ , le positon [47], découvert à posteriori par Anderson [6]. Ensuite, par la présence du spin, les vecteurs propres ont 4 composantes, que nous pouvons écrire, en fonction du spin (α, β) sous la forme :

$$\Psi = \begin{pmatrix} \psi^{L,\alpha} \\ \psi^{L,\beta} \\ \psi^{S,\alpha} \\ \psi^{S,\beta} \end{pmatrix} \quad (1.71)$$

Les deux premières composantes sont nommées *Large*, associées à l'énergie

20. i.e. respect du principe de la relativité restreinte : invariance des lois par changement de référentiel galiléen.

positive, ce qui est notre cas, et les secondes sont nommées *Small*. Quand nous considérons les énergies positives associées à une composante ψ^L importante, nous pouvons montrer que les composantes *Small* sont « petites », telles que :

$$\psi^S < \frac{p}{2mc} \psi^L, \quad (1.72)$$

d'où leur nom. Ce spineur est particulier à plusieurs niveaux ; nous ne l'étudions pas en détail [56] mais nous pouvons en donner quelques caractéristiques. Premièrement, par l'utilisation des matrices de PAULI le spineur peut être complexe. Deuxièmement, au contraire du cas non-relativiste, nous ne pouvons plus séparer les composantes spatiales et de spin comme simplement le produit de deux fonctions. Enfin, des changements dans la symétrie imposent qu'une rotation de 4π soit nécessaire pour que le spineur retrouve son état initial.

En présence d'un potentiel extérieur, par exemple celui généré par les noyaux, l'équation de DIRAC prend la forme ²¹ :

$$\left[c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2 - \mathbf{V} \right] \Psi = E\Psi = E \begin{pmatrix} \Psi_L \\ \Psi_s \end{pmatrix} = \begin{pmatrix} V & c(\boldsymbol{\sigma} \cdot \mathbf{p}) \\ c(\boldsymbol{\sigma} \cdot \mathbf{p}) & V - 2mc^2 \end{pmatrix} \begin{pmatrix} \Psi_L \\ \Psi_s \end{pmatrix} \quad (1.73)$$

Sous cette forme, avec 1 électron et 1 noyau nous avons déjà l'apparition du couplage *spin-orbit*. Il conviendrait d'appliquer le couplage minimal, i.e. la transformation de $\mathbf{p} \rightarrow \boldsymbol{\pi} = \mathbf{p} + e\mathbf{A}$. Le noyau, avec un magnéton nucléaire $\mu_N \sim \frac{q \cdot \hbar}{2m_p}$ inversement proportionnel à sa masse, a une influence mille fois inférieure à celle produite par le moment magnétique de l'électron. L'influence du potentiel vecteur généré par le noyau est donc abordée sous forme perturbative. Comme nous ne traitons pas des propriétés magnétiques dans cette thèse, nous nous bornerons à écrire que l'avantage de négliger le champ magnétique est d'obtenir une équation de DIRAC symétrique par rapport au temps. Dans *Dirac*, pour des systèmes *couches fermées* et en l'absence de champ magnétique, cela nous permettra d'utiliser les « paires de KRAMERS » [108].

21. Une translation d'énergie est effectuée pour que les énergies positives soient comprises dans l'intervalle $[0; +\infty[$.

1.6.2 Solution pour un système poly-électronique

Avec Dirac [1, 165], nous écrivons [163], l'hamiltonien de DIRAC-COULOMB sous le formalisme de la seconde quantification :

$$\hat{H}_N = \sum_{pq} f_q^p \{\hat{a}_p^\dagger \hat{a}_q\} + \frac{1}{4} \sum_{pqrs} V_{rs}^{pq} \{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\}, \quad (1.74)$$

avec f_q^p l'opérateur de Fock et V_{rs}^{pq} , les intégrales 2-électrons

$$V_{rs}^{pq} = \langle pq | \hat{g} | rs \rangle - \langle pq | \hat{g} | sr \rangle, \quad \hat{g}(i, j) = \frac{1}{r_{ij}}. \quad (1.75)$$

Cette expression a la même structure que dans le domaine non-relativiste, mais le couplage Spin-Same-Orbit (SSO) est inclus.

Les modifications à apporter sont nombreuses pour obtenir des solutions convergentes et il a fallu prendre en compte de nombreux éléments. Dès le début des calculs relativistes, le problème des solutions d'énergie négative s'est posé; en effet cet hamiltonien n'est pas borné en énergie [28]. Dans cette thèse, nous resterons dans l'approximation « no-pair »; ainsi, nos déterminants ne contiendront que des spineurs d'énergie positive [5].

Pour montrer les autres défis du formalisme relativiste, nous pouvons réécrire le système linéaire (eq.1.73) sous la forme :

$$\psi^S = \left[2mc^2 + E - V \right]^{-1} c(\boldsymbol{\sigma} \cdot \mathbf{p}) \psi^L = \frac{(\boldsymbol{\sigma} \cdot \mathbf{p})}{2mc} \left[1 + \frac{E - V}{2mc^2} \right]^{-1} \psi^L \quad (1.76)$$

La première partie de l'équation montre le problème qui peut apparaître si le potentiel nucléaire est mal défini. Généralement, nous remplaçons le modèle ponctuel par un modèle continu gaussien [202, 7].

De plus, les composantes *Large* et *Small* ne sont pas indépendantes et cela a conduit à la notion de « *kinetic balance* ». Si nous pouvons voir que ψ^S est nul dans le cas non-relativiste, i.e. quand c tend vers l'infini, ici, ce lien existe. Nous pouvons écrire :

$$\lim_{c \rightarrow \infty} c \psi^S = \frac{1}{2m} (\boldsymbol{\sigma} \cdot \mathbf{p}) \psi^L \quad (1.77)$$

Nous devons respecter ce critère au niveau des *OA* qui correspondront aux composantes *Small* [204]²², il n'y a pas de convergence *HF* sans respect de cette règle.

Enfin, indépendamment des critères de convergence, les composantes *Small* ont une densité localisée proche du noyau et ces intégrales (SS|SS) ne contribuent qu'à hauteur de α^4 dans l'énergie. Elles ne sont généralement pas calculées mais remplacées par une valeur approximée [201].

1.6.3 Hamiltonien de BREIT - Électrodynamique Quantique

Pour améliorer la qualité des résultats obtenus par rapport aux expériences, nous devons améliorer l'hamiltonien de DIRAC-COULOMB. Nous montrerons notamment l'importance des effets de l'hamiltonien de BREIT dans le chapitre 3 [79] pour les électrons de cœur.

Correction de l'hamiltonien : hamiltonien de BREIT

Nous pouvons améliorer la qualité de l'hamiltonien en ajoutant des termes, négligés jusqu'à présent, dans les intégrales 2-électrons. En effet, certains effets relativistes (d'ordre plus petit en $1/c^2$) apparaissent quand nous prenons en compte les effets des champs magnétiques créés par les *boucles de courants* des noyaux dans les référentiels des électrons et celles des électrons dans leurs orbites respectives. Ces termes²³ peuvent être présentés dans le terme de BREIT :

$$g^{\text{Breit}}(i, j) = -\frac{c\alpha_i \cdot c\alpha_j}{2c^2 r_{ij}} - \frac{(c\alpha_i \cdot \mathbf{r}_{ij})(c\alpha_j \cdot \mathbf{r}_{ij})}{2c^2 r_{ij}^3} \quad (1.78)$$

qui peut être modifié comme :

$$g^{\text{Breit}}(i, j) = g^{\text{Gaunt}}(i, j) + g^{\text{gauge}}(i, j) = -\frac{c\alpha_i \cdot c\alpha_j}{c^2 r_{ij}} - \frac{(c\alpha_i \cdot \nabla_i)(c\alpha_j \cdot \nabla_j) r_{ij}}{2c^2} \quad (1.79)$$

22. De nouvelles difficultés apparaissent alors dans les cas *restricted*- ou *unrestricted*- kinetic balance, i.e. le nombre de composantes *Small* égal ou non au nombre de composantes *Large*; mais cela dépasse le cadre de cette thèse.

23. Termes que nous retrouvons également en électrodynamique quantique.

Alors que le couplage *spin-orbite* (SSO : Spin-Same-Orbit) est intégré dans l'interaction coulombienne, le terme de Gaunt prend en compte l'interaction entre des orbites différentes (SOO : Spin-Other-Orbit)²⁴. Le terme de *jauge* (gauge-term) qui intègre le retard à la transmission de l'interaction, n'est pas inclus dans nos calculs.

Correction de l'Électrodynamique Quantique.

En ce qui concerne les électrons de cœur, pour lesquels les effets relativistes sont importants [155], il faut ajouter les corrections que sont la « self-energy » et la « vacuum polarization » [121, 119, 120]. Alors que la première correspond à l'interaction entre l'électron et son propre champ radiatif, la seconde correspond à la création d'une paire particule-antiparticule (fig.1.6). Les calculs moléculaires des effets de la QED n'étant pas disponibles avec *Dirac*, nous avons utilisé une correction atomique basée sur les travaux de Koziol *et al.* [107] (cf.3.3), avec la limitation que la modification de la fonction d'onde n'a pas été considérée.

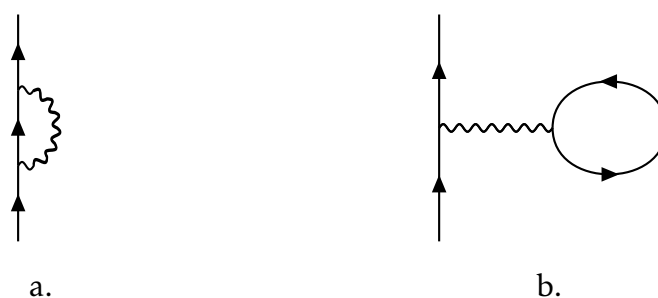


FIGURE 1.6 – Représentation de Feynman de la « self-energy » (a.) et de la « polarisation du vide » (b.).

1.6.4 Approximation à 2-composantes de l'hamiltonien du Dirac-Coulomb : ${}^2\text{DC} - {}^2\text{DC}^{\text{M}}$

L'équation 4-composantes (eq.1.73) étant coûteuse computationnellement [56, p.204], des approximations ont été créées pour passer de 4- à 2- composantes,

²⁴. L'ajout du terme de GAUNT rendant les calculs plus coûteux ce n'est pas une option par défaut dans *Dirac*.

basées sur la suppression du calcul des *Small* car très localisées [56, 122, 163, 44]. Même si différentes approches ont été conçues pour découpler les composantes, nous utiliserons les « *eXact 2-Component* » : X2C. Et deux approximations sont utilisées dans cette thèse pour transformer cette expression en une équation relativiste 2-composantes, que nous nommerons : ${}^2\text{DC}$ et ${}^2\text{DC}^{\text{M}}$.

Dans les deux cas, c'est une transformation unitaire de type FOLDY-WOUTHUYSEN, où nous cherchons une transformation unitaire U , telle que :

$$U^\dagger \begin{bmatrix} h_{LL} & h_{LS} \\ h_{SL} & h_{SS} \end{bmatrix} U = \begin{bmatrix} \tilde{h}_{++} & 0 \\ 0 & \tilde{h}_{--} \end{bmatrix} \quad (1.80)$$

$$h^{2c} = [U^\dagger h^{4c} U]_{++} \quad (1.81)$$

Si cette transformation a lieu dès le début du processus sur l'hamiltonien 1-électron (bare hamiltonian) nous obtenons X2C-1e [91]. Celui-ci n'est pas assez précis puisque l'interaction coulombienne est prise de façon non-relativiste ; nous appliquons généralement une correction atomique : *Atomic Mean-Field Integrals (AMFI)* [85, 166], permettant de retrouver une partie du couplage *spin-orbitale*. Nous avons donc, X2C-AMFI, nommé ${}^2\text{DC}$ dans le corps de cette thèse.

Si cette transformation se fait après le cycle *SCF* 4-composantes à convergence de l'opérateur de Fock, nous arrivons à l'hamiltonien X2C « molecular mean-field », nommé ${}^2\text{DC}^{\text{M}}$ [183]. Le coût est supérieur, mais l'erreur est moindre (*picture-change error* [165, 163]).

Les différentes problématiques

PROBLÉMATIQUE N 1 : OBTENIR LES POTENTIELS D'IONISATION ET LES ÉNERGIES D'EXCITATION DE COEUR

La chimie quantique a notamment pour but de détecter les espèces en présence, étudier leurs interactions et obtenir leurs propriétés physico-chimiques. L'équipe PCMT du PhLAM a déjà une grande expérience dans ces domaines : obtenir *les potentiels d'ionisation (IP)* de valence [179], étudier l'influence de l'environnement : Embedding [73, 23], obtenir les propriétés [143, 78] et travailler en dynamique moléculaire aux interfaces pour étudier les réactivités [161].

Nous voulons, ici, montrer différentes techniques permettant d'atteindre les propriétés des électrons de cœur. Nous donnons principalement des informations relatives aux potentiels d'ionisation, mais nous aurons l'occasion de donner des informations sur les énergies d'excitation.

Cette section présentera une expérience de *X-ray Photoelectron Spectroscopy*, nous indiquerons quelques éléments techniques et expérimentaux. Nous poursuivrons, naturellement, par donner quelques méthodes de chimie quantique permettant d'atteindre les propriétés *IP* et *EE*. Cette problématique s'inscrit dans le travail que nous présentons au chapitre 3.

2.1 *X-ray Photoelectron Spectroscopy (XPS)*

2.1.1 Présentation d'une expérience XPS

L'une des expériences qui montre l'utilité de la *X-ray Photoelectron Spectroscopy* est celle présentée par Kong *et al.* [106]¹. Cet article présente l'interaction entre l'acide chlorhydrique gazeux et la glace d'eau (fig. 2.1-a). Entre autres choses, les *IP* 1s et 2p du chlore sont mesurés (fig. 2.1-b); mais nous avons également dans ce travail la présentation de tous les intérêts de cette spectroscopie.

En effet, nous détectons facilement les espèces en présence : le chlore est la seule espèce présente à avoir des *IP* à ce niveau d'énergie (~2800 eV). Nous différencions Cl⁻ de HCl : l'hydrogène a une influence sur les *IP* du chlore (shift de 2,2 eV). Nous pouvons travailler sur une surface, dans un liquide ou dans une interface (physisorption); nous obtenons un profil en profondeur ($\Delta d=0,5$ nm) et nous observons alors l'influence de l'environnement (phase aqueuse ou phase gazeuse). De plus, nous parvenons à des résultats très précis : les raies ont ici des largeurs à mi-hauteur de l'ordre de 1,0 eV.

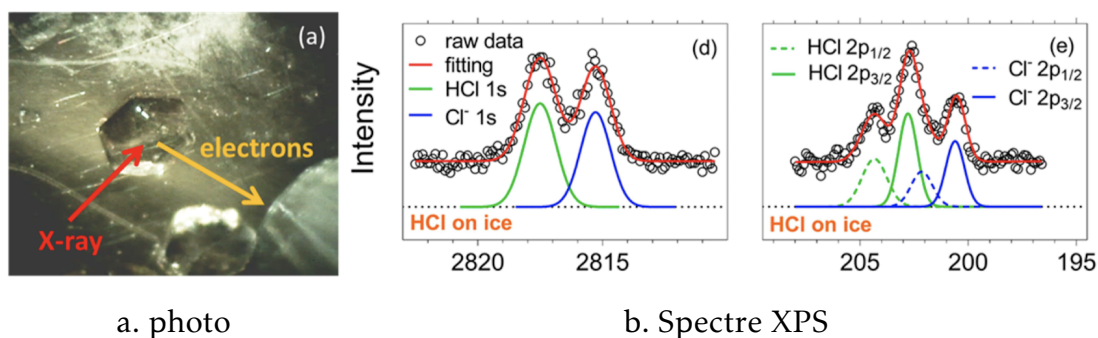


FIGURE 2.1 – Spectre XPS du Chlore (niveaux 1s et 2p) durant l'absorption de HCl sur la glace (eV) [106]².

1. Nous notons également le travail de Pelimanni *et al.* [148] mais il convient de remarquer qu'en 2017 Greczynski *et al.* [74] comptent plus de 9000 articles sur la méthode XPS.

2. Reprinted (adapted) with permission from Xiangrui Kong and Astrid Waldner and Fabrizio Orlando *et al.* - Coexistence of Physisorbed and Solvated HCl at Warm Ice Surfaces - The Journal of Physical Chemistry Letters 8.19 (sept. 2017), p4757-4762. Copyright 2017 American Chemical Society.

2.1.2 Expérimentalement

Il ne convient pas de reprendre l'historique de la spectroscopie ni depuis NEWTON, ni depuis HERTZ [89], ni de montrer toutes les avancées [60, 49]. De la même façon, nous ne traiterons pas des difficultés inhérentes à l'obtention d'un spectre XPS, comme décrit notamment par Shard [176] et par Stevie *et al.* [190]. Cependant, il paraît intéressant de donner quelques éléments sur les avancées technologiques en lien direct avec ce que nous sommes capables de calculer pour les *IP*.

Les énergies

Au contraire de la *Photoelectron Spectroscopy (PES)* pour la valence, *Ultraviolet Photoelectron Spectroscopy (UPS)*, qui demande des énergies de 10 à 45 eV, pour les électrons de cœur, il nous faut des énergies de l'ordre de 200eV à 2000eV : *soft x-rays* à des valeurs supérieures à 2000eV : *hard x-rays*, technologie *X-ray Photoelectron Spectroscopy (XPS)*. Les technologies *synchrotron* nous permettent d'obtenir ces gammes d'énergies (fig.2.2).

Les échantillons

Les échantillons peuvent se présenter sous plusieurs formes : *Composés solides*, historiquement les métaux et des cristaux [11, 21, 173], mais également en chimie organique sur des polymères [27]; *Composés liquides* [209, 59]; *Composés gazeux* [13] dont CH₃I, CF₃I [24], Xe [145] qui sont étudiés dans notre article [79]. Enfin, aux *interfaces*, comme nous l'écrivions en introduction de ce chapitre, les réactions sont d'un grand intérêt (HCl sur glace [106], interaction gaz/liquide [52]).

Qualité des résultats

La précision dépend à la fois de la qualité du rayonnement incident, des monochromateurs (la précision relative pour ces deux éléments peut être de l'ordre $\Delta E/E = 10^{-9}$ [49]), des capteurs et de l'échantillon. Il est donc difficile de comparer les données provenant de deux expériences différentes. Cependant,

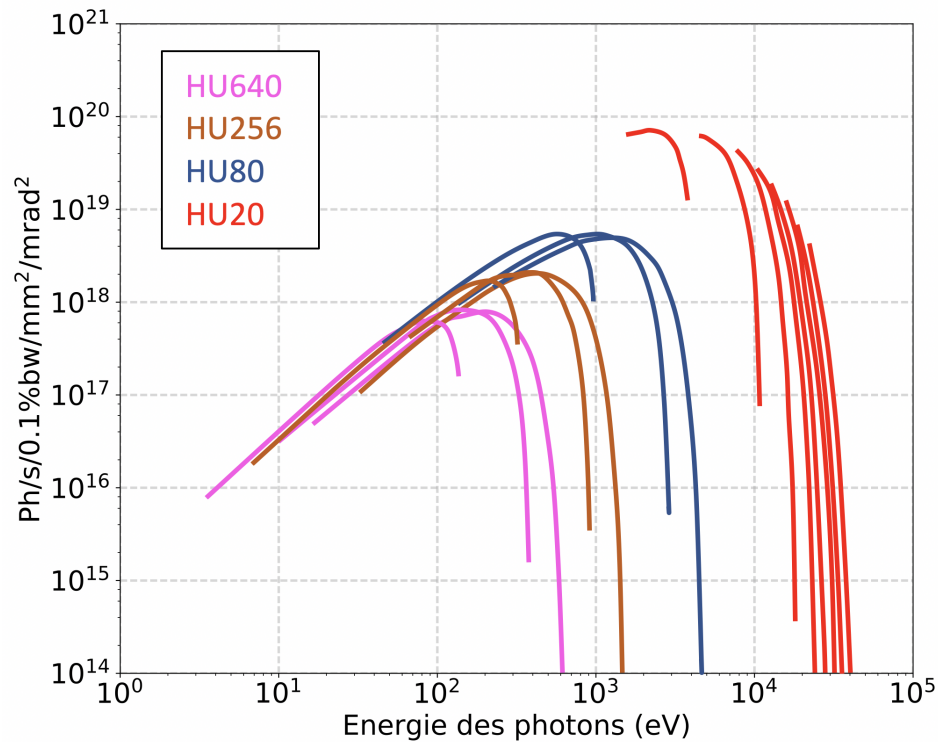


FIGURE 2.2 – Exemple d'énergie accessible par la technologie synchrotron (eV). SOLEIL : Source Optimisée de Lumière d'Énergie Intermédiaire du LURE (Laboratoire d'utilisation du rayonnement électromagnétique, Paris) [49, 42]. Les différents termes et les couleurs représentent les différentes lignes de travail.

pour des expériences *XPS* des isoélectroniques du xénon et pour des électrons de cœur d'énergie proche de 30 keV, nous pouvons obtenir une largeur à mi-hauteur proche de 10 eV et une précision de l'ordre de 1×10^{-2} eV [145].

Spectre expérimental

Le traitement des données expérimentales peut s'avérer complexe [136, 190]. Nous ne donnerons ici que la description sommaire d'un spectre *XPS*.

Sur la figure (fig.2.3), nous observons en ordonnée une unité arbitraire (parfois en *counts*). En abscisse, l'énergie est généralement donnée en eV. C'est

l'énergie cinétique E_k des électrons émis qui est mesurée selon la relation :

$$E_k = h\nu - BE - \Phi_s, \quad (2.1)$$

où, $h\nu$ est l'énergie des photons, BE est la *Binding Energy*, i.e. IP et Φ_s est la *work function* ; c'est une valeur caractéristique du spectromètre ³.

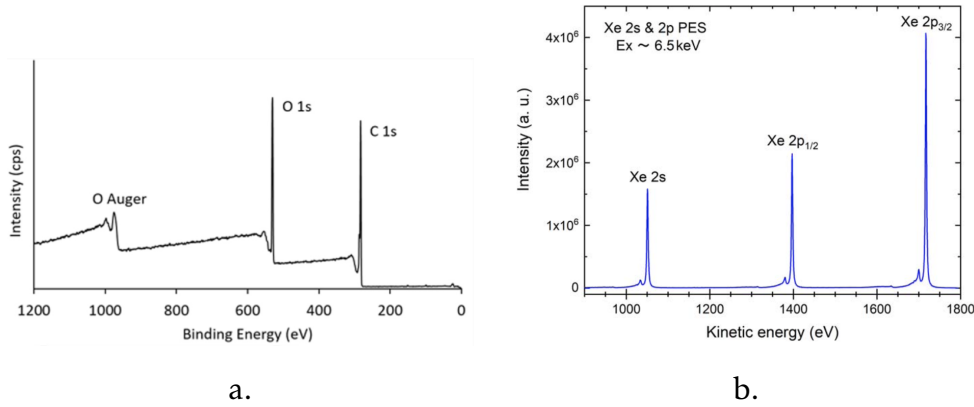


FIGURE 2.3 – Spectres XPS expérimentaux, a. Spectre du PET - Stevie *et al.* [190] ⁴, b. Spectre du xénon - Oura *et al.* [145].

La dénomination des pics se fait soit par le nom de *OM* d'origine $nL_{j=l+s}$ soit par la dénomination des couches électroniques K, L, M [71] (tab.2.1).

Couche	K	L ₁	L ₂	L ₃	M ₁	M ₂	M ₃	M ₄	M ₅
Configuration électronique	$1s^{-1}$	$2s^{-1}$	$2p_{1/2}^{-1}$	$2p_{3/2}^{-1}$	$3s^{-1}$	$3p_{1/2}^{-1}$	$3p_{3/2}^{-1}$	$3d_{3/2}^{-1}$	$3d_{5/2}^{-1}$

TABLEAU 2.1 – Correspondance entre la notation par couche électronique (*edge*) et la notation par orbitale.

Comme nous nous intéressons principalement aux *OM* de cœur, nous restons dans le chapitre 3 en notation atomique, même s'il peut y avoir de légères

3. Pour la calibration, nous pouvons utiliser typiquement la raie 4f de l'or ou les raies 2p et 3p du cuivre [136, 174], mais la calibration présente de nombreux défis [172, 74].

4. Fred A. Stevie, Carrie L. Donley, Introduction to x-ray photoelectron spectroscopy, Journal of Vacuum Science & Technology A, American Vacuum Society, Dec 1, 2020, Rights managed by AIP Publishing. Reprinted with permission of Copyright Clearance Center. All rights reserved.

levées de dégénérescence, en symétrie atomique, pour les énergies d'ionisation.

2.2 Chimie quantique et potentiels d'ionisation des électrons de cœur

2.2.1 Lien entre les calculs et les expériences XPS

Commençons par regarder si, par un calcul *HF*, il est possible de retrouver les caractéristiques d'un spectre *XPS* (cf.2.1.1).

Prenons l'exemple des anions iodure I^- , et hypoiodite IO^- (fig.2.4) et intéressons-nous aux couches K et L. Nous retrouvons bien tous les intérêts et les facteurs influençants : nous pouvons distinguer facilement la présence de l'iode, seule espèce à ce niveau d'énergie ($\sim 33\text{keV}$), les OA de l'oxygène étant à plusieurs keV. Nous différencions I^- de IO^- et l'ajout d'une molécule d'eau a de l'influence : l'environnement chimique de l'atome cible provoque bien des modifications dans la valeur de l'*IP*.

Nous pouvons donc, en première approximation, utiliser le théorème de Koopmans⁵ pour déterminer les *IP*.

Cependant, si cette approximation est correcte concernant les électrons de valence, la disparition d'un électron de cœur entraîne un changement important du *screening* du noyau. La relaxation de la fonction d'onde est très importante et doit être prise en compte pour que nos résultats soient comparables à l'expérience. Il est donc intéressant d'utiliser des méthodes pour lesquelles la corrélation est correctement représentée et pour lesquelles la relaxation est également prise en compte de façon acceptable. La méthode *Equation of Motion Coupled-Cluster Singles Doubles (EOM-CCSD)* présente ces deux avantages à la fois par sa fonction de référence *CCSD* et par le fait de prendre en compte des déterminants de « relaxation »⁶.

5. De la même façon, utiliser le théorème de Janak [94] (théorie de la fonctionnelle de la densité).

6. (2h/2p) pour les simples excitations *EE* et (2h/1p) et (1h/2p) pour respectivement les simples *IP* et les simples *EA*.

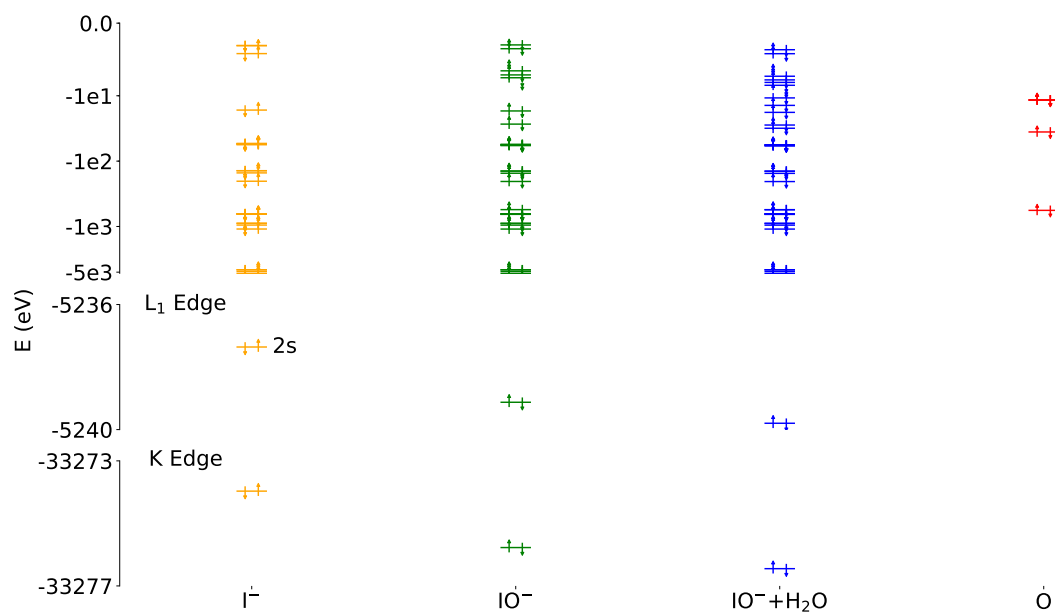


FIGURE 2.4 – Les intérêts de la XPS dans un calcul HF. Énergies (en eV) des OM pour IO^- ; I^- ; $\{\text{IO}^-+\text{H}_2\text{O}\}$ et O - $^2\text{DC}^{\text{M}}+\text{Gaunt}$ - dyall-av3 ζ pour I et aug-cc-pV3 ζ pour H et O. IO^- : $d_{\text{I-O}}=1,875 \text{ \AA}$; $\{\text{IO}^-;\text{H}_2\text{O}\}$ $d_{\text{O-H}}=0,975 \text{ \AA}$, $\text{H}\hat{\text{O}}\text{H}=110,565^\circ$, IO^- —HO alignés, $d=1,5 \text{ \AA}$.

2.2.2 Exemples de quelques classes de méthodes permettant d'accéder aux propriétés de cœur

Cette thèse traite principalement des *Simple-IP* de cœur, mais nous allons présenter ici quelques classes de méthodes de chimie quantique permettant d'obtenir les *IP* et les *EE*.

Il existe deux types de techniques : les méthodes « indirectes » nécessitant deux calculs, un pour l'état *fondamental*, un autre pour l'état ciblé et les méthodes « directes » capables d'atteindre l'énergie de l'état excité ou ionisé directement.

Commençons par les méthodes « indirectes », comme ΔSCF qui regroupe à la fois ΔHF , ΔDFT et ΔMP2 . Dans ces méthodes, il y a l'apparition (ou la disparition) d'un radical, dans lequel au moins une *OM* est incomplète. Nous pouvons penser à l'équation (eq.2.2.2) pour la réaction d'ionisation d'une espèce

X^7 :



Ces méthodes nécessitent généralement deux calculs, dont un calcul *couche ouverte*. La représentation est donnée en figure (fig.2.5). Il faut calculer les énergies des OM pour l'état fondamental, trouver l'OM d'intérêt et effectuer un second calcul *open-shell*.

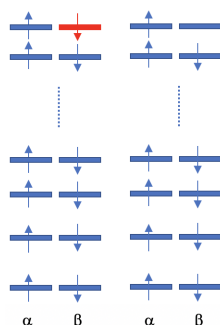


FIGURE 2.5 – Représentation des deux étapes nécessaires à l'obtention des *IP* pour les méthodes Δ SCF; un premier calcul pour identifier l'OM d'intérêt, puis un deuxième calcul *open-shell*.

Avec Δ HF (Bagus [10, 12]), Naves de Brito *et al.* [137] obtiennent des *IP* de cœur pour l'oxygène et le carbone dans des molécules organiques avec une précision de quelques dixièmes d'électronvolt.

Dans la même lignée, nous avons la possibilité d'utiliser la *Density Functional Theory (DFT)* : Δ DFT comme l'article de Besley *et al.* [20] pour les orbitales 1s du carbone, de l'azote et de l'oxygène. Mais, Fouda *et al.* [64] ou encore Pueyo Bellafont *et al.* [152] montrent que cette méthode est sensible à la *fonctionnelle*; ces derniers parviennent tout de même, pour la *IP-1s* de l'azote de différentes molécules, dans les limites de 0,3 eV d'erreur en utilisant B3LYP. Nous noterons l'intérêt de cette méthode pour étudier des molécules organiques de taille conséquente, telle que la 1,3,5,2,4-benzotrithiadiazepine $C_6H_4S_3N_2$ de l'article de Takahata *et al.* [196] toujours pour les *IP-1s*.

7. De la même façon, nous aurions, à partir d'un anion : $X^- \xrightarrow{h\nu} X + e^-$.

Enfin, il existe Δ MP2 comme dans l'article de South *et al.* [184]. À noter, pour cette dernière méthode, le travail de Shim *et al.* [181] où pour le calcul de Δ MP2 une *Base* est utilisée pour l'atome d'intérêt et un pseudo-potentiel pour les atomes « spectateurs » (Δ MP2-MIX).

Mais il existe également Δ CC et nous allons en présenter deux variations. Dans un premier temps, comme pour Δ SCF, nous pouvons faire un calcul direct en séparant les *spin-orbitales*, Watts *et al.* [206]. Dans un deuxième temps, nous pouvons réaliser une projection lors du calcul *HF* pour laisser libre l'orbitale correspondant au radical à considérer : c'est la technique *Core-Valence Separation (CVS)*. CVS est un terme *générique* qui indique le fait de marquer une différence entre les électrons de cœur et les électrons de valence [33]. Cette méthode donne de bons résultats pour les *OM* de cœur de par leur nature « localisée ». Comme l'indiquent Zheng *et al.* [212], en Δ CCSD avec les *IP-1s* des composés contenant les atomes {C; N; O; F}, l'erreur moyenne est de 0,17 eV et l'écart-type est de 0,21 eV par rapport aux valeurs expérimentales.

Nous pouvons nous intéresser maintenant à quelques méthodes « directes », comme avec la technique *Complete Active Space : (CAS)SCF - CASPT2*. Dans l'article de Lundberg *et al.* [126] où le complexe de fer $[\text{Fe}(\text{H}_2\text{O})_6]^{2+}$ a été étudié, et dont les résultats correspondent aux spectres expérimentaux *XAS*, le choix de l'espace actif reste toujours complexe et pour l'instant l'étude est limitée à un seul atome. Il existe également, pour le même système, l'article de Grell *et al.* [75] avec l'étude du couplage *spin-orbitale*⁸.

Maganas *et al.* [131] utilisent les méthodes *Multi-Reference Configuration Interaction (MRCI)* et *Multi-Reference Equation of Motion (MREOM)* pour simuler les spectres *XAS* et notamment la couche L du fer pour les complexes $[\text{FeCl}_4]^{2-}$ et $[\text{FeCl}_4]^-$. Pour la méthode *MRCI*, l'erreur provient surtout des manques *Size-Extensive* et relaxation, alors que la méthode *MREOM* est capable d'un meilleur résultat et avec un espace actif plus petit.

Pour la *Time-Dependent Density Functional Theory (TD-DFT)*, les travaux de Besley [18, 19] montrent la possibilité d'utiliser cette méthode dans différents

8. Pour limiter les coûts nous pourrions utiliser la méthode *RASSCF/RASPT2*. Cette méthode est utilisée, pour la première ligne des métaux de transition, dans l'article de Vancoillie *et al.* [199], pour la valence, mais avec cette méthode c'est alors le choix de l'espace actif qui devient encore un peu plus compliqué.

domaines de la spectroscopie. La difficulté est le choix de la *fonctionnelle*, bien que *TD-DFT* soit moins sensible que la Δ DFT [64]. Cela offre l'opportunité de travailler avec des systèmes comprenant un grand nombre d'atomes (comme dans l'article de Brena *et al.* [26] avec metal-(free)-phthalocyanines : Ni-(H₂)-C₃₂N₈H₁₆) pour expliquer les shake-up lines⁹.

Nous pouvons citer *Algebraic Diagrammatic Construction (ADC)*; cette méthode est basée sur un *état fondamental* MÖLLER-PLESSET. Elle donne de bons résultats par le fait d'intégrer dans sa construction matricielle les configurations doublement excitées comme le montrent Wenzel *et al.* [208] pour les *EE* de molécules organiques. De plus, nous pouvons appliquer *CVS* sur cette méthode [50].

Enfin, nous parlons de *Linear Response Coupled-Cluster (LR-CC)*; cette méthode est basée sur le même principe que *EOM*, mais diagonalise la matrice \mathcal{H} comme dans l'article de Koch *et al.* [105]. Dans l'article de Coriani *et al.* [41], nous remarquons l'utilisation de *CVS* pour obtenir les *EE* de cœur. Les *IP* peuvent être obtenus en combinant *CVS* avec une excitation restreinte vers une orbitale simulant le continuum; ce qui est possible par l'utilisation de *Base* possédant une primitive très diffuse.

2.2.3 Vers le chapitre 3

Après cette présentation, forcément incomplète, de quelques méthodes, nous pouvons passer à *Equation of Motion (EOM)*. Cette méthode « directe » pouvant être utilisée de façon « black-box », contrairement aux méthodes nécessitant le choix d'un espace actif, est capable de donner les *IP*, les *EE* et les *EA*. Et elle ne manque ni de corrélation par son *état fondamental CC* ni de relaxation par les configurations (2h/2p) ou (2h/1p) qu'elle intègre. L'article de Shee *et al.* [179] a déjà montré les avantages de cette méthode dans un cadre relativiste, mais les propriétés du cœur restaient encore à étudier. Dans ce sens, la méthode *CVS* a été implémentée comme nous le montrons au chapitre 3.

9. Ce sont des raies « d'excitations » qui apparaissent dans un spectre *XPS*.

PROBLÉMATIQUE 2 : QUELLES APPROXIMATIONS RÉALISER ?

Nous voulons obtenir les EE , EA , IP en utilisant la méthode *Equation of Motion Coupled-Cluster Singles Doubles (EOM-CCSD)* (cf.1.5.2). Résoudre ces équations aux valeurs propres occasionne des coûts computationnels respectivement en $\mathcal{O}(N^6)$, $\mathcal{O}(N^5)$ et $\mathcal{O}(N^5)$.

Avant de présenter nos travaux, nous présenterons quelques techniques permettant de simplifier les calculs et de réduire les temps computationnels. Trois grandes familles existent : nous pouvons simplifier la façon d’obtenir les intégrales ; nous pouvons diminuer le nombre d’intégrales intervenant en réduisant l’espace des OM ; nous pouvons utiliser des développements perturbatifs pour simplifier le nombre de termes présents dans chaque bloc de la matrice (développements perturbatifs).

Nous ne présenterons uniquement que les méthodes en rapport avec cette thèse, déjà implémentées dans ExaCorr¹⁰ ou en passe de l’être.

Nous sommes intéressés par les IP , mais les méthodes seront présentées de façon générale comme nous l’avons fait pour l’introduction de l’opérateur \hat{R} . Nous finirons ce travail sur les coûts computationnels.

2.3 Approximations

2.3.1 Approximations dans l’origine des intégrales

Décomposition de CHOLESKY.

La première méthode dont nous parlerons est la décomposition de CHOLESKY ; cette méthode est basée sur les approximations des intégrales 2-électrons. L’article de Beebe *et al.* [16] présente à la fois la technique et les erreurs introduites. Cette méthode est déjà présente dans le module ExaCorr de *Dirac* ; nous l’avons utilisée, notamment dans la comparaison de nos résultats avec les codes MRCC [99], quand nous avons travaillé à l’implémentation de $CC2$ (cf.5.4). À

10. Nouvelle implémentation dans *Dirac* du code CC pour des architectures de calculs hautement parallélisées (cf.5).

l'avenir, cette méthode pourra être appliquée aux équations *CCSD* et *EOM-CC* comme dans l'article Epifanovsky *et al.* [58].

Dans la base des *OA*, l'intégrale 2-électrons s'écrit :

$$(\mu\nu | \lambda\sigma) = \int \chi_\mu(r_1) \chi_\nu(r_1) \frac{1}{|r_1 - r_2|} \chi_\lambda(r_2) \chi_\sigma(r_2) d^3r_1 d^3r_2 \quad (2.3.1)$$

est réécrite comme :

$$(\mu\nu | \lambda\sigma) \approx \sum_{P=1}^M B_{\mu\nu}^P B_{\lambda\sigma}^P, \quad (2.3.2)$$

où « M » est le rang de la décomposition. L'obtention des tenseurs *B* est basée dans un premier temps sur le calcul des intégrales $(\lambda\sigma|\lambda\sigma)$ (la "diagonale" de (eq.2.3.1)). Généralement « M » est de 3 à 10 fois le nombre de fonctions de *Base* « N », cela dépend du seuil de précision demandé, de l'ordre de 1×10^{-9} . Mais M reste toujours très inférieur à la taille de la matrice $(\mu\nu | \lambda\sigma)$ en $N(N+1)/2$ [103]. Il y a également un gain sur l'enregistrement car ces tenseurs ont une taille en $\mathcal{O}(N^3)$. En utilisant la décomposition de *CHOLESKY*, l'erreur en énergie peut être de l'ordre de $1 \times 10^{-4} E_h$ pour un gain de temps de 15 à 20 %.

Réduction de l'espace des virtuelles

En pratique, nous avons toujours la possibilité de faire l'approximation de ne pas transformer toutes les *OM* virtuelles (non plus que toutes les occupées) pour les calculs corrélés. Les calculs sur la valence ne nécessitent pas toujours la prise en compte de toutes les orbitales virtuelles [72] et nous avons montré au chapitre précédent qu'il en est de même pour les orbitales occupées du cœur.

La seconde technique dont nous parlerons est la technique *Frozen Virtual Natural Orbital (FVNO)*. Cette technique est basée sur la réduction de l'espace des *OM* virtuelles canoniques, remplacées alors par des *OM* naturelles et nous calculerons l'énergie de corrélation par l'utilisation d'une base plus compacte.

Nous pouvons citer les travaux de Landau *etal* [112] et de Taube *et al.* [197]. Ils décrivent ainsi *Frozen Virtual Natural Orbital (FVNO)* : dans un premier temps, la matrice densité de la partie virtuelle (*single-particle*) est calculée au niveau MBPT(2), les blocs occupé-virtuel et virtuel-occupé étant mis à zéro. Il n'y a pas

de changement dans les orbitales occupées.

Le bloc virtuel-virtuel de la matrice densité est diagonalisé, pour obtenir les *Natural Orbitals* et nous choisissons de garder seulement celles possédant un « nombre d'occupation » suffisamment élevé.

Nous retransformons cette base tronquée sur la base des OA. Nous pouvons alors construire la matrice EOM. Landau *et al.* [112] s'intéressent également à la façon de sélectionner n_{crit} , soit par un certain pourcentage du nombre total des virtuelles soit par un certain pourcentage de l'occupation totale. Enfin, Höfener *et al.* [87] travaillent sur la quantité de virtuelles pouvant être éliminées.

Cette méthode est depuis peu implémentée dans *Dirac* par Yuan *et al.* [211] dans une version encore en développement.

2.3.2 Approximations perturbatives des termes de la matrice

$$\mathcal{H}_{\text{eff}}$$

Nous pouvons commencer par reprendre le schéma introduit par de Izsák [93] pour hiérarchiser différentes approximations possibles pour la matrice hamiltonienne. Nous reproduisons la figure 1 de cet article sur la figure (fig.2.6) modifiée pour correspondre à nos besoins. Le cadre « A » de cette figure sera traité dans le chapitre 4, ce sont les approximations *Partitioned - Equation of Motion Coupled-Cluster (P-EOMCC)*, *Equation of Motion Coupled-Cluster - Many-Body Perturbation 2d order (EOMCC-MBPT(2))* et *Partitioned - Equation of Motion Coupled-Cluster - Many-Body Perturbation 2d order (P-EOMCC-MBPT(2))*.

Intéressons-nous au cadre « B », qui représente d'autres approximations couramment utilisées avec lesquelles nous n'avons pas réalisé de calcul.

La partie CC2 sera traitée au niveau de l'état *fondamental* dans le chapitre 5 (cf.5.4). Nous pouvons déjà introduire l'idée maîtresse : les équations pour les amplitudes t_1 restent identiques ; les amplitudes t_2 sont traitées à l'ordre 1 en perturbation. D'un point de vue de la matrice EOM, les blocs $\{\mathcal{H}_{DS}; \mathcal{H}_{DD}\}$ sont donc modifiés en conséquence.

Si nous développons également les équations t_1 à l'ordre 1 (comme pour EOM-MBPT(2)), les blocs $\{\mathcal{H}_{SS}; \mathcal{H}_{SD}\}$ sont modifiés en conséquence. La méthode

devient alors MP2-CC2 par Tajti *et al.* [194]¹¹. Enfin, si ce bloc $\mathcal{H}_{DD}^{[0]}$ est résolu en utilisant la première itération d'un calcul CI, nous parlons de CIS(D) [83].

Pour *Algebraic Diagrammatic Construction (ADC)*, l'idée initiale a été introduite par Schirmer [167], puis développée notamment dans les articles [168, 208, 138, 50, 195]. Partant d'un *état fondamental MP2*, les blocs sont développés à des ordres différents pour un coût computationnel en $\mathcal{O}(N^5)$. Il est à noter que le bloc \mathcal{H}_{DD} est diagonal et surtout que la matrice est hermitienne, ce qui est une simplification importante par rapport aux méthodes CC.

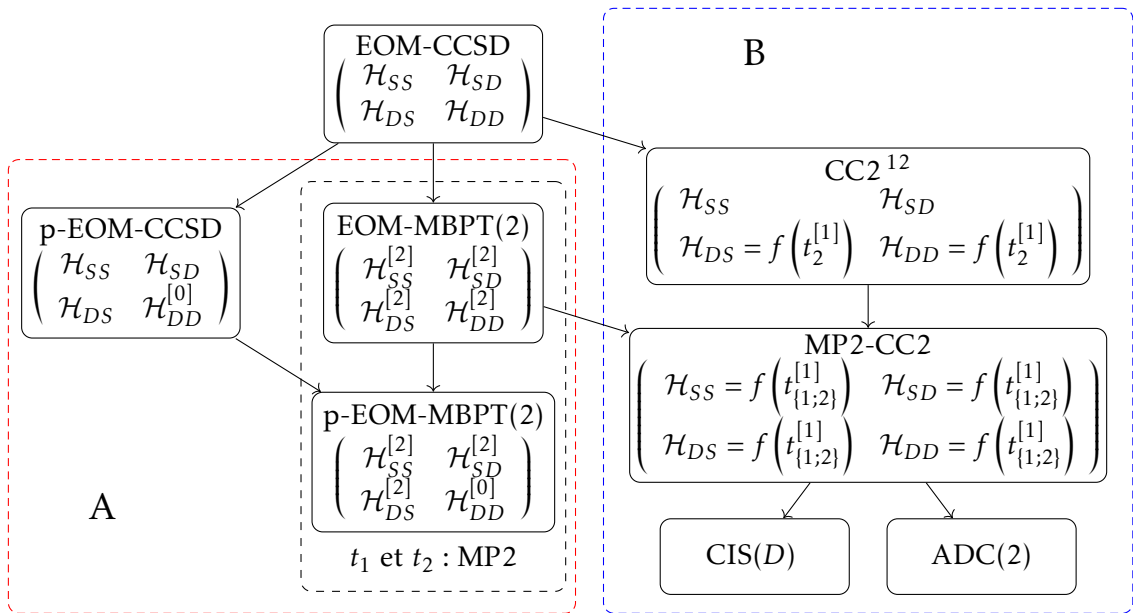


FIGURE 2.6 – Les différentes approximations de la matrice EOM. Les termes sont donnés dans le texte.

Rappelons le scaling pour les EE de quelques méthodes [123] sur le tableau (tab.2.2)¹³.

11. $CISD_{\infty}$ [82].

12. Il est intéressant de noter que CC2 impose pour le bloc \mathcal{H}_{DD} que t_2 soit défini à l'ordre ⁽¹⁾ ce qui impose que le bloc soit diagonal et identique au bloc $\mathcal{H}_{DD}^{[0]}$ de *P-EOMCC*, i.e. : $CC2[\mathcal{H}_{DD}(t_2^{(1)})] = P-EOMCC[\mathcal{H}_{DD}^{(0)}]$.

13. Reprinted (adapted) with permission from Pierre-François Loos, Anthony Scemama, Denis Jacquemin - The Quest for Highly Accurate Excitation Energies : A Computational Perspective - Journal of Physical Chemistry Letters 2020, 11, 2374-2383. Copyright 2020 American Chemical

Méthode	Coût Computationnel	Erreur typique (eV)
TD-DFT	$\mathcal{O}(N^4)$	0,2 - 0,4
CIS	$\mathcal{O}(N^4)$	$\sim 1,0$
CIS(D)	$\mathcal{O}(N^5)$	0,2-0,3
ADC(2)	$\mathcal{O}(N^5)$	0,1-0,2
CC2	$\mathcal{O}(N^5)$	0,1-0,2
ADC(3)	$\mathcal{O}(N^6)$	0,1-0,2
EOM-CCSD	$\mathcal{O}(N^6)$	0,1-0,2
CC3	$\mathcal{O}(N^7)$	$\sim 0,04$
EOM-CCSDT	$\mathcal{O}(N^8)$	$\sim 0,03$
EOM-CCSDTQ	$\mathcal{O}(N^{10})$	$\sim 0,01$
CASPT2	$\mathcal{O}(N^5); N!$	0,1-0,2
FCI	$\mathcal{O}(N!)$	0,0

TABLEAU 2.2 – Coût de quelques méthodes permettant d’obtenir *les énergies d’excitation*, associé à l’erreur typique en (eV) pour des molécules organiques. N est le nombre d’orbitales moléculaires [123].

2.3.3 Vers le chapitre 4

Ainsi, après avoir vu qu’il était possible, par différentes techniques de réduire les coûts computationnels, nous présenterons, dans le chapitre 4, les différentes approximations que sont *Partitioned-EOMCC (P-EOMCC)*, *EOMCC - Many-Body Perturbation Theory (2) (EOMCC-MBPT(2))* et la combinaison des deux approximations précédentes *P-EOMCC-MBPT(2)*.

PROBLÉMATIQUE 3 : EXPLOITER DE FAÇON EFFICACE DES PLATEFORMES *HPC* POUR DES CALCULS CORRÉLÉS

La chimie quantique, notamment dans les méthodes post-*HF*, a besoin d'une puissance de calcul considérable. Lors de la section précédente, nous avons évoqué les coûts de calcul de différentes méthodes en indiquant les ordres de grandeur, mais sans entrer dans les détails du nombre d'opérations réalisées ou de la mémoire utilisée. Les équations des méthodes *CC* et *EOM-CC*, sont basées sur une succession de produits tensoriels [178]. De nombreuses bibliothèques existent pour les calculer, notamment sur des architectures du type de celles des supercalculateurs : *High-Performance Computing (HPC)*. Psarras [151] n'en dénombre pas moins de 75¹⁴.

Avant de montrer l'avancée de nos travaux, nous présentons brièvement quelques-uns des éléments nécessaires aux *HPC* d'un point de vue hardware puis software.

2.4 High-Performance Computing (HPC)

HPC

Les besoins croissants en calculs ont nécessité la création de *High-Performance Computing (HPC)*. Ils sont utilisés pour les prévisions météorologiques (BELENOS pour MétéoFrance [17]), les simulations physiques, la cryptanalyse et l'intelligence artificielle¹⁵... À l'heure actuelle, un système est considéré comme une *HPC* s'il est capable de réaliser 1×10^{12} FLOPS¹⁶.

Les défis sont donc nombreux dans la réalisation de *HPC* d'un point de vue Hardware : miniaturisation, communication entre les microprocesseurs, problème thermique, *fault tolerance*... et Software : parallélisation, partage des données, synchronisation...

14. Et cela en comptant uniquement les bibliothèques pouvant fonctionner indépendamment, omission de TMM et TCE incluses dans NWChem, omission de BLIS sur QChem.

15. Projet Grands Challenges.

16. *floating-point operations per second*.

FLOPS et évolution des HPC

Il est intéressant de noter quelques-unes des évolutions de la puissance de calcul des *HPC*, notamment parce que l'équipe de PCMT a travaillé sur certains de ces systèmes. L'évolution des *HPC* est connue; nous noterons que le gigaFLOPS est atteint dans les années quatre-vingts, le téraFLOPS dans les années quatre-vingt-dix, le pétaFLOPS et la dizaine de pétaFLOPS sont atteints respectivement en 2008 et 2011. En mars 2012, Curie, créé pour le GENCI [68], dispose d'une puissance de 1,36 pétaFLOPS. En 2013, un ordinateur personnel peut dépasser les 200 gigaFLOPS (équivalent d'un *HPC* de 1995), voire les 5,6 téraFLOPS (équivalent d'un *HPC* de 2001). En janvier 2015, le système Occigen, pour le GENCI (installé sur le site du CINES [39]) est doté d'une puissance de 2,1 pétaflops. Les 200 pétaFLOPS sont obtenus en 2018 avec Summit [127, 142, 191]. En 2020, le supercalculateur Jean Zay [95] (IDRIS [90]) développe 29 pétaFLOPS. Cette même année, NVIDIA commercialise la plateforme Ampère d'une puissance de 19,5 téraFLOPS (équivalent d'un *HPC* de 2007 - 11 k\$) et un système DGX A100 de 5 pétaFLOPS pour les besoins de l'intelligence artificielle (équivalent d'un *HPC* de 2011 - 200 k\$).

L'exaFLOPS, prochaine étape, est en cours de développement. CRAY l'annonce pour la fin de 2021 à Argonne [8](500 M\$).

Travailler sur ces *HPC* présente des difficultés d'ordre pratique notamment d'avoir des durées de calculs inférieures à 24 heures. L'équipe PCMT du PhLAM possède un cluster nommé **Sakura**, utilisé pour s'affranchir de cette limite de temps mais également pour du développement [79, 179, 143]. La puissance de ce calculateur n'est pas mesurable puisque les presque 900 coeurs répartis sur une quarantaine de nœuds ne travaillent pas de concert.

2.4.1 Hardware

Depuis l'invention des transistors à la fin des années quarante, des microprocesseurs dans les années soixante-dix et des multicœurs dans les années deux mille, les technologies n'ont cessé d'évoluer. Le principe est simple : placer, dans un minimum de place, un maximum de *Central Processing Unit (CPU)* pouvant

travailler rapidement et de concert.

Au milieu de cette même décennie, il est apparu que les Cartes Graphiques *Graphics Processing Unit (GPU)*, alors spécialisées pour les jeux et les applications multimédias, étaient capables de développer plusieurs centaines de gigaFLOPS (~350 pour la G80GL quadri-cœur), alors que les *CPU*, à caractère généraliste, se cantonnaient à sept fois moins (~50 pour Intel Core2 Duo). Les cartes graphiques ont donc été de plus en plus utilisées dans les calculateurs et font désormais partie des *HPC*. Nous pouvons regarder, à titre d'exemple, l'architecture de connexion pour deux nœuds dans Summit, où nous pouvons observer à la fois des CPU et des GPU (fig.2.7).

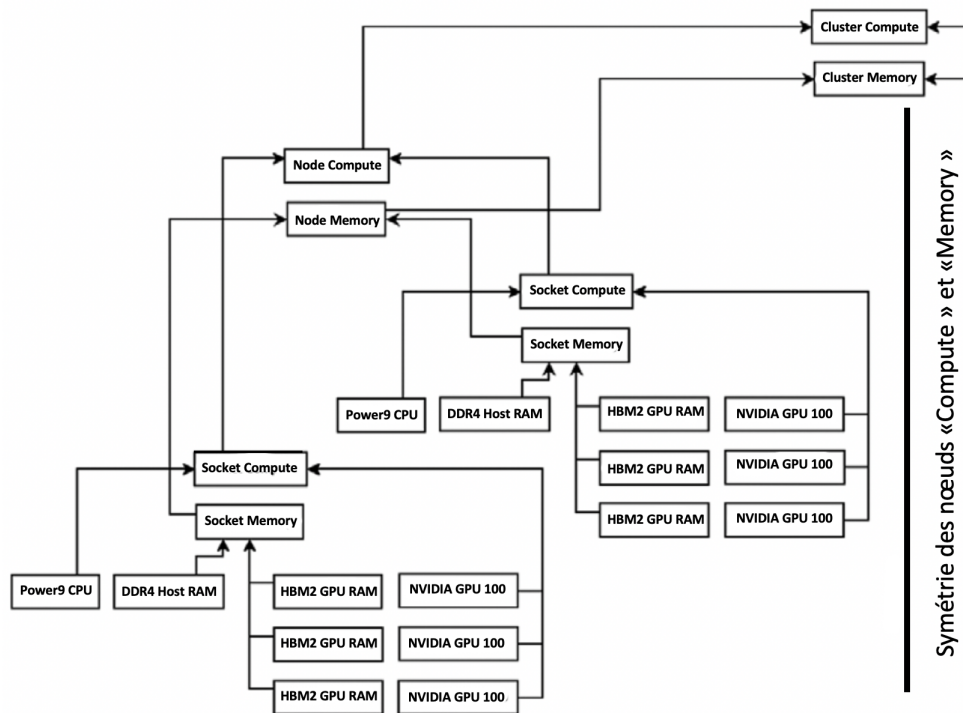


FIGURE 2.7 – Représentation de l'architecture hardware de 2 nœuds de Summit [129]¹⁷.

17. Dmitry I. Lyakh, Domain-specific virtual processors as a portable programming and execution model for parallel computational workloads on modern heterogeneous high-performance computing architectures, International Journal of Quantum Chemistry, Mar 12, 2019, 119 :e25926. Reprinted with permission of Copyright Clearance Center. All rights reserved.

Une fois les *CPU* et *GPU* placés, le but a été de les relier. Nous sommes passés alors à l'interconnexion de ces *nœuds* de calcul à l'aide de *switchs*. Les communications dans Summit se font avec des *switchs* IBTM 2 EDR Mellanox à 23 Gbit/s.

Ensuite, arrivent des problèmes de consommation énergétique : "Tous les appareils transforment une partie de l'énergie électrique qu'ils reçoivent en énergie thermique : c'est l'effet Joule [97]". Summit fonctionne avec une puissance de 13 MW, Jean Zay consomme une énergie de 2 MW h. Il est alors nécessaire de refroidir ces systèmes; des « armoires » spécialisées existent à cet effet comme la HPE SGI 8600 [88].

Enfin, il est à noter l'apparition de *Field-Programmable Gate Array (FPGA)* qui est un réseau de portes programmables in situ. Ces *FPGA* servent pour de l'électronique embarquée. Les vitesses d'horloge sont plus faibles mais la puissance consommée rend ces processeurs intéressants : la puissance électrique par opération est au moins divisée par 2 (tab.2.3).

Chip [77]	IntelSkyLake [92]	NVidia Pascal [140]	Xilinx XCVU9P [210]
Technologie	14 nm	16 nm	16 nm
Puissance	205 W	300 W	<50 W
Fréquence	2,7 GHz	1,5 GHz	0,1-0,5 GHz
DRAM	128-768 GB	0	48 GB

TABLEAU 2.3 – Exemple de CPU, GPU et de FPGA.

2.4.2 Software

Taxonomie de Flynn

En 1966, Flynn [62] décrit ce que seront les possibilités du transfert des données et des instructions, appelées depuis Taxonomie de Flynn (tab.2.4).

Les processus SISD sont purement séquentiels ; les SIMD sont utilisés par les

		Flux d'instructions	
		unique	multiple
Flux de données	unique	SISD	MISD
	multiple	SIMD	MIMD

TABLEAU 2.4 – Taxonomie de Flynn : représentation des différentes distributions possibles. Respectivement *S* et *M* pour *Single* et *Multiple*; *I* et *D* pour *Instruction* et *Data*.

processus vectoriels¹⁸. L'utilisation de MISD est très spécifique; nous pouvons utiliser cette technique pour vérifier, par redondance, la présence d'erreurs [186].

Dans MIMD, la mémoire peut être partagée. Pour cela, il faut un mécanisme de contrôle afin d'éviter les conflits d'écriture. De plus, il doit y avoir une organisation physique adéquate des différents nœuds pour que la communication soit efficace; il faut prêter attention au *degré* et au *diamètre* qui sont respectivement le nombre de liaisons sortantes de chaque nœud et la distance minimum, en nombre de nœuds, entre 2 nœuds communicants.

Pour simplifier la programmation, une même application est exécutée sur tous les processeurs : *Single Program Multiple Data (SPMD)*. L'exécution se fait sur plusieurs processeurs avec la même copie du code. Chaque copie a un rang différent; nous avons alors parallélisme des données et du contrôle. C'est cette technique qui est généralement utilisée pour les calculs scientifiques avec MPI et OpenMP et sur les *Application Programming Interface (API)* travaillant avec des GPU.

Communication {intra; inter}-nœuds.

Pour pouvoir communiquer les informations entre différents processeurs, nous pouvons utiliser les systèmes à mémoire partagée ou à mémoire distribuée programmables respectivement avec OpenMP et MPI. Les inconvénients et les avantages sont propres à chacun des systèmes.

18. Ensemble de données de même taille et de même type et chaque vecteur est traité indépendamment : traitement de l'image.

OpenMP : Open Multi-Processing La mémoire est partagée ; la programmation est plus facile car il y a juste à spécifier les parties du programme à paralléliser. Le programme est une succession de régions séquentielles et parallèles. Le problème est de deux ordres : le réseau de connexions entre les *données* et les *processeurs* limite la bande passante et il peut y avoir des problèmes de *race condition*¹⁹ ; il faut alors parfois imposer une synchronisation.

L'utilisation courante se fait sur un système de plusieurs cœurs d'une même machine, ce qui est pratique, par exemple, pour découper une boucle *for*. Cet *API* est utilisé pour les processus SIMD.

MPI : Message Passing Interface MPI est également un *API* de communication, mais entre différents processeurs (nœuds), la mémoire est distribuée. Théoriquement, il n'y a pas de limite au nombre de nœuds qui peuvent travailler de concert ; la puissance de calcul est donc plus importante que pour les processus OpenMP. La programmation est plus complexe, car c'est au programmeur de gérer la communication entre les nœuds.

GPU : Graphics Processing Unit Les cartes graphiques *GPU* ont la particularité de pouvoir réaliser de nombreux calculs parallélisés ; ceci est permis grâce à leur ingénierie optimisée pour le transfert d'informations entre leurs cœurs.

Différents *API* ont été développés pour utiliser cette puissance de calcul (communication entre les nœuds et gestion des *threads*), notamment *Compute Unified Device Architecture (CUDA)* développé par NVIDIA™ et *OpenCL* pour les cartes AMD.

Hybride Une application peut être hybride ; la communication entre nœuds est assurée par MPI et le travail dans chaque nœud multi-cœur par Open-MP.

2.4.3 Vers le chapitre 5

Devant ces éléments, il a été décidé d'introduire dans *Dirac* la bibliothèque ExaTensor [129] créée par Dmitry Lyakh, permettant de réaliser ces calculs

19. Deux processus peuvent fonctionner à des vitesses différentes, ce qui peut provoquer des conflits en sortie ; de même, des problèmes peuvent apparaître s'ils sont interdépendants.

tensoriels de façon à ce qu'ils soient hautement parallélisés. Cette bibliothèque gère en interne (*API*) la distribution des données, les calculs et la récupération des résultats. Dans *Dirac* cette bibliothèque a été intégrée dans un nouveau module *ExaCorr*.

Equation of Motion Coupled-Cluster et Core-Valence Separation

Dans ce chapitre, nous montrerons comment obtenir *le potentiel d'ionisation (IP)* et *l'énergie d'excitation (EE)* pour n'importe quelle *orbitale moléculaire (OM)* de cœur par l'utilisation de la technique *Core-Valence Separation (CVS)*. Notre intérêt se portera également sur les hamiltoniens relativistes et la correction que peut nous apporter la *Quantum Electrodynamics (QED)*. Nous avons travaillé notamment sur différents composés halogénés ou iso-électroniques de ces composés. Ces travaux ont conduit à l'article : « relativistic EOM-CCSD for Core-Excited and Core-Ionized State Energies Based on the Four-Component Dirac-Coulomb (-Gaunt) Hamiltonian » [79], reproduit en annexe C.

Nous voulons ici discuter de quelques étapes qui nous semblent importantes, premièrement dans l'étude de *CVS* et de différentes options et deuxièmement dans l'étude de différents hamiltoniens. La première partie nous donne accès aux *IP* et aux *EE* des électrons de cœur ; la deuxième étudie les voies à emprunter pour améliorer les résultats par rapport aux valeurs expérimentales.

Nous montrerons un exemple au travers de l'ion iodure I^- . L'iode est un composé intervenant dans des processus atmosphériques [63], mais également dans risques radioactifs [30]. La chimie de l'iode est complexe par ses différents degrés d'oxydation. Son noyau est suffisamment lourd pour que la vitesse relativiste de ses électrons de cœur ait de l'influence sur les énergies des *OM* ; de plus des

corrections QED sont à appliquer.

3.1 Utilisation de CVS avec EOM-CC

Dans le premier chapitre, nous avons parlé de l'algorithme de Davidson comme pouvant donner de façon itérative les valeurs propres (et les vecteurs propres associés) de la matrice *EOM*. Cependant, il y a un problème inhérent à l'algorithme de Davidson : il ne peut donner les valeurs propres que dans l'ordre croissant ou décroissant. Nous ne pouvons donc pas obtenir, par exemple pour *IP*, les énergies des électrons des *OM* de cœur sans avoir au préalable calculé toutes les énergies de plus hautes valeurs. Le problème n'est pas tant les *Simples* que les *Doubles*.

Prenons comme exemple le spectre complet de l'anion iodure I^- dans les limites $[0; 35]keV$ (fig.3.1)¹. Ce graphique est issu de la construction d'un histogramme de pas de 10 eV. L'axe des ordonnées de gauche représente le nombre de solutions dans chacun de ces intervalles. Celui de droite représente \log_{10} du cumul du nombre de solutions calculées. L'axe des abscisses représente l'énergie en eV. Pour cet exemple, une solution est considérée comme *Simple-IP* si au moins un des coefficients de r_i est tel que $\%SI > 1$. Nous remarquons que le nombre de solutions à calculer avant d'arriver à la *K edge* (proche de 35 keV) est de l'ordre de 1×10^5 ². Cette façon de faire est donc inenvisageable pour de plus gros systèmes : il nous faut alors une autre méthode.

Pour pallier ce problème, nous avons utilisé la technique : *Core-Valence Separation (CVS)*. C'est en s'appuyant sur ce concept que Dreuw *et al.* [50] et Coriani *et al.* [41], ont, indépendamment, formulé une modification des approches de diagonalisation (en l'occurrence avec *ADC* et avec *Linear Response Coupled-Cluster (LR-CC)*) de façon à cibler des états de cœur³. Dans ces deux cas, ils utilisent des

1. Ce spectre est également présenté dans [79].

2. Il y a plus de 3×10^5 solutions en tout avec cette base, il y en aurait une infinité en base complète.

3. Nous noterons que ce sont les difficultés rencontrées par Coriani *et al.* avec l'utilisation de l'algorithme de Lanczos [111] (problèmes de convergence pour les excitations de cœur et lors de l'utilisation de *Base importante*) qui les ont justement conduits à leur travail sur *Core-Valence Separation (CVS)*.

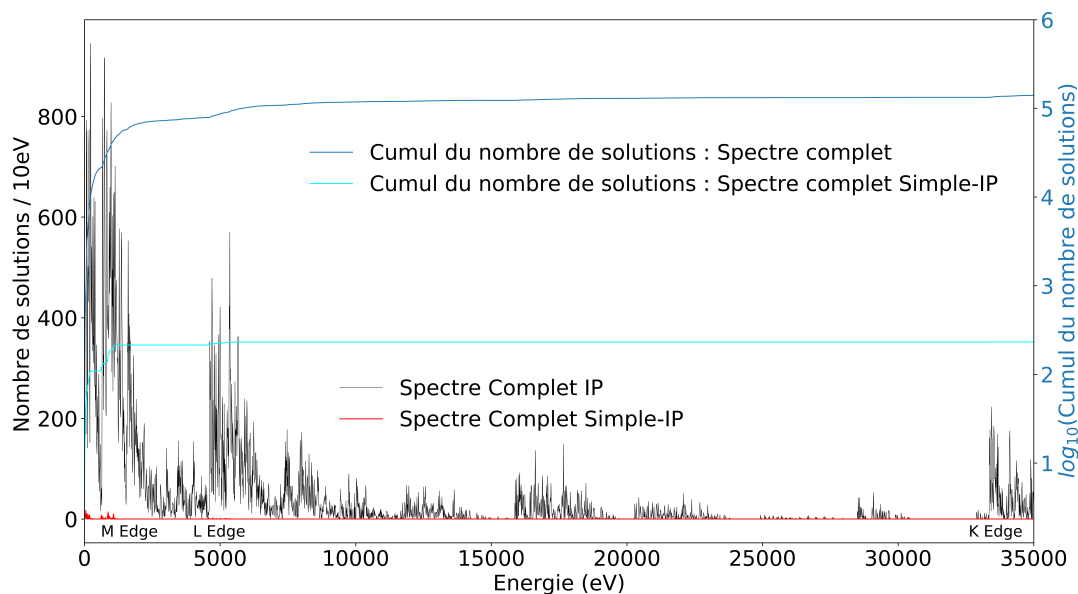


FIGURE 3.1 – Spectre des potentiels d’ionisation pour Γ . Représentation du spectre complet (en noir) et du spectre complet pour les simples ionisations (en rouge) (dans les limites $[0; 35]$ keV). Base : dyall-acv3 ζ . Hamiltonien ${}^2\text{DC}^{\text{M}}$. En abscisse : les énergies des différents potentiels d’ionisation (en eV). En ordonnée : sur la gauche le nombre de solutions par intervalles de 10eV, sur la droite le \log_{10} du cumul du nombre de solutions.

techniques de projection pour éliminer les contributions aux vecteurs propres venant de déterminants ayant au moins un des indices occupés dans la valence. Ainsi, avec une approche de type Davidson, il est désormais possible de cibler des racines intérieures ayant le bon caractère (*IP* ou *EE* de cœur). Nous avons utilisé l’approche de Coriani *et al.* pour implémenter *CVS-EOM* dans *Dirac* à partir de l’implémentation de *EOM* de Shee *et al.* [179].

Pour *EOM*, c’est une technique de projection (\mathcal{P}) qui retire les états de valence du vecteur d’essai ⁴.

4. Dans *Dirac* la limite entre le cœur et la valence se fait par le choix d’une valeur d’énergie.

Si I représente les OM de cœur, la projection s'effectue telle que :

$$\begin{cases} \mathcal{P}_I^{CVS} r_i^a = 0 & \text{si } i \neq I \\ \mathcal{P}_I^{CVS} r_{ij}^{ab} = 0 & \text{si } i \neq I \text{ ou } j \neq I \end{cases} \quad (3.1.1)$$

L'équation aux valeurs propres devient alors :

$$\mathcal{P}_I^{CVS} (\mathbf{H} \mathcal{P}_I^v \mathbf{r}_k) = E_k^{CVS} \mathcal{P}_I^{CVS} \mathbf{r}_k \quad (3.1.2)$$

Nous ne changeons donc pas la taille de matrice $EOM-CC$; nous ne faisons que mettre tous les termes de valence à 0. Vidal *et al.* [200] présentent, a contrario dans Q-Chem [175], la possibilité de ne travailler qu'avec les OM d'intérêt dès le début du calcul et gagner d'autant sur les coûts computationnels. Distinguer les intégrales ayant des orbitales de cœur et de valence n'est pas implémenté dans *Dirac* et nécessitera encore du travail ; notre intérêt, ici, se porte sur l'application de cette méthode pour des éléments lourds ; nous pourrions toujours intégrer cette fonctionnalité après en avoir vérifié les qualités.

En utilisant cette technique de projection, nous sommes alors en mesure d'obtenir les *potentiels d'ionisation (IP)* de cœur avec $CVS-EOM$. Nous obtenons alors la figure (fig.3.2). À noter que la limite en énergie pour CVS a été choisie à une valeur de $2 E_h$ au-dessus de l' OM d'intérêt.

Nous remarquons que les valeurs données par CVS concordent fidèlement avec le spectre complet *Simple-IP*.

3.2 Utilisation des différentes options pour CVS : *Frozen Core (FC)* et *No-Double (ND)*

Nous avons voulu par la suite étudier certaines options pour simplifier encore les calculs et gagner d'autant en coût. Nous avons appliqué deux options, la première est *Frozen-Core (FC)*, avec laquelle nous imposons que les amplitudes t_i^a et t_{ij}^{ab} soient nulles si i et j correspondent à des OM du cœur :

$$\mathcal{P}^{cœur} |\Psi_{CC}\rangle \Rightarrow t_i^a = t_{ij}^{ab} = 0 \quad \text{si } i, j \in \text{cœur} \quad (3.2.1)$$

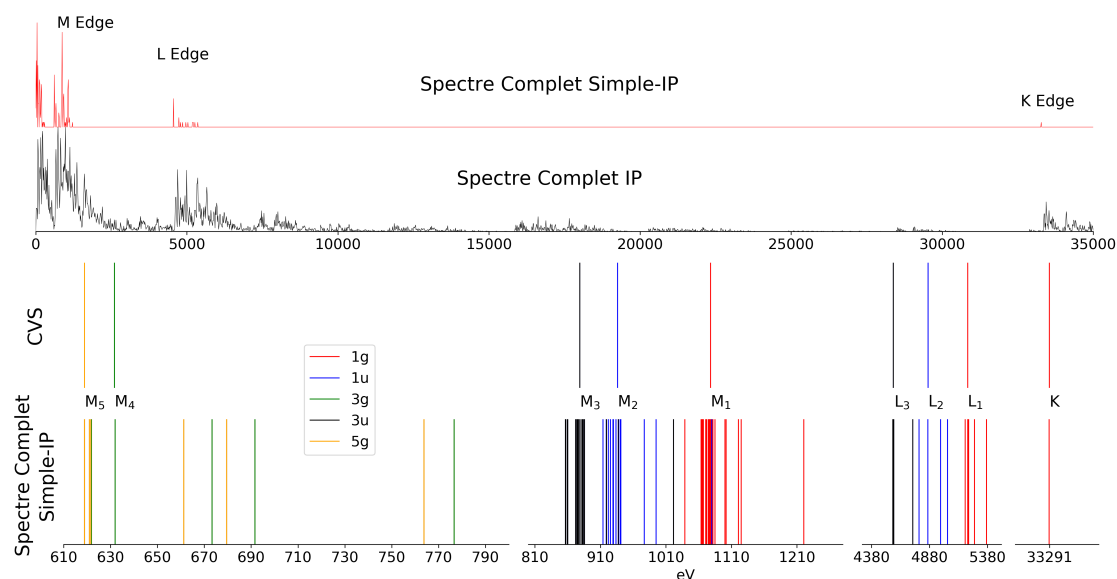


FIGURE 3.2 – Énergies des ionisations de cœur de l’ion iodure I^- à partir de la diagonalisation complète de \bar{H} . (En haut) Représentation du spectre des valeurs propres jusqu’à 35 keV, séparée entre tous les états (noir) et ceux ayant une contribution Simple-ionisation non négligeable (rouge). Les valeurs en ordonnée correspondent au nombre d’états dans un intervalle de 10 eV, avec la plus grande valeur ramenée à 1. (En bas) États Simple-ionisation obtenus par la diagonalisation complète de \bar{H} et par l’approche *Core-Valence Separation*. Les labels ng/nu indiquent la valeur absolue de la projection du moment angulaire ($|m_j| = n/2, n = 1, 3, 5$), pour les composantes des spineurs de symétrie gerade (g) ou ungerade (u). Tous les calculs sont réalisés avec l’hamiltonien ${}^2DC^M$.

La seconde option est *No Doubly core hole determinants* (ND); avec cette option nous imposons que les configurations avec deux indices de cœur soient nulles :

$$\mu r_{ij}^{ab} = 0 \quad \text{si } i, j \in \text{cœur} \quad (3.2.2)$$

Les zones d’influences de ces trois options CVS, FC et ND sont schématisées pour la matrice *EOM-IP* sur la figure (fig.3.3).

5. Adapté d’un schéma de [50] Michael Wormit, Andreas Dreuw, The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states, International Journal of Quantum Chemistry, WILEY INTERDISCIPLINARY REVIEWS : COMPUTA-

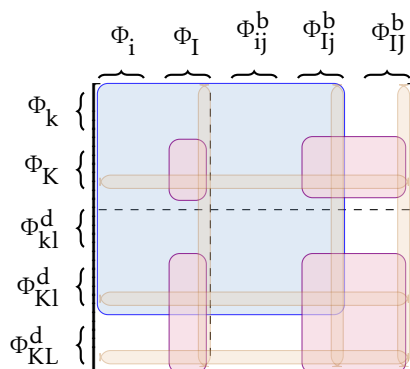


FIGURE 3.3 – Représentation de la matrice *EOM* pour les *potentiels d'ionisation* (*IP*) pour les différentes options (*CVS*), (*FC*) et (*ND*) - Les lettres capitales représentent les électrons de cœur. Les parties colorées violettes, bleues et marron représentent respectivement (*CVS*), (*ND*) et (*FC*)⁵.

Pour l'option *FC*, la limite séparant la partie à « geler » est donnée par une valeur en énergie, indépendante de la limite *CVS*. Nous pouvons également borner les orbitales à « dégeler ». Nous avons utilisé quatre limites différentes. Premièrement, nous gelons dès la séparation valence / cœur (*FC-V*), « *V* » pour *valence* (pour l'iodure I^- cela correspond à la limite $4d/5s$). Deuxièmement, nous gelons à la même énergie que *CVS* (*FC-f*), « *f* » pour *follow*. Troisièmement, nous gelons dès la valence, mais nous dégelons l'*OM* d'intérêt (*FC-V-Ex*), « *Ex* » pour *Except*. Quatrièmement, nous gelons toutes les *OM* ayant une énergie inférieure à l'*OM* d'intérêt (*FC-fpMO*), « *fpMO* » pour *follow previous Molecular Orbital*. Les différentes limites sont données dans le tableau (tab.3.1).

Nous obtenons pour l'iodure I^- les résultats présentés sur le graphique (fig.3.4).

Dans tous les cas, geler des *OM* de cœur, option (*FC*), i.e. exclure ces orbitales du calcul *CCSD*, conduit à une diminution des *IP* par rapport à *CVS*. Cette baisse est moins sensible dans le cas (*FC-fpMO*). Dans le cas où nous n'excluons pas l'*OM* d'intérêt, (*FC-V-Ex*), les résultats ne sont pas meilleurs et se révèlent même pires dans le cas des *OM* de cœur.

		Limite pour le vecteur d'essai		Limites pour les OM à geler				
		Core-Valence Limit	CVS Limit	FC	FC follow	FC Except Start	FC Except End	FC fpMO
LUMO								
HOMO								
HOMO	-p							
OM	+q	✓		✓				
OM	+1	✓		✓				
<u>OM</u>		✓	✓	✓	✓			
OM	-1	✓	✓	✓	✓	✓	✓	✓
OM	-r	✓	✓	✓	✓	✓	✓	✓

TABLEAU 3.1 – Représentation des différentes limites pour les approximations utilisées pour EOM-IP-CVS. OM est l'orbitale moléculaire d'intérêt et $\{p; q; r\}$ des nombres entiers.

L'option d'exclure les déterminants doublement excités, (ND), conduit à une augmentation des énergies, mais ne compense pas la diminution introduite par (FC-V) pour les OM les plus proches du noyau $\{K; L\}$. Par contre, nous pouvons voir une amélioration des résultats dans les variants de (FC) par compensation d'erreur. Des résultats corrects et stables sur la série sont obtenues par la combinaison (FC-ND-f).

Quoi qu'il en soit, ces variants sont toujours moins bons que le schéma classique CVS. Et les coûts computationnels ne seraient avantageux que pour de larges zones de (FC).

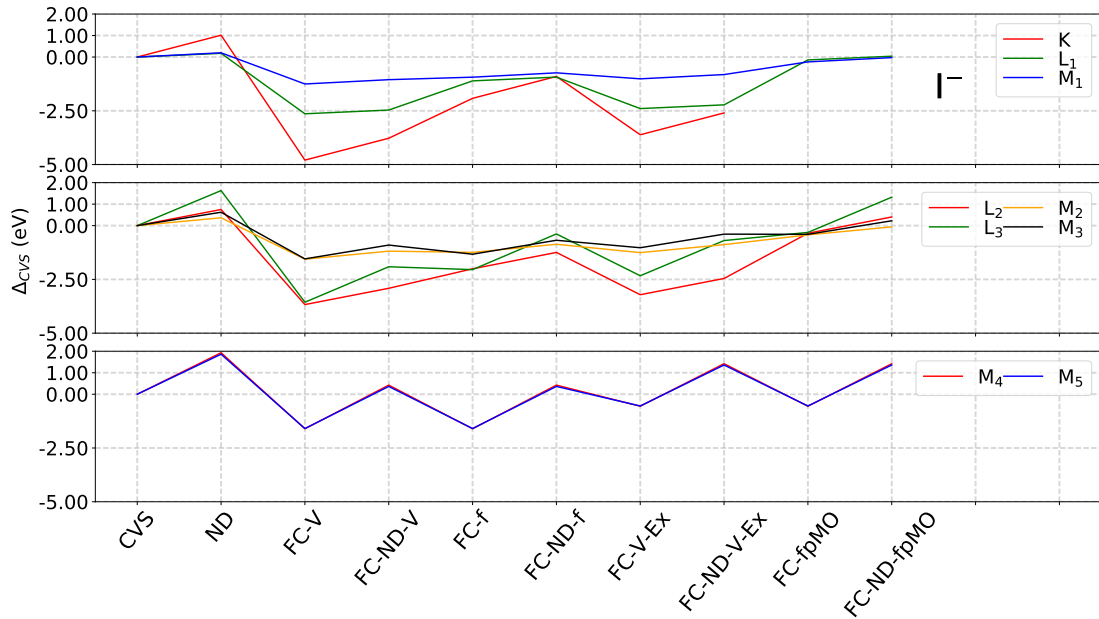


FIGURE 3.4 – Effet des différentes options appliquées à CVS pour Γ . Les valeurs en (eV) sont relatives à CVS. Les calculs sont obtenus avec l’hamiltonien ${}^2\text{DC}^M$ et dyall-acv3 ζ .

3.3 Influence de l’hamiltonien

En plus de la définition des projecteurs, et donc de la flexibilité accordée à la fonction d’onde pour décrire les états de cœur, l’hamiltonien joue un rôle très important dans la qualité des énergies calculées. Dans cette partie, nous étudions l’influence de l’hamiltonien et de l’ajout du terme de GAUNT (fig.3.5). Nous travaillons avec six hamiltoniens, deux à 4-composantes (${}^4\text{DC}$, ${}^4\text{DC}^{PF}$), quatre à 2-composantes (${}^2\text{DC}^M$, X2C-AMFI) et (${}^2\text{DCG}^M$, X2C-G-AMFI), ces 2 derniers comprenant le terme de GAUNT⁶. Les IP s’étalant sur plus de 30 keV, nous présentons, comme dans l’article, les différences entre les hamiltoniens X et l’hamiltonien DIRAC-COULOMB à 4-composantes ${}^4\text{DC}$ telles que : $\Delta E_m(X) = E_m(X) - E_m({}^4\text{DC})$.

L’hamiltonien ${}^4\text{DC}$ est considéré comme la référence car il n’y a pas d’approxi-

6. Nous avons conservé la notation de l’article pour les hamiltoniens mais X2C est écrit ${}^2\text{DC}$ dans le premier chapitre de cette thèse.

mation (hormis l'omission des intégrales (SS|SS)). ${}^4\text{DC}^{PF}$ présente la particularité de voir les énergies négatives supprimées dès le calcul sur l'hamiltonien de DIRAC *bare-nucleus one-electron*, ce qui le rapproche d'un calcul 2-composantes, sans pour autant introduire l'erreur de *picture change* pour la répulsion électronique comme le font les méthodes à 2-composantes⁷. Les hamiltoniens à 2-composantes sont décrits au chapitre 1 (cf.1.6.4).

La première chose est de donner une estimation de l'erreur du *picture-change*, i.e. estimer l'erreur due à la non transformation des intégrales 2-électrons. Dans le tableau (tab.3.2), nous regardons l'erreur de corrélation entre les méthodes ${}^4\text{DC}$, ${}^2\text{DC}^M$ et ${}^2\text{DCG}^M$. Deux éléments nous donnent des informations sur cette erreur, le premier vient de l'avantage d'avoir travaillé avec *MP2*. Avec cette technique les orbitales canoniques et les énergies associées sont les mêmes pour ${}^2\text{DC}^M$ et ${}^4\text{DC}$, ainsi toutes les erreurs sont dues aux erreurs de *picture-change*, ici $-0,53$ eV. Comme *CCSD* et *MP2* suivent les mêmes tendances, nous pouvons estimer que les erreurs pour *CCSD* sont du même ordre de grandeur.

Le deuxième argument vient en comparant ${}^4\text{DC}$ et ${}^4\text{DC}^{PF}$; comme dans ces calculs les erreurs de *picture change* sont éliminées et que les énergies de corrélations sont presque les mêmes $< 3 \times 10^{-4}$ eV, la différence $\Delta E_c(\text{CCSD})$ est due au *picture-change*. Nous pouvons considérer que ${}^2\text{DC}^M$ est une méthode fiable présentant un coût computationnel avantageux.

Nous avons également indiqué dans le tableau (tab.3.2) les valeurs pour ${}^2\text{DCG}^M$. Les différences étant sensiblement identiques, nous considérons ici que les erreurs sont du même ordre, en prenant garde évidemment au fait que nous n'avons pas les données auxquelles nous référer pour ${}^4\text{DCG}$ (Ajout du terme de Gaunt).

Nous pouvons enfin comparer X2C-AMFI et X2C-G-AMFI avec ${}^2\text{DCG}^M$; même si les courbes se suivent, la différence pour la couche K est proche de 50 eV.

L'hamiltonien le plus précis, au-delà de ${}^4\text{DCG}$ qui nous est rapidement inaccessible avec *Dirac*, devient alors ${}^2\text{DCG}^M$ puisque nous avons estimé le *picture change* comme faible et pris en compte le terme de GAUNT. Les résultats sont clairs puisqu'à 81.8 eV de la valeur obtenue par ${}^4\text{DC}$.

7. Projection dans le sens de ce que propose Furry [65] (option .VEXTPJ sous *Dirac*).

	${}^4\text{DC}^{PF}$	${}^2\text{DC}^M$	${}^2\text{DCG}^M$
$E_c(\text{MP2})$	-48,54	-49,07	-49,05
ΔE_c	$< 3 \times 10^{-4}$	-0,53	-0,51
$E_c(\text{CCSD})$	-46,13	-46,65	-46,64
ΔE_c	$< 3 \times 10^{-4}$	-0,52	-0,51

TABLEAU 3.2 – Γ^- Énergies de corrélation pour (MP2) et (CCSD) et différence ΔE_c par rapport à ${}^4\text{DC}$ (en eV).

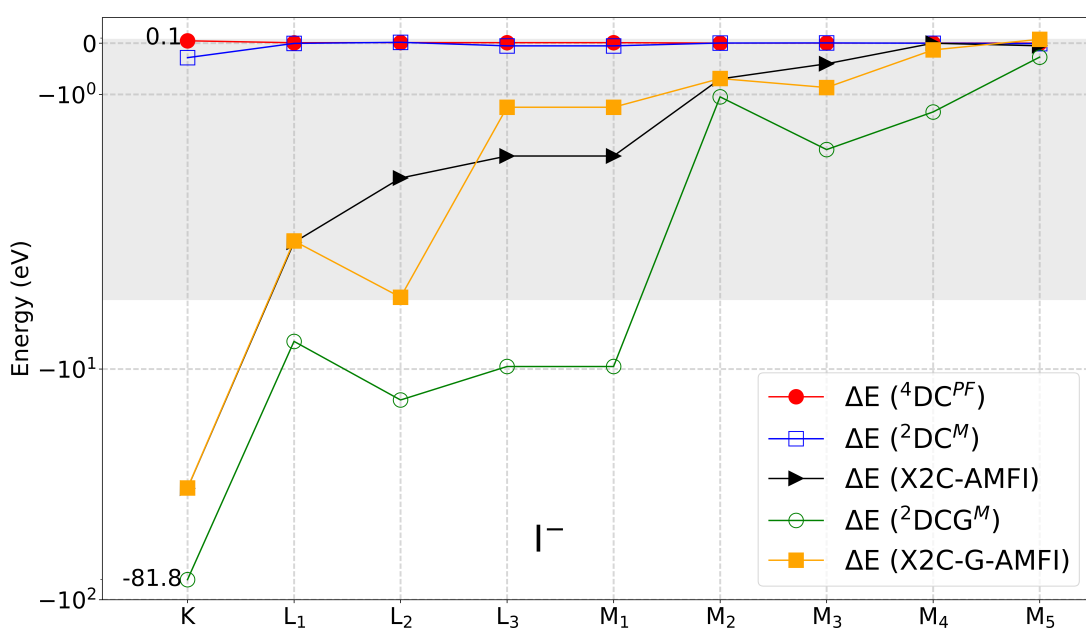


FIGURE 3.5 – Influence de l'hamiltonien sur les (IP) de cœur de l'ion iodure Γ^- . Les énergies (en eV) sont les différences, pour chaque *edge* m , entre l'hamiltonien X et ${}^4\text{DC}$ telles que : $\Delta E_m(X) = E_m(X) - E_m({}^4\text{DC})$. L'échelle des ordonnées est logarithmique, excepté pour la zone grisée où l'échelle est linéaire.

Il est important de s'intéresser ensuite à l'inclusion des intégrales de type $(SS|SS)$ pour l'ionisation de la 1s avec Γ^- ; l'énergie est alors plus basse de 2,58 eV. Relativement à l'énergie de l'ionisation de 33 209,16 eV cela semble négligeable. Mais ces intégrales sont importantes pour atteindre la précision expérimentale et nous avons pu voir que plus l'élément est lourd et plus l'influence de ces

intégrales se propage sur les couches L et M (cf. Annexe C).

Une dernière remarque concerne l'hamiltonien de BREIT et la *Quantum Electrodynamics (QED)*. Avec Dirac il n'était pas possible, jusqu'alors, de calculer la correction QED [192]; nous avons donc utilisé, pour l'article, les travaux de Southworth *et al.* [185] et de Koziol *et al.* [107]. À partir des travaux de ces derniers, il est possible de calculer pour la 1s de l'iode une énergie de $\varepsilon_{\text{QED}} = 36,17 \text{ eV}$ et un total de $\varepsilon_{\text{BREIT+QED}} = 104,88 \text{ eV}$ en ajoutant l'hamiltonien de BREIT, ce qui est loin d'être négligeable.

3.4 Conclusion

La conclusion reste celle de l'article, nous en rappelons les idées principales. Premièrement, *Core-Valence Separation (CVS)* est une très bonne approximation pour étudier ionisations, ou excitations, de cœur, comme nous pouvons le voir avec la comparaison des valeurs obtenues par la diagonalisation complète de la matrice *EOM-CC*. Nous pouvons dire que CVS fonctionne comme un filtre, en retirant les contributions (2h/1p) concernant la valence et ayant des énergies dans la gamme de celles des Simple-ionisations de cœur. Dans les différentes options pour CVS, celle qui fonctionne le mieux est celle pour laquelle nous avons utilisé simultanément *Frozen-Core*, en suivant la limite CVS, et retiré les déterminants double-excitations contenant uniquement les spineurs des électrons de cœur ND. Deuxièmement, pour les hamiltoniens, ${}^2\text{DC}^{\text{M}}$ donne des résultats corrects. Les intégrales de type (SS|SS) doivent être incluses à partir de la cinquième ligne de la classification périodique des éléments. Enfin, les effets qu'impliquent la QED et l'hamiltonien de BREIT ne sont pas négligeables à ce niveau, même si le terme de GAUNT, utilisé seul, permet déjà de retrouver une partie importante des effets relativistes d'ordre supérieur à l'effet SSO.

Méthodes EOM-CC approximées : Partitioned-EOM, EOM - Many-Body-Perturbation(2) et P-EOM-MBPT(2)

Les objectifs de ce travail sont d'étudier le comportement d'approches approximées de la matrice EOM-CCSD *Partitioned - Equation of Motion Coupled-Cluster (P-EOMCC)* et *Equation of Motion Coupled-Cluster - Many-Body Perturbation 2d order (EOMCC-MBPT(2))* pour obtenir les propriétés de molécules contenant des éléments lourds, qui présentent des intérêts dans la chimie atmosphérique [63, 61] ou dans le domaine nucléaire. Ces méthodes ont été implémentées, avec symétrie, dans module RELCCSD de *Dirac*; c'est le travail que nous présentons ici. Elles ont également été implémentées sans symétrie dans le module ExaCorr. Ce travail s'inscrit dans la continuité de l'article de Shee *et al.* [179].

L'atout que nous avons ici est la possibilité de pouvoir travailler avec la valence comme avec le cœur, mais également de pouvoir faire varier l'hamiltonien relativiste.

Les systèmes étudiés pour les *IP* de valence sont des composés halogénés I_3^- , XO^- avec $X \in \{Cl; Br; I; At\}$. Nous avons également travaillé avec le tennes-

sine¹ Ts (élément super-lourd Z=117). Pour les *IP* de cœur, nous avons étudié {HBr; HCl}, la taille de ces systèmes nous permettant de réaliser la diagonalisation complète de la matrice *EOM-CC*. Nous avons pu comparer ces résultats à ceux obtenus par la méthode *CVS* (cf.3.1). Enfin, *EE* et *EA* sont également décrits, validant nos équations et cette méthode.

4.1 Les différentes approximations

L'obtention des valeurs propres de la matrice *EOM*, de coût en $\mathcal{O}(N^6)$ pour *EE*, peut faire l'objet d'approximation. Comme dans le chapitre 1 (cf.1.5.1), nous nommerons chaque terme de la matrice *EOM* : « \mathcal{H} », i.e. *Similarity-Transformed Normal-Ordered*.

La première idée est *Partitioned - Equation of Motion Coupled-Cluster (P-EOMCC)*. Elle rend le bloc \mathcal{H}_{DD} diagonal en développant les termes à l'ordre⁽⁰⁾. Cette idée vient des travaux de Lödwin [125].

Reprenons l'équation (eq.1.63) :

$$\begin{pmatrix} \mathcal{H}_{SS} & \mathcal{H}_{SD} \\ \mathcal{H}_{DS} & \bar{\mathcal{H}}_{DD} \end{pmatrix} \begin{pmatrix} R_{k_a} \\ R_{k_b} \end{pmatrix} = \omega_k \begin{pmatrix} R_{k_a} \\ R_{k_b} \end{pmatrix} \quad (4.1.1)$$

développons :

$$\begin{cases} \mathcal{H}_{SS} \cdot R_{k_a} + \mathcal{H}_{SD} \cdot R_{k_b} = \omega_k R_{k_a} \\ \mathcal{H}_{DS} \cdot R_{k_a} + \bar{\mathcal{H}}_{DD} \cdot R_{k_b} = \omega_k R_{k_b} \end{cases} \quad (4.1.2)$$

Nous pouvons alors définir un Hamiltonien effectif :

$$\begin{aligned} \mathcal{H}_{SS}(\omega_k)_{\text{eff}} &= \bar{\mathcal{H}}_{SS} + \mathcal{H}_{SD} (\omega_k \cdot \mathbb{1} - \bar{\mathcal{H}}_{DD})^{-1} \bar{\mathcal{H}}_{DS} \\ \mathcal{H}_{SS\text{eff}} R_{k_a} &= \omega_k R_{k_a} \end{aligned} \quad (4.1.3)$$

En faisant un développement perturbatif à l'ordre 0 [125, 67], le bloc \mathcal{H}_{DD} est remplacé par $\mathcal{H}_{DD}^{(0)}$ qui est diagonal pour les orbitales (semi-)canoniques HARTREE-FOCK [76].

1. Tennessee.

La matrice *P-EOMCC* devient alors :

$$\begin{pmatrix} \mathcal{H}_{SS} & \mathcal{H}_{SD} \\ \mathcal{H}_{DS} & \mathcal{H}_{DD}^{(0)} \end{pmatrix} \quad (4.1.4)$$

Ensuite, en suivant les travaux de Stanton *et al.* [188], si les termes sont développés à l'ordre ⁽²⁾ en utilisant la théorie *Many-Body Perturbation Theory 2d order (MBPT(2))*, nous arrivons à *Equation of Motion Coupled-Cluster - Many-Body Perturbation 2d order (EOMCC-MBPT(2))* :

$$\begin{pmatrix} \mathcal{H}_{SS}^{(2)} & \mathcal{H}_{SD}^{(2)} \\ \mathcal{H}_{DS}^{(2)} & \mathcal{H}_{DD}^{(2)} \end{pmatrix} \quad (4.1.5)$$

Les amplitudes ne sont pas calculées par la méthode *CCSD* mais par un calcul *MP2*, ce qui génère un gain computationnel important :

$$t_i^{a(1)} = \frac{f_i^a}{\epsilon_i - \epsilon_a} \quad (4.1.6)$$

$$t_{ij}^{ab(1)} = \frac{\langle ab || ij \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \quad (4.1.7)$$

Il faut noter que pour un calcul *HF*, les t_1 sont donc nuls avec cette approximation.

Enfin, nous pouvons combiner les deux approximations précédentes dans la matrice qui devient alors *Partitioned - Equation of Motion Coupled-Cluster - Many-Body Perturbation 2d order (P-EOMCC-MBPT(2))* [76] :

$$\begin{pmatrix} \mathcal{H}_{SS}^{(2)} & \mathcal{H}_{SD}^{(2)} \\ \mathcal{H}_{DS}^{(2)} & \mathcal{H}_{DD}^{(0)} \end{pmatrix} \quad (4.1.8)$$

Pour toutes ces méthodes, les équations sont données en Annexe (cf.A.2). Il convient de noter les travaux de Goings *et al.* [70] sur les *EE* de valence de molécules organiques, comme la Pyridine, Dutta *et al.* [54] pour leur travail sur *EA* également pour des molécules organiques et l'adénine. Nous définirons par

une « erreur de relaxation » les erreurs introduites par l'utilisation de *Partitioned* comme le font Gwaltney *et al.* [76] dans leur étude sur les excitations de la valence; et nous définirons par « une erreur de corrélation » les erreurs introduites par l'utilisation de *Many-Body Perturbation Theory 2d order (MBPT(2))*. La disparition des amplitudes t_1 pose effectivement un problème de relaxation des orbitales mais ne voulant pas employer le même terme, nous utiliserons la remarque de Goings *et al.* [70, p 4] dans leur étude sur la différence entre P-EOM-MBPT(2) et LR-CC2 :

The \hat{T}_1 operator has long been understood as accounting for orbital response to electron correlation and the \hat{T}_2 operators are understood as accounting for the dynamic (pairwise) electron correlation.

4.2 Détails computationnels et présentation des résultats

Les géométries des molécules sont décrites dans l'article de Shee *et al.* [179] pour XO^- et I_3^- , et dans l'article de Visscher *et al.* [203] pour HX, rappelées en Annexe B.2.

Pour les hamiltoniens, nous avons utilisé Dirac-Coulomb et Dirac-Coulomb-Gaunt, respectivement ${}^2\text{DC}^{\text{M}}$ et ${}^2\text{DCG}^{\text{M}}$ (en ajoutant le terme de Gaunt cf.1.6.3). Lorsque les calculs ont été faits avec les deux hamiltoniens, pour la valence notamment avec EE et EA, seuls les résultats avec GAUNT sont présentés, les différences étant de l'ordre du meV. Les intégrales (SS|SS) sont approximées par un modèle de charges ponctuelles [201]. Une distribution gaussienne est utilisée pour la charge nucléaire [202].

Quand la totalité des spineurs sera corrélée lors des calculs, nous l'indiquerons par la notation $[\infty E_{\text{h}}]$. Dans le cas contraire, la notation : [limite inférieure; limite supérieure] E_{h} indiquera les bornes retenues pour la transformation des OA en OM qui serviront dans les calculs.

Pour les tableaux, le terme 'T' signifie, de manière générique : « transitions », les énergies sont en eV. Nous avons utilisé %SI comme indiqué au chapitre 1 (cf.1.5.4), $\%SI_k=100\left(r_i\right)_k^2$, il en sera de même avec %SA avec r^a et %SE avec r_i^a .

Enfin, pour faciliter la lecture, nous omettrons l'acronyme « EOM - CCSD ». Les différentes parties traitant respectivement des *IP*, *EA* et *EE* dans le cadre *EOM-CCSD*, *IP* (respectivement *EE*, *EA*) désignera toujours le calcul sans approximation (hormis *CVS* pour *IP* de cœur). Pour les méthodes approximées, IP_{App} (EE_{App} , EA_{App}), nous avons les trois approximations décrites précédemment : P-*IP* (P-*EE*; P-*EA*) pour le *Partitioned*, MBPT(2)-*IP* (MBPT(2)-*EE*, MBPT(2)-*EA*) pour l'approximation au second ordre et P-MBPT(2)-*IP* (P-MBPT(2)-*EE*, P-MBPT(2)-*EA*) pour la double approximation. Pour les tableaux, l'écriture se simplifiera encore en P-(*IP*; *EA*; *EE*);, PT(2) et P-PT(2).

Les *Bases* sont dyall-av3 ζ [55] pour les halogènes et les aug-cc-pVTZ [100] de Dunning pour l'oxygène, le carbone et l'hydrogène, toutes sont non-contractées.

Pour les *IP* (1h/0p). Dans un premier temps pour la valence, les quatre premières solutions sont calculées, les systèmes étudiés sont I_3^- [-3; 12] E_h et XO^- avec $X \in \{Cl; Br; I; At; Ts\}$ [∞E_h]. Puis pour les *Potential energy curves (PEC)*, XO^- avec $X \in \{Cl; Br; I; At\}$ [-10; 100] E_h , les courbes sont obtenues par un ajustement polynomial des énergies absolues en fonction de la distance de séparation des 2 noyaux X-O; les distances varient entre 1,4 Å et 2,4 Å. Pour les *IP* de cœur, les systèmes étudiés sont HCl et HBr, [∞E_h].

Pour *EA* (0h/1p), nous avons travaillé avec les dihalométhanés CH_2I_2 et CH_2IBr , pour lesquels nous présentons les six premières valeurs. Les électrons corrélés sont dans les limites [-3; 6] E_h pour CH_2I_2 et [-4; 6] E_h pour CH_2IBr .

Pour *EE* (1h/1p), nous présentons ici les résultats obtenus pour les systèmes I_3^- et CH_2I_2 . Les électrons corrélés sont dans les limites [-3; 12] E_h pour I_3^- et [-3; 6] E_h pour CH_2I_2 .

4.3 Potentiels d'ionisation

4.3.1 Triiodure I_3^-

Le premier système que nous étudions est I_3^- . Cet ion hypervalent est intéressant par sa stabilité et par les multiples voies qu'il offre dans sa photodissociation. Il a été étudié théoriquement [72, 205, 179] et expérimentalement [213, 34]. Les résultats sont donnés dans le tableau (tab.4.1).

T	Δ_{IP}					Δ_{exp}			
	IP	P-IP	PT(2)	P-PT(2)	exp [34]	IP	P-IP	PT(2)	P-PT(2)
$IP_1^{1/2g}$	4,46	-0,05	0,11	0,06	4,53	-0,06	-0,12	0,04	0,00
$IP_2^{3/2g}$	5,00	-0,14	0,01	-0,10	4,93	0,07	-0,07	0,08	-0,03
$IP_3^{1/2u}$	4,91	-0,10	0,04	-0,03	4,87	0,04	-0,06	0,08	0,00
$IP_4^{3/2u}$	4,28	-0,10	0,04	-0,04	4,25	0,03	-0,07	0,07	-0,01

TABLEAU 4.1 – I_3^- - Potentiel d'ionisation (eV). Δ_{IP} et Δ_{exp} sont respectivement les différences par rapport à IP et par rapport aux valeurs expérimentales [34]. Pour tous les cas %SI est dans l'intervalle [91%; 93%].

Dans un premier temps, comme il en est discuté dans les articles de référence, nous pouvons voir que sans approximation, $EOM-IP$ donne de bons résultats par rapport à l'expérience.

Dans un deuxième temps nous pouvons voir que les valeurs données par les méthodes approximées IP_{App} sont correctes par rapport aux calculs IP . Le comportement global est que P-IP sous-estime les valeurs, MBPT(2)-IP les surestime et P-MBPT(2)-IP est compris entre les deux; il compense les erreurs. Tous les %SI sont supérieurs à 90%; ce sont des ionisations simples et les méthodes approximées reproduisent bien ce comportement. Le vecteur propre est dominé par une ionisation venant d'une seule OM (1h/0p) et la faible différence entre ces pourcentages ne permet pas d'expliquer les différences observées.

4.3.2 Les ions hypohalogénites XO^- , $X \in \{Cl; Br; I; At; Ts\}$

Nous avons réalisé le même travail, pour les IP de valence de XO^- , $X \in \{Cl; Br; I; At; Ts\}$. Ces composés présentent l'halogène à l'état d'oxydation (+I). Les trois premiers composés {Cl; Br; I} ont un rôle dans la dégradation de l'ozone [29] et l'astate présente des intérêts en médecine nucléaire [118], le tennesse complète la série.

Nous allons regarder les quatre premiers IP ($X^{1/2}$, $X^{3/2}$, $A^{1/2}$ et $A^{3/2}$) des XO^- par $EOM-CCSD$ pour en réalité obtenir l'état fondamental et le premier état excité des radicaux correspondants. Les deux états sont des doublets et présentent chacun un *splitting*. Ce *splitting* sera d'autant plus important que nous descendons dans la classification périodique des éléments. Shee *et al.* [179] ont

en plus montré que l'état « X » le plus bas en énergie est localisé sur l'halogène, contrairement à l'état « A » qui est porté par l'oxygène; cela aura forcément une conséquence sur nos approximations.

Dans un premier temps, les quatre premiers IP ont été calculés, nommés : $X^{1/2}$, $X^{3/2}$, $A^{1/2}$ et $A^{3/2}$. Nous avons ensuite fait la différence entre les approximations IP_{App} et IP.

Nous avons fait la même chose pour les cinq composés; nous obtenons alors le graphique de la figure (fig.4.1) où les différences sont classées selon les 4 états.

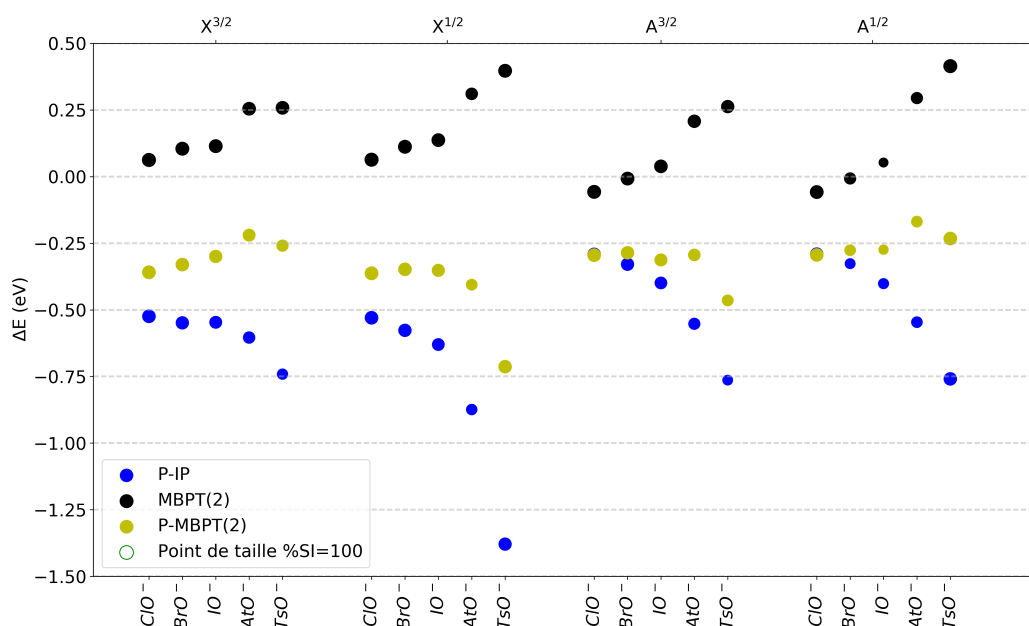


FIGURE 4.1 – XO^- $X \in \{Cl; Br; I; At; Ts\}$. ΔE Différence entre les méthodes approximations IP_{App} (P-IP, MBPT(2) et P-MBPT(2)) et EOMCC-IP en (eV) pour les quatre premières ionisations de valence : $X^{1/2}$, $X^{3/2}$, $A^{1/2}$ et $A^{3/2}$. Les points de ce graphique ont une taille proportionnelle à leurs %SI.

De façon générale, l'évolution de la différence entre les méthodes approximations et IP est la même pour les quatre premières solutions au fur et à mesure que nous traversons les états excités du plus bas au plus haut : MBPT(2)-IP est croissante, P-IP décroissante et P-MBPT(2)-IP reste relativement stable sur la série XO^- .

Commençons par MBPT(2)-IP; une discontinuité claire apparaît dans l'erreur entre l'iode et l'astate, mais cette méthode montre les plus faibles écarts avec IP.

En ce qui concerne le chlore, pour les états $X : \{X^{1/2}; X^{3/2}\}$, les valeurs sont légèrement surestimées (0,06 eV) et pareillement sous-estimées (-0,06 eV) pour les états $A : \{A^{1/2}; A^{3/2}\}$. Évidemment, le couplage spin-orbite est très faible mais l'erreur semble indépendante.

En ce qui concerne le brome et l'iode, pour les solutions X l'erreur est proche de 0,11 eV, pour les A elles sont inférieures à 0,05 eV. Là encore, le couplage spin-orbite est calculé avec la même précision. Il n'y a, en effet, qu'une faible différence entre les solutions $X^{1/2}$ et $X^{3/2}$ respectivement de 0,01 eV et 0,02 eV pour le brome et l'iode. Et cette différence, entre les solutions $A^{1/2}$ et $A^{3/2}$, est inférieure à 1 meV pour le brome et de 0,02 eV pour l'iode.

Ainsi, pour ces 3 composés, du chlore à l'iode, la moyenne des erreurs absolues $\bar{\delta}$ pour les solutions X est de $\bar{\delta} = 0,1$ eV avec un écart-type de $\sigma = 0,03$ eV; pour les solutions A , ces valeurs sont de $\bar{\delta} = 0,04$ eV et $\sigma = 0,02$ eV. Les faibles valeurs des écart-types semblent indiquer que l'erreur est systématique.

Pour l'astate et le tennessine, l'erreur devient plus importante mais reste inférieure à 0,5 eV. Le couplage spin-orbite n'est plus calculé avec la même qualité que ce soit pour les états X ou A .

P-IP présente les moins bons résultats : les erreurs sont toutes négatives; cette approximation sous-estime donc les IP . Cependant, il est visible que pour des solutions différenciées uniquement par le couplage *spin-orbitale* (les couples : $\{X^{3/2}; X^{1/2}\}$ et $\{A^{3/2}; A^{1/2}\}$), du chlore à l'iode, l'erreur n'est qu'un décalage, les valeurs ne sont pas très dispersées. Pour l'astate, les valeurs sont toujours relativement proches de 0,5 eV, mais le couplage spin-orbite pour les états X n'est pas calculé avec la même qualité; les erreurs sont respectivement de -0,60 eV et -0,87 eV. Pour le tennessine (SuperHeavy Element), l'erreur devient beaucoup plus importante et atteint même -1,38 eV pour $X^{1/2}$.

Enfin, la méthode P-MBPT(2)-IP compense les erreurs des approximations dont elle est issue. Du chlore à l'iode, les différences sont incluses dans les limites : $[-0.36; -0.27]$ eV, dans ce cas-là $\bar{\delta} = 0,34$ eV et $\sigma = 0,02$ eV pour les solutions X et $\bar{\delta} = 0,29$ eV, $\sigma = 0,01$ eV pour les solutions A . Les écarts-types nous montrent que l'erreur est quasiment constante.

En raison de la structure même de l'approximation P-IP, c'est le manque de relaxation (disparition des déterminants (2h/1p) non-diagonaux dans le

bloc \mathcal{H}_{DD}) qui est la cause des différences observées. Pour MBPT(2)-IP, c'est le manque de corrélation qui se produit. Il est à noter que ces deux énergies varient dans des sens opposés lorsque nous descendons la colonne des halogènes. Ainsi, l'approximation P-MBPT(2)-IP profite de cette compensation d'erreur; paradoxalement la méthode la plus approchée devient, dans ce cas, la plus stable sur cette série.

La voie qui s'ouvre pour expliquer l'erreur croissante est que la symétrie X est portée par une fonction d'onde principalement halogénée [179] qui nécessite plus de relaxation. Les %SI évoluant pour chaque composé et pour chaque état, il n'est pas possible d'en tirer une conclusion générale. Pour P-IP, ils sont généralement inférieurs à quelques pourcents de ceux obtenus avec IP mais cette différence ne semble pas être une caractéristique de l'erreur. Par exemple, l'état $A^{1/2}$ de l'iode a un %SI à 49% et montre un caractère où plusieurs déterminants (1h/0p) interviennent dans le vecteur propre. Les méthodes approximées présentent des %SI respectivement de {55%; 43%, 47%} pour P-IP, MBPT(2)-IP et P-MBPT(2)-IP, mais la comparaison avec l'état $A^{3/2}$, pour lequel les erreurs sont équivalentes, montre des %SI compris entre 74% et 86%. Ainsi, la lecture seule de la valeur %SI ne permet pas de caractériser l'erreur.

Courbes d'énergie potentielle pour les ions hypohalogénites

Nous avons poursuivi en étudiant les *Potential Energy Curves* (PEC) pour les quatre composés XO^- $X \in \{Cl; Br; I; At\}$. Comme précédemment, IP nous donne accès à l'état fondamental du radical et à son premier état excité. Les électrons corrélés sont dans l'intervalle $[-10; 100]E_h$.

Nous donnons un exemple pour AtO^- sur la figure (fig.4.2) où le 0 en énergie est donné par rapport aux minimums des courbes $X^{3/2}$, les autres courbes étant données en annexe B.

Les solutions sont clairement séparées en deux groupes, l'état fondamental et le premier état excité. Les solutions $\{X^{3/2}; X^{1/2}\}$ appartiennent au premier groupe centré sur l'halogène pour lequel l'erreur est liée au *splitting*; en effet l'état $X^{3/2}$ est correctement représenté mais pour l'état $X^{1/2}$ les courbes sont dispersées en énergie. Le deuxième groupe, $\{A^{3/2}; A^{1/2}\}$, a un comportement

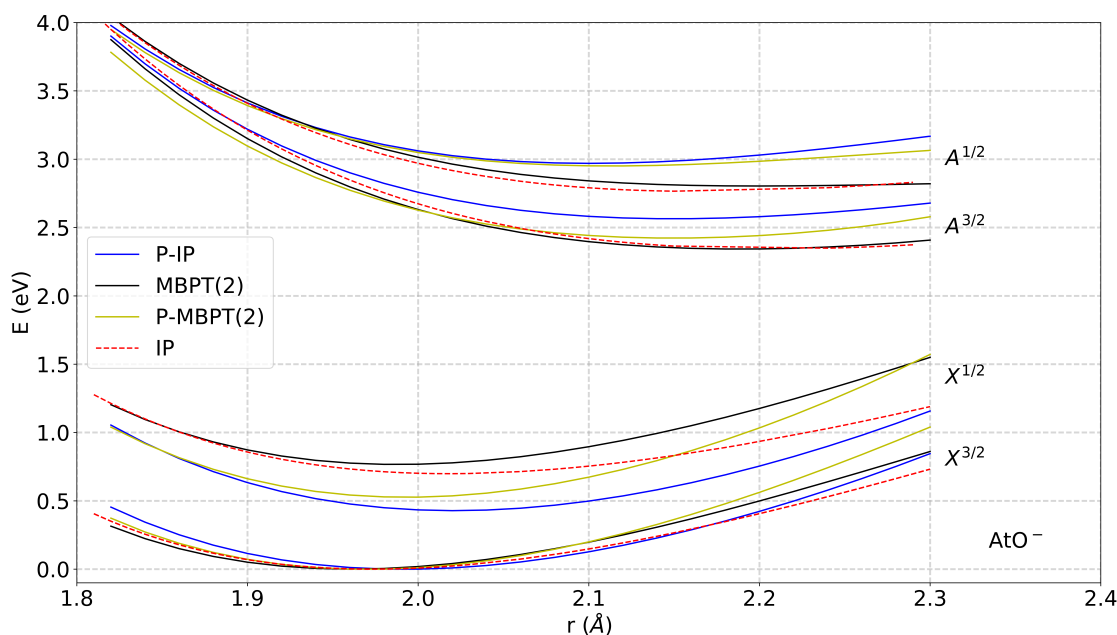


FIGURE 4.2 – Courbes d'énergie potentielle pour AtO^- dans la limite $[1,8; 2,3]\text{\AA}$. Par la méthode IP, P-IP, MBPT(2)-IP et P-MBPT(2)-IP, et cela pour les 4 états ($X^{1/2}$, $X^{3/2}$, $A^{1/2}$ et $A^{3/2}$). Hamiltonien ${}^2\text{DCG}^M$ (dyall-av3 ζ) pour At et aug-cc-pV3 ζ pour O. Les électrons corrélés sont dans la limite $[-10; 100]E_h$.

différent ; les solutions A semblent toutes provenir d'un même point et c'est la différence de courbure qui montre la différence en énergie. MBPT(2)-IP (en noir) présente la meilleure correspondance avec IP, P-IP (en bleu) s'en écarte le plus vite. Le comportement de P-MBPT(2)-IP (en jaune), entre les deux courbes, est encore clairement dans une position de compensation de l'erreur.

Ce graphique est intéressant mais il place les courbes relativement à 0 eV par rapport au minimum de $X^{3/2}$. Nous avons voulu comparer les positions des minimums (pour les énergies et pour les distances d_{X-O}) (les minimums en position et en énergie sont donnés en annexe).

Dans cette optique, nous présentons deux graphiques représentant les erreurs en énergie ΔE en fonction des erreurs en position Δr . Dans la première figure (fig.4.3), les données sont classées selon les IP_{App} ; dans la seconde figure (fig.4.4), le classement se fait selon le composé XO^- .

Prenons la première figure (fig.4.3), nous remarquons immédiatement que

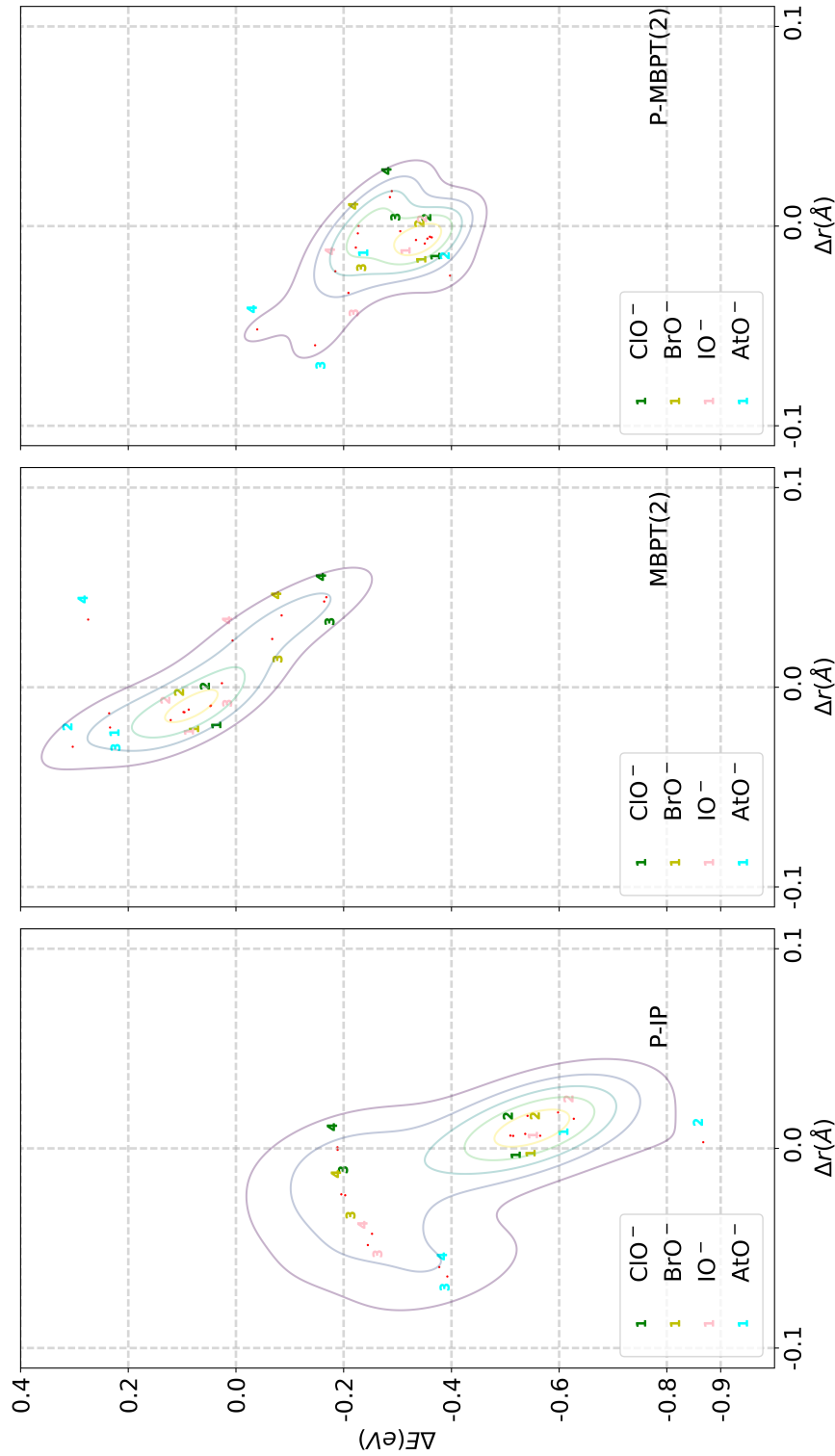


FIGURE 4.3 – Différences en énergies et en positions pour les $PEC - \Delta E$ (eV) en fonction de Δr (Å) pour les 3 IP_{App} et les 4 composés, classées par chaque IP_{App} . Pour des raisons de lisibilité : les nombres de 1 à 4 représentent les solutions $X^{3/2}$, $X^{1/2}$, $A^{3/2}$ et $A^{1/2}$. Les 5 différentes isodensités ont été calculées avec l'algorithme *Gaussian kde* [171], respectivement dans les limites $[-0,1; 0,1]$ (Å), $[-1,0; 0,4]$ (eV) pour Δr et ΔE .

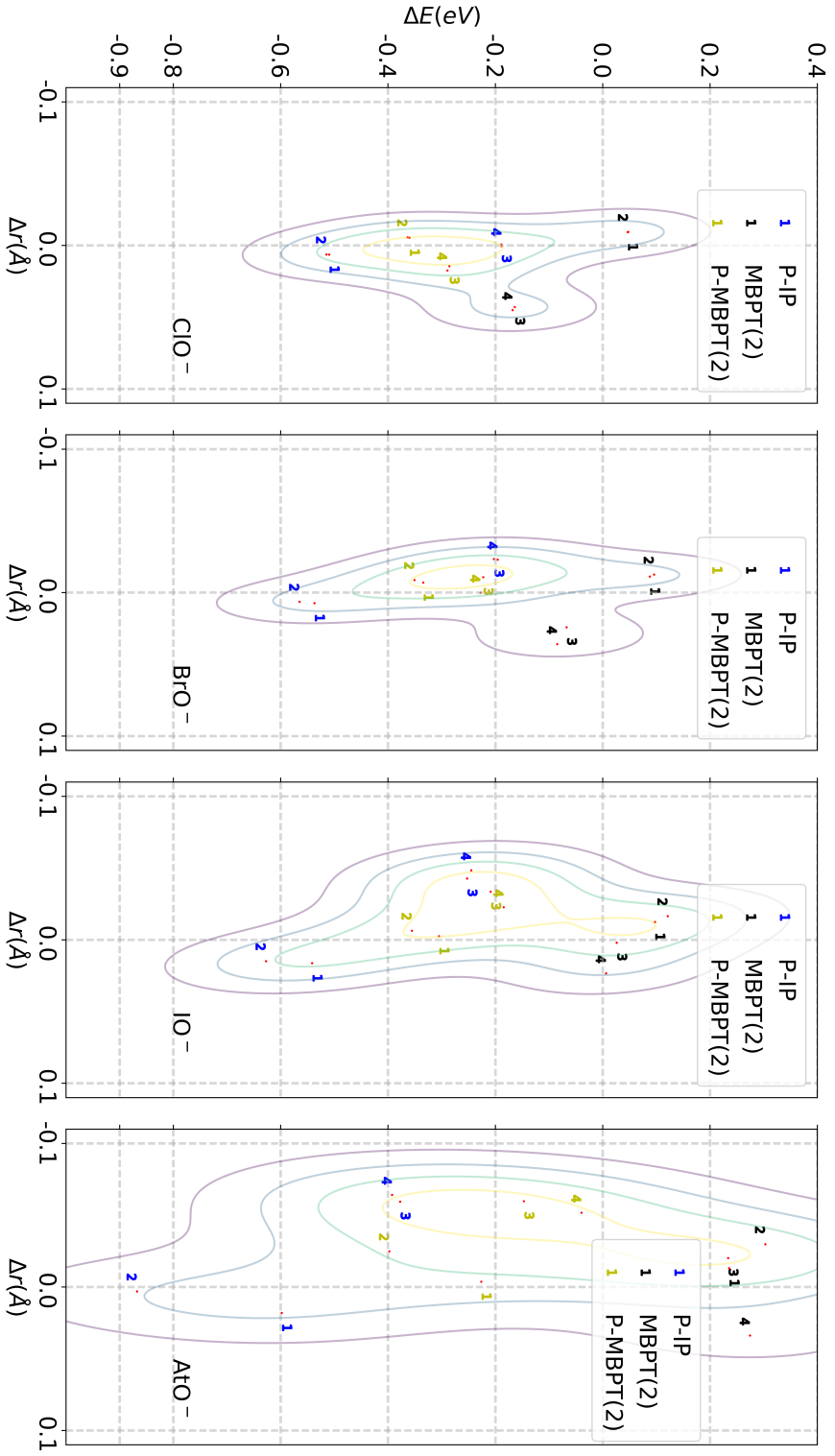


Figure 4.4 – Différences en énergies et en positions pour les PEC - ΔE (eV) en fonction de Δr (Å) pour les 3 IP $_{App}$ et les 4 composés, classés par composé. Pour des raisons de lisibilité : les nombres de 1 à 4 représentent les solutions $X^{3/2}$, $X^{1/2}$, $A^{3/2}$ et $A^{1/2}$. Les 5 différentes isodensités ont été calculées avec l'algorithme *Gaussian kde* [171], respectivement dans les limites $[-0,1; 0,1]$ (Å), $[-1,0; 0,4]$ (eV) pour Δr et ΔE .

P-IP est plus dispersé, que MBPT(2)-IP est relativement correct et centré sur (0; 0) et que P-MBPT(2)-IP compense les erreurs. Ensuite, nous observons que les valeurs pour AtO^- sont toujours les plus excentrées. Sur ces trois graphiques, nous remarquons deux zones. La première zone, pour les états $X^{3/2}$ et $X^{1/2}$, présente des valeurs concentrées. Si pour P-IP les énergies sont proches de $-0,5$ eV, les erreurs Δ_r sont faibles. Les valeurs sont correctes pour MBPT(2)-IP. La deuxième zone concerne les états $A^{3/2}$ et $A^{1/2}$: pour P-IP, Δ_E est constante jusqu'à l'iode ; pour MBPT(2)-IP l'évolution est plus complexe ; l'erreur est importante pour le chlore et cette erreur tend à diminuer jusqu'à l'iode. Pour P-MBPT(2)-IP, nous voyons encore la compensation d'erreur (en excluant les solutions $A^{3/2}$ et $A^{1/2}$ de l'astate) : toutes les valeurs sont comprises dans une zone de $\pm 0,1$ eV et $\pm 0,025$ Å.

Pour le deuxième graphique (fig.4.4), les informations sont évidemment identiques, mais à la lecture nous voyons facilement l'augmentation de l'erreur pour P-IP et la solution $X^{1/2}$ au fur et à mesure que nous descendons la colonne des halogènes. Nous remarquons que les erreurs pour $A^{3/2}$ et $A^{1/2}$ sont indépendantes de la magnitude du couplage spin-orbite.

Ainsi, les valeurs des méthodes approximées ont des écarts importants pour les atomes les plus lourds par rapport à *EOM* mais montrent de bons résultats pour des atomes plus légers. Encore une fois, P-MBPT(2)-IP a un comportement systématique intéressant, il ne montre que peu de dispersion en comparaison de P-IP.

Enfin, la différence entre MBPT(2)-IP et IP peut s'expliquer par une description plus pauvre de l'état fondamental par cette approximation [76]. Une correction peut être apportée en comparant les énergies de corrélation CCSD par rapport à MP2. Nous avons corrigé nos résultats par la différence entre ces deux énergies (par rapport à leurs minimums respectifs) le long de la courbe de dissociation ; cela semble parfois améliorer les résultats mais sans qu'une conclusion générale n'apparaisse. Cette méthode de compensation peut être intéressante dans l'explication mais peu avantageuse pour obtenir les énergies, le but de MBPT(2)-IP et P-MBPT(2)-IP étant d'éviter de faire des calculs CCSD. Cette correction aurait également pu être apportée dans l'étude précédente sur XO^- mais un changement de la ligne de base ne conduit qu'à un changement de la moyenne et non de l'écart-type (comme la distance est fixée entre les IP).

Énergie de transition

Nous avons discuté des énergies de façon isolée, mais c'est également important de voir comment les méthodes représentent l'espace entre les états électroniques. Si ce comportement est correct, il est alors possible de discuter de la transition électronique entre l'état fondamental et l'état excité. Bien évidemment, nous avons vu sur les *PEC* (Annexe B) que les erreurs ne sont pas constantes et que nous avons donc plutôt intérêt à travailler autour de la distance d'équilibre de l'état fondamental.

Concernant les énergies de transition, l'état $X^{3/2}$, le plus bas en énergie, est pris comme référence. La différence entre ces différents niveaux $\Delta_{IP_{App}} - \Delta_{IP}$ est étudiée; ce sont ces différences d'énergies qui sont représentées graphiquement dans la figure (fig.4.5).

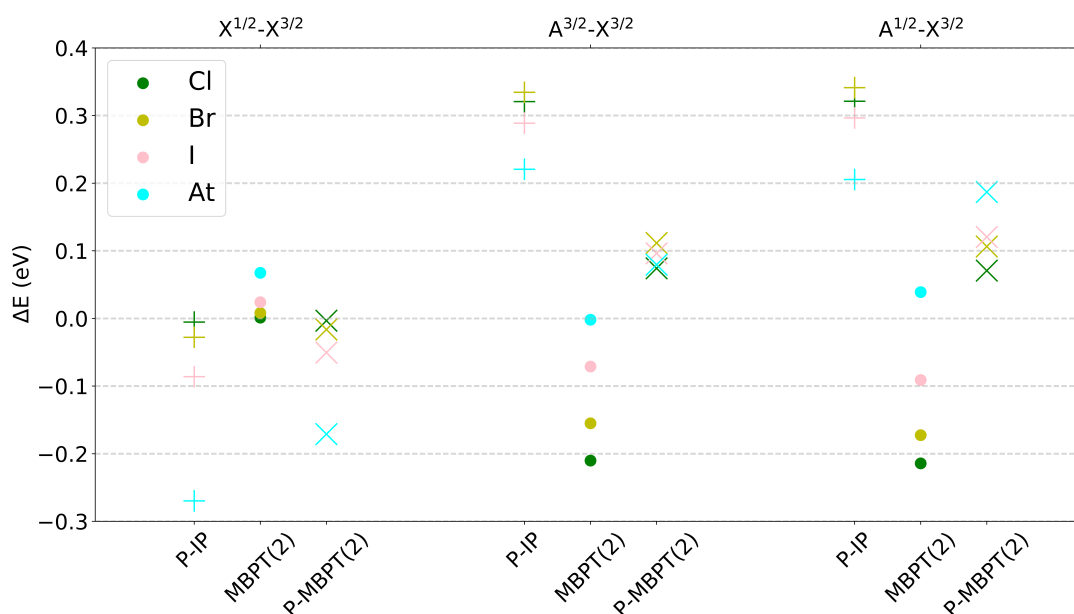


FIGURE 4.5 – Erreur dans les énergies $\Delta E = \Delta_{IP_{App}} - \Delta_{IP}$ (eV) pour les transitions : $X^{1/2}-X^{3/2}$, $A^{3/2}-X^{3/2}$ et $A^{1/2}-X^{3/2}$, pour les différents XO^- et selon les différentes IP_{App} .

Pour la première transition ($X^{1/2}-X^{3/2}$), concernant le chlore et le brome, la *splitting* est faible; toutes les méthodes montrent sensiblement les mêmes

résultats. Pour les deux autres composés, l'erreur augmente pour P-IP, pour MBPT(2)-IP, les énergies sont correctes et restent inférieures à 0,1 eV. Pour P-MBPT(2)-IP, l'erreur reste faible jusqu'à l'iode pour lequel l'erreur est de $-0,05$ eV. Pour les deux transitions suivantes ($A^{3/2}-X^{3/2}$ et $A^{1/2}-X^{3/2}$), pour un même IP_{App} et le même composé, les erreurs sont sensiblement les mêmes. P-IP surestime les énergies, MBPT(2)-IP, au contraire, les sous-estime mais les résultats s'améliorent de ClO^- à AtO^- , la mauvaise représentation des minimums peut expliquer ce fait. P-MBPT(2)-IP compense les erreurs.

Il est donc possible d'utiliser les méthodes IP_{App} pour déterminer les énergies de transition entre les différents niveaux. Nous pouvons comparer ces énergies de transition par rapport aux valeurs expérimentales (tab. 4.2). Dans ce tableau est indiquée l'énergie de transition $\Delta_T = X^{1/2} - X^{3/2}$ pour les différentes approximations. À la lecture de ce tableau, les conclusions restent les mêmes, nous pouvons également remarquer que la méthode P-MBPT(2)-IP reste correcte jusqu'à l'iode avec une erreur de $\sim 0,04$ eV.

$\Delta_T X^{1/2}-X^{3/2}$	exp [69]	Δ_{IP}	Δ_{P-IP}	$\Delta_{MBPT(2)-IP}$	$\Delta_{P-MBPT(2)-IP}$
ClO^-	0,0397	0,0396	0,0343	0,0408	0,0363
BrO^-	0,1270	0,1256	0,0977	0,1336	0,1092
IO^-	0,2593	0,2717	0,1856	0,2956	0,2211

TABLEAU 4.2 – Énergies de transition Δ_E entre $X^{1/2}-X^{3/2}$ (en eV) pour ClO^- , BrO^- et IO^- , comparaison avec les données expérimentales exp [69].

Méthodes approximées pour les IP des électrons de cœur, utilisation de CVS.

Enfin, nous avons voulu étudier les IP des électrons de cœur en utilisant également CVS. Nous avons travaillé avec les composés HCl et HBr car une diagonalisation complète de la matrice $EOM-CC$ est possible.

Cela nous permet de vérifier la validité de nos méthodes par rapport au maximum de précision numérique que nous sommes capables d'atteindre. Cependant la diagonalisation n'étant possible que pour de « petits » systèmes, nous comparerons les énergies entre CVS- IP_{App} et CVS- IP^2 .

2. Puisque nous avons utilisé CVS pour les méthodes approximées, nous n'ajouterons pas ce

HCl - Le tableau (tab.4.3), représente les IP des orbitales moléculaires pour les couches {K; L} de HCl. Les symboles Δ_f et Δ_{IP} signifient respectivement les différences entre CVS-{IP} avec full-IP et CVS-IP_{App} avec CVS-IP.

T	full-IP E_f	CVS-IP E_{IP}	Δ_f	P-IP Δ_{IP}	PT(2) Δ_{IP}	P-PT(2) Δ_{IP}	exp
$1s_{\sigma_{1/2}}^{-1}$	2835,46	2835,76	0,30	-0,45	0,23	-0,20	
$2s_{\sigma_{1/2}}^{-1}$	280,25	280,27	0,02	-0,14	0,12	-0,01	
$2p_{\sigma_{1/2}}^{-1}$	209,66	209,59	-0,07	-0,06	0,13	0,09	208,96 [†] 208,7 ^{††}
$2p_{\pi_{1/2}}^{-1}$	208,02	207,93	-0,09	-0,15	0,13	0,09	207,1 ^{††}
$2p_{\pi_{3/2}}^{-1}$	207,94	207,84	-0,10	-0,06	0,13	0,08	207,32 [†]

TABLEAU 4.3 – HCl - Énergie E (en eV) - Δ_f la différence entre IP- full-IP - Δ_{IP} celle entre IP_{App}-IP (en eV) - Les %SI IP_{App} sont compris dans l'intervalle [87%; 88%]. [†] [57] (précision de $\pm 0,05$ eV), ^{††} [81] (précision de $\pm 0,5$ eV.)

D'abord, les résultats obtenus par la méthode CVS-IP sont satisfaisants par rapport à full-IP, comme nous avons pu le voir lors du chapitre précédent (cf.3.1). Dans le tableau, Δ_f est de 0,3 eV pour la couche L_1 et n'excède pas -0,1 eV pour la couche L_2 .

Ensuite, comparons les méthodes approximées avec IP : Δ_{IP} . Tout au long de la série des OM, les erreurs sont faibles; pour la couche L, elles ne sont jamais supérieures à 0,15 eV quelle que soit la méthode IP_{App} par rapport à IP. Les %SI, qui ne sont pas reportés ici, sont tous compris dans l'intervalle [87%; 88%] et chaque ionisation est portée par le même déterminant (1h/0p). Dans le cas de HCl, MBPT(2)-IP surestime les énergies, P-IP les sous-estime et P-MBPT(2)-IP est le plus proche de IP n'excédant jamais 0,09 eV. Lors de la comparaison des valeurs, il y a une compensation d'erreur pour P-MBPT(2)-IP.

Enfin, les valeurs sont également proches de celles obtenues expérimentalement par Hayes *et al.* [81] et Eland *et al.* [57]. Les différences $\Delta E = E(2p_{\sigma_{1/2}}) - E(2p_{\pi_{1/2}})$ et $\Delta E = E(2p_{\sigma_{1/2}}) - E(2p_{\pi_{3/2}})$ sont respectivement {1,60 eV; 1,64 eV}, pour full-IP, elles sont égales à {1,65 eV; 1,73 eV}. La faible dispersion des valeurs par les méthodes approximées que ce soit par la nature de CVS ou par la préfixe, pour ne pas alourdir le texte inutilement, par exemple P-IP signifiera ici CVS-P-IP.

nature des IP_{App} nous indique que ces méthodes permettent, pour des éléments légers comme le chlore, d'atteindre la précision expérimentale avec des écarts relatifs ne dépassant pas $0,1\sim 0,2$ eV.

HBr - Les résultats sont donnés dans le tableau (tab.4.4). La différence entre IP et full-IP est plus importante ici que dans le cas précédent, mais CVS-IP reste une méthode fiable. Pour les méthodes approximées, l'erreur absolue est également beaucoup plus grande, mais le comportement général de compensation d'erreur par P-MBPT(2)-IP est toujours visible. L'élément important à observer est que les méthodes approximées présentent des erreurs constantes; il n'y a qu'un décalage entre ces méthodes et IP : proche de $-1,15$ eV, de $0,23$ eV et de $-0,78$ eV pour P-IP, MBPT(2)-IP et P-MBPT(2)-IP respectivement.

Pour comparer avec les données expérimentales, nous pouvons regarder en particulier la différence entre les niveaux $3d_{\sigma_{1/2}}^{-1}$ et $3d_{\delta_{1/2}}^{-1}$ qui est de $1,1$ eV. C'est un peu plus faible pour full-IP et IP avec respectivement $1,04$ eV et $1,08$ eV. Pour les méthodes approximées, nous obtenons : $1,05$ eV pour P-MBPT(2)-IP ; $1,08$ eV pour MBPT(2)-IP et $1,06$ eV pour P-IP. La comparaison avec l'expérience est donc possible.

T	full-IP E_f	CVS-IP E_{IP}	Δ_f	P-IP Δ_{IP}	PT(2) Δ_{IP}	P-PT(2) Δ_{IP}	exp[141]
$3p_{\sigma_{1/2}}^{-1}$	200,77	200,92	0,16	-1,15	0,15	-0,85	
$3p_{\pi_{1/2}}^{-1}$	194,21	194,00	-0,21	-1,13	0,15	-0,83	
$3p_{\pi_{3/2}}^{-1}$	194,15	193,84	-0,31	-1,13	0,14	-0,84	
$3d_{\sigma_{1/2}}^{-1}$	78,03	78,13	0,10	-1,15	0,23	-0,78	78,5
$3d_{\delta_{1/2}}^{-1}$	77,00	77,05	0,06	-1,13	0,23	-0,77	77,4
$3d_{\pi_{3/2}}^{-1}$	77,84	77,90	0,06	-1,16	0,23	-0,80	78,3
$3d_{\delta_{3/2}}^{-1}$	77,01	76,93	-0,08	-1,13	0,23	-0,78	77,3
$3d_{\delta_{5/2}}^{-1}$	76,88	76,74	-0,13	-1,14	0,22	-0,79	77,1

TABLEAU 4.4 – HBr - Énergie E (en eV) et %SI pour $3p^{-1}$ et $3d^{-1}$ - Δ_f la différence entre IP- full-IP - Δ_{IP} celle entre IP_{App} -IP (en eV) - Les %SI IP_{App} sont compris dans l'intervalle [89%;91%]. Pour les valeurs expérimentales [141] (précision de $\pm 0,33$ eV).

Ainsi, pour la technique CVS associée à P-IP, MBPT(2)-IP et P-MBPT(2)-IP, nous pouvons conserver la même conclusion que pour les ions XO^- comme cela a été vu dans la partie précédente. La perte de précision par la méthode *Partitioned*

due à la perte de relaxation est compensée en partie par les erreurs de corrélation de la méthode MBPT(2)-IP.

4.4 Énergie d'excitation et affinité électronique

Les méthodes *EOM* approximées peuvent être aussi utilisées pour obtenir les énergies d'excitation (*EE*) et les affinités électroniques (*EA*). Pour explorer ces propriétés, nous avons réalisé des calculs sur CH_2I_2 et CH_2IBr pour *EA* et sur I_3^- et CH_2I_2 pour *EE*.

Affinité électronique CH_2I_2 - Les données sont reportées dans le tableau (tab.4.5). Nous ne reviendrons pas sur la discussion entre les données expérimentales de Modelli *et al.* [134] et le travail théorique réalisé par Shee *et al.* [179], nous nous contenterons de comparer les méthodes EA_{App} .

T	EA		P-EA		PT(2)		P-PT(2)	
	E	%SA	%SA	Δ_{EA}	%SA	Δ_{EA}	%SA	Δ_{EA}
EA ₁	-0,33	48	50	-0,06	48	-0,06	49	-0,10
EA ₂	0,41	57	60	0,01	58	-0,01	60	0,00
EA ₃	0,75	83	81	-0,02	82	-0,01	80	-0,02
EA ₄	0,86	98	98	0,01	98	-0,01	98	0,00
EA ₅	1,49	87	86	0,01	86	0,00	85	0,01
EA ₆	1,70	67	62	-0,01	65	-0,01	61	-0,01
exp[134]			Δ_{exp}					
DA ₂	0,46	-0,05		-0,05		-0,06		-0,05

TABLEAU 4.5 – CH_2I_2 - Affinité électronique en (eV) - Δ_{EA} différence entre EA_{App} -EA - Δ_{exp} la différence avec l'expérience [134].

Premièrement, nous remarquons que la première *EA* pour obtenir CH_2I_2^- est négative. Toutes les EA_{App} le sont également, même si les énergies correspondantes sont sous-estimées; -0,1 eV au pire pour P-MBPT(2)-EA. La deuxième information nous indique que, hormis EA₁, toutes les méthodes EA_{App} sont correctes avec des erreurs absolues toujours inférieures à 0,02 eV. La troisième information est que le %SA, que nous n'avons pas reporté ici, nous montre qu'hormis EA₄ tous les vecteurs propres sont multi-déterminantaux (0h/1p).

Ce caractère étant, nous remarquons que les %SA sont identiques d'une méthode à l'autre. Enfin, la méthode P-MBPT(2)-EA ne compense pas les erreurs comme nous l'avons observé avec IP. La méthode la plus stable de EA₂ à EA₆ est MBPT(2)-EA.

CH₂I₂ - Les résultats sont donnés sur le tableau (tab.4.6). Comme précédemment, hormis EA₁, les résultats sont très satisfaisants et inférieurs à 0,01 eV en valeur absolue. Les %SA sont faibles ; les vecteurs propres sont multi-déterminantaux, composés par des déterminants (0h/1p) mais équivalents d'une EA_{App} à l'autre. Les résultats obtenus pour E₁, inférieurs au cas CH₂I₂, s'expliquent par la substitution entre l'iode et le brome. Les écarts sont tellement faibles qu'il est difficile de classer qualitativement les méthodes les unes par rapport aux autres. Cependant, nous pouvons noter que MBPT(2)-EA présente uniquement des valeurs négatives, l'écart-type est de proche de 2 meV.

T	EA		P-EA		PT(2)		P-PT(2)	
	E	%SA	%SA	Δ_{EA}	%SA	Δ_{EA}	%SA	Δ_{EA}
EA ₁	-0,03	52	53	-0,04	52	-0,06	53	-0,08
EA ₂	0,45	66	70	0,00	67	-0,01	70	0,00
EA ₃	0,86	97	97	0,01	97	-0,01	97	0,00
EA ₄	0,99	87	85	0,00	86	-0,01	85	-0,01
EA ₅	1,62	70	69	0,01	68	-0,01	67	0,01
EA ₆	1,92	52	49	0,00	55	-0,01	48	0,00

TABLEAU 4.6 – CH₂I₂ - Affinité électronique (en eV) - Δ_{EA} différence entre EA_{App}-EA.

Énergies d'excitation I₃⁻ - Les données sont reportées sur le tableau (tab.4.7).

D'abord, nous pouvons voir que les EE_{App} fonctionnent bien. Les écarts-types sont faibles, ce qui signifie que les erreurs pour P-EE et P-MBPT(2)-EE ont des décalages ; ces méthodes surévaluent *les énergies d'excitation*. Comme dans le cas de EA, nous n'avons pas la compensation des erreurs. La méthode P-MBPT(2)-EE n'améliore pas résultats obtenus par P-EE (comme pour nos observations sur EA et contrairement à nos observations sur IP), les erreurs s'additionnent. Nous remarquons une inversion de symétrie (en caractère gras dans le tableau) pour P-EE avec les niveaux $\{1_g; 0_u^-; 1_u\}$, mais ils sont tellement proches en énergie qu'il n'est pas improbable que de telles inversions apparaissent.

Par contre, nous devons noter la présence de valeurs hors-normes pour les

T	EE		P-EE		PT(2)-EE		P-PT(2)-EE		CASPT2[72]	
	EE	%SE	%SE	Δ_{EE}	%SE	Δ_{EE}	%SE	Δ_{EE}	E	Δ_{EE}
2_g	2,25	43	45	0,27	43	0,00	45	0,27	2,24	-0,01
1_g	2,38	23	24	0,25	23	0.11	24	0.37	2,32	-0,06
0_u	2,38	41	42	0.27	41	0.00	42	0.27	2,47	0,09
1_u	2,38	46	47	0.25	46	0,11	47	0,37	2,47	0,09
0_g	2,84	21	22	0,27	22	0,00	22	0,28	2,76	-0,08
0_g^+	2,89	21	22	0,27	21	0,00	22	0,27	2,82	-0,07
1_g	3,07	41	43	0,28	41	0,01	43	0,28	2,85	-0,22
2_u	3,33	48	49	0,26	48	-0,01	49	0,24	3,10	-0,23
1_u	3,42	48	49	0,26	48	-0,01	49	0,25	3,11	-0,31
0_u^+	3,67	18	19	0,24	17	-0,01	19	0,23	3,52	-0,15
2_g	4,08	24	24	0,27	24	-0,01	24	0,26	3,98	-0,10
0_u	4,10	42	43	0,29	42	-0,01	43	0,28	3,79	-0,31
1_g	4,18	47	48	0,28	47	-0,01	49	0,27	4,06	-0,12
1_u	4,22	40	41	0,30	40	0,00	40	0,29	3,80	-0,42
0_u^+	4,49	15	16	0,24	15	0,00	17	0,24	4,51	0,02
0_g^-	4,69	20	21	0,30	20	-0,01	21	0,29	4,51	-0,18
0_g^+	4,70	20	21	0,30	20	-0,01	21	0,29	4,53	-0,17
1_g	4,90	40	36	0.30	40	0,01	26	0.33	4,60	-0,30
\bar{x}				0,27		0,01		0,28		-0,14
σ				0,02		0,04		0,04		0,14

TABLEAU 4.7 – I_3^- - Énergies d'excitation (eV)- Δ_{EE} différence entre EE_{App} -EE (en eV) - En gras les niveaux inversés (Voir le texte) - \bar{x} la moyenne et σ l'écart-type - données expérimentales pour (0_u^+) : 3,43 4,25 [34] 3,45 4,28 [213] - Données CASPT(2) [72].

premières 1_g et 1_u par les méthodes MBPT(2)-EE et P-MBPT(2)-EE. En effet, alors que MBPT(2)-EE présente une différence d'au maximum 0,01 eV en valeur absolue, l'erreur pour ce niveau est de 0,11 eV. Il en est de même pour P-MBPT(2)-EE qui passe d'une erreur de 0,27 eV à 0,37 eV. Celle-ci est la cause de l'inversion de symétrie pour ces méthodes et nous n'en avons pas encore élucidé la cause.

Ensuite, nous remarquons en général que les %SE sont faibles. D'après les vecteurs propres, qui ne sont pas reproduits ici, ces excitations sont principalement dominées par un caractère multi-déterminantal $2 * (1h/1p)$. Mais quoi qu'il en soit, les pourcentages sont sensiblement les mêmes pour tous les EE_{App} .

Enfin pour la comparaison avec, CASPT(2) [72], la différence avec EE n'est pas constante. De manière générale CASPT(2) est inférieure à EE. Il y a certainement un mélange d'état proche de l'équilibre. Il n'y a pas de corrélation visible EE_{App} et CASPT(2).

CH₂I₂ - Les données sont reportées sur la tableau (tab.4.8).

Pour ce système, les écarts-types sont encore une fois satisfaisants, de l'ordre

T	EE		P-EE		PT(2)-EE		P-PT(2)-EE		CASPT2	
	EE	%SE	%SE	Δ_{EE}	%SE	Δ_{EE}	%SE	Δ_{EE}	E	Δ_{EE}
a	3,60	20	21	0,29	20	-0,07	21	0,23	3,76	0,16
b	3,62	20	20	0,29	20	-0,07	20	0,23	3,78	0,16
a	3,62	20	21	0,29	20	-0,06	21	0,23	3,78	0,15
b	3,84	16	17	0,30	15	-0,07	16	0,23	4,03	0,19
a	3,87	19	19	0,29	19	-0,07	19	0,23	4,27	0,40
b	3,94	13	13	0,31	12	-0,08	13	0,24	4,27	0,33
a	3,99	16	17	0,30	16	-0,08	17	0,23	4,31	0,32
b	4,06	16	16	0,31	16	-0,07	16	0,24	4,38	0,32
b	4,22	18	19	0,30	18	-0,06	19	0,25	4,50	0,28
a	4,32	15	16	0,31	15	-0,07	16	0,24	4,60	0,28
b	4,35	15	15	0,31	15	-0,06	15	0,26	4,62	0,27
a	4,49	15	15	0,31	15	-0,07	16	0,25		
b	4,63	15	15	0,33	15	-0,07	15	0,27		
a	4,68	17	17	0,35	17	-0,06	18	0,30		
b	4,74	15	14	0,34	15	-0,07	15	0,28		
a	4,91	14	15	0,36	15	-0,06	15	0,30		
\bar{x}				0,31		-0,07		0,25		0,26
σ				0,02		0,01		0,02		0,08

TABLEAU 4.8 – CH_2I_2 - Énergies d'excitation (eV)- Δ_{EE} différence entre EE_{App} -EE (en eV) - \bar{x} la moyenne et σ l'écart-type - Données CASPT(2) [117]

de 0,2 eV ; les moyennes sont respectivement de 0,31 eV, -0,07 eV et 0,25 eV pour P-EE, MBPT(2)-EE et P-MBPT(2)-EE. Pour CH_2I_2 P-MBPT(2)-EE retrouve des valeurs comprises entre P-EE et MBPT(2)-EE. Nous remarquerons cette fois-ci la stabilité des différences avec CASPT(2) dont les données sont extraites du travail de Liu *et al.* [117]. Les %EE sont encore une fois faibles et dominés par des simples excitations.

4.5 Conclusion

Nous avons des implémentations fonctionnelles pour {p; MBPT; P-MBPT}-EOM-CCSD dans un cadre relativiste Dirac-Coulomb-Gaunt dans l'approximation *2-component molecular mean field*. Les erreurs sont assez systématiques de l'EOM-MBPT(2) par rapport à l'EOM, et entre *P-EOMCC-CCSD* et *P-EOMCC-MBPT(2)*. Toutes ces méthodes sont robustes et n'ont pas présenté de problème particulier de convergence.

Ces méthodes approchées (IP_{App}) fonctionnent bien dans les cas étudiés ici. Pour les potentiels d'ionisation de valence, *P-EOMCC* et *EOMCC-MBPT(2)*

introduisent des erreurs dont les effets sont opposés. La compensation d'erreur observée pour *P-EOMCC-MBPT(2)* devient intéressante en raison des faibles valeurs de dispersion pour l'ionisation des électrons de valence. En ce qui concerne les *Courbes d'énergie potentielle*, les IP_{App} peuvent être utilisées afin d'obtenir les distances d'équilibre, les minimums en énergies et les énergies de transition et ce avec les mêmes précautions que pour les autres études. Pour les *OM* de cœur, IP_{App} couplé à la méthode *Core-Valence Separation (CVS)*, les énergies obtenues sont correctes par rapport à *EOM-CC-IP* et correctes par rapport à l'expérience.

Pour *Partitioned-IP*, c'est évidemment le problème de la relaxation qui intervient, seules les composantes diagonales du bloc \mathcal{H}_{DD} de la matrice EOM-CCSD étant prises en compte dans cette approximation. Pour MBPT(2)-IP, la différence résulte de la description de l'énergie de corrélation. De manière générale, P-IP sous-estime les énergies d'ionisation alors que MBPT(2)-IP les surestime.

Les méthodes approximées, utilisées pour les énergies d'excitation *EE* et les affinités électroniques *EA*, donnent de très bons résultats pour les systèmes étudiés, ici en présence d'iode. Il faut noter quelques inversions pour des symétries dont les niveaux sont très proches en énergie. Même lorsque les différences sont relativement plus importantes, les écarts-types sont suffisamment faibles pour ne parler que d'un décalage.

La facilité d'emploi de ces méthodes, sans avoir à choisir d'espace actif notamment, les rend intéressantes pour l'étude des propriétés proches de l'équilibre.

EOM-CCSD se révèle une méthode précise mais coûteuse; *P-EOMCC* et *EOMCC-MBPT(2)* ont un coût moindre et peuvent facilement être utilisées comme techniques d'exploration. Il serait également intéressant de voir comment ces méthodes fonctionnent avec des techniques d'embedding comme cela a été réalisé dans l'équipe PCMT par Bouchafra *et al.* [23]. Pour les éléments lourds, un travail sur les *IP* de cœur de UO_2^{2+} est en cours. Tout semble indiquer des différences importantes de plusieurs eV mais la compensation d'erreur limite les problèmes et les énergies de transition sont inférieures à 0,5 eV.

4.6 Remarques sur les coûts computationnels

Pour terminer, nous pouvons faire quelques remarques sur les coûts computationnels. Cependant, il devient compliqué de se limiter au simple critère $\mathcal{O}(N^p)$ pour les méthodes approximées. En effet, 3 éléments sont à prendre en compte. Premièrement, pour plus de rigueur, il convient de développer $\mathcal{O}(N^p)$ en $(n_{occ}^m n_{vir}^n)^3$. Deuxièmement, il faut se montrer vigilant pour l'utilisation d'intermédiaires de calcul et troisièmement, il faut identifier les produits tensoriels inclus dans les boucles d'itérations.

Pour la notion d'intermédiaire, il faut rappeler que les techniques *CC* sont basées sur une succession de produits tensoriels. Mais, quand des opérations identiques sont réalisées de la même manière mais à plusieurs endroits dans les équations, comme le produit $A_{o_1 o_2}^{v_1 v_2} = t_{o_1}^{v_1} \otimes t_{o_2}^{v_2}$, par exemple, il convient alors de calculer l'intermédiaire A une seule fois et de l'utiliser à convenance.

Pour *EOM-CC*, le calcul le plus cher est en $(n_{occ}^2 n_{vir}^4)$, mais également en $(n_{occ}^3 n_{vir}^3)$ et $(n_{occ}^4 n_{vir}^2)$ impliquant que cette méthode soit : $\mathcal{O}(N^6)$. Pour *EOM-CC-IP*, l'usage est de donner $\mathcal{O}(N^5)$. En effet, le bloc H_{DD} présente un calcul $(n_{occ}^2 n_{vir}^3)$. Pour *EOM-CC-EA*, également en $\mathcal{O}(N^5)$, le bloc H_{DD} présente un calcul en $(n_{occ}^1 n_{vir}^4)$; déterminer les *EA* sera toujours plus cher que de calculer *IP*⁴. Par contre, pour *IP* et *EA*, l'utilisation d'intermédiaires fait apparaître certains calculs en $\mathcal{O}(N^6)$ mais ils n'entrent pas dans les boucles d'itérations.

Pour *EOMCC-MBPT(2)-EE*, le coût reste en $\mathcal{O}(N^6)$, mais il n'est pas nécessaire de calculer les amplitudes *CCSD* : t_1 et t_2 qui sont en $\mathcal{O}(N^6)$. Pour *IP* et *EA*, les coûts sont respectivement en $(n_{occ}^2 n_{vir}^3)$ et $(n_{occ}^1 n_{vir}^4)$. De plus, pour les systèmes à couches fermées, les $t_1^{(1)}$ sont nuls.

Pour le *Partitioned - Equation of Motion Coupled-Cluster (P-EOMCC)*, nous sommes en $\mathcal{O}(N^5)$ [76]; mais en regardant en détail, pour *EE* nous avons $(n_{occ}^2 n_{vir}^3)$. Pour les *IP* et *EA*, le coût est en $\mathcal{O}(N^4)$ pour le bloc H_{DD} et c'est le bloc H_{DS} qui prend alors le coût du calcul en $\mathcal{O}(N^5)$. Pour *Dirac*, l'utilisation d'intermédiaires (toujours en $\mathcal{O}(N^6)$) fait passer le calcul itératif en $\mathcal{O}(N^4)$.

3. n_{occ} et n_{vir} représentant respectivement le nombre d'OM occupées et virtuelles.

4. Ce qui montre l'importance de décomposer $\mathcal{O}(N^5)$.

Enfin, pour *P-EOMCC-MBPT(2)-EE*, le coût est effectivement en $\mathcal{O}(N^5)$, le bloc H_{DS} a le même coût, l'avantage résultant du fait d'éviter les calculs $\mathcal{O}(N^6)$ des amplitudes *CCSD*.

Vers la mise en œuvre de méthodes *Coupled-Cluster* relativistes approximées et de propriétés moléculaires pour architectures hybrides et massivement parallèles

Ce chapitre présentera des exemples de travaux réalisés avec le nouveau module ExaCorr qui intègre sous *Dirac* la nouvelle bibliothèque de gestion des calculs parallèles ExaTensor. Ainsi, pour les calculs *CC*, *Dirac* possède désormais, en plus de l'implémentation « historique » (RELCCSD), une implémentation visant des architectures *HPC* de nouvelle génération¹, construite pour les calculs hautement parallélisés.

Après une brève introduction de cette bibliothèque de calcul, nous montrons deux exemples extraits de l'article de Pototschnig *et al.* : « Implementation of relativistic coupled cluster theory for massively parallel GPU-accelerated computing architectures » [150], réalisés sur Summit, le supercalculateur d'Oak Ridge [127, 142, 191]. Cet article présente notamment les performances pour *CC*, nous le reproduisons en Annexe D. Nous poursuivrons par la présentation

1. Avec ou sans accélération par des *GPU*.

de CC2 (CC2). Enfin, nous finirons par une introduction de la méthode *Linear Response Coupled-Cluster (LR-CC)* appliquée à la polarisabilité (α).

5.1 Une nouvelle bibliothèque : ExaTensor

ExaTensor est une bibliothèque de gestion des produits tensoriels pour des applications nécessitant des calculs hautement parallélisés. Elle a été écrite par Lyakh [129]; elle peut travailler sur des clusters hétérogènes comme sur des supercalculateurs *HPC*. Elle dépend de bibliothèques mathématiques comme BLAS et LAPACK et des bibliothèques usuelles de distribution de données comme OpenMP, MPI (pour le travail sur plusieurs nœuds) et CUDA (pour le travail avec des GPU). Cette bibliothèque ExaTensor gère la parallélisation en s'appuyant sur une autre bibliothèque TAL-SH qui effectue le travail sur chaque nœud².

ExaTensor n'est pas propre à *Dirac* et peut fonctionner indépendamment pour tout type de calculs tensoriels (les produits, les additions, la transposition et les multiplications par des scalaires) et cela sur tous les types de tenseurs (sparse/sparse block).

Cette bibliothèque est basée sur un DVSP (domain-specific virtual processing), qui est un processeur virtuel intermédiaire pouvant interpréter tout type d'expression tensorielle algébrique spécifique à un domaine de recherche.

TAL-SH

TAL-SH [130] : *Tensor Algebra Library for SHared memory computers* est un API prévu pour des opérations sur des CPU multi-cœurs à mémoire partagée et un accélérateur optionnel GPU. Les calculs sont opérés de façon asynchrone pour des systèmes matériels hétérogènes. Le partage du tenseur se fait par un processus : OpenMP multithreading.

2. Les codes sources sont accessibles - ExaTensor possède une licence : *GNU Lesser General Public License*.

ExaTensor

ExaTensor, également asynchrone, se distingue donc de TAL-SH par les processus MPI entre les nœuds. Le but est de devenir « target-agnostic »³. Il y a deux types de communication MPI. Premièrement un processus aller/retour qui transfère les métadonnées (les tâches); deuxièmement un processus aller (1GB) pour les données, la demande se faisant par la vérification d'un registre de requêtes.

5.2 Exemple d'expression informatique pour les calculs tensoriels

La portabilité a été grandement améliorée, notamment d'un point de vue de la programmation, Lyakh [129] écrit à ce propos :

Such direct translation of quantum many-body equations into a human-readable code drastically accelerates the implementation of new coupled cluster methods in the ExaCorr module.

L'écriture des produits tensoriels est extrêmement simple⁴. Reprenons le produit tensoriel : $A_{ij}^{ab} = \frac{1}{4} V_{ef}^{ab} t_{ij}^{ef}$, ExaTensor est quasi-wysiwyg, pourvu que les tenseurs $\{A; V; T\}$ soient déclarés au préalable; cette contraction s'écrit comme elle se lirait⁵ :

```
ierr=talsh_tensor_contract("A(a,b,i,j)+=V(a,b,e,f)*T(e,f,i,j)",A,V,T, &
    & scale=ONE_QUARTER)
```

Cela facilite grandement la lecture en cas de débogage. Le retour de « ierr », qui est un nombre entier, donne également des informations sur les problèmes éventuellement rencontrés lors de l'exécution du produit tensoriel.

A titre de comparaison, voici une contraction dans RELCCSD mais incluant la symétrie :

3. Pouvant s'exécuter sur n'importe quelle architecture : « inter-plateformes ».
 4. Même si d'autres bibliothèques, comme Tensor.jl [32] ou Jutho/TensorOperations.jl [98] sous Julia, permettent une implémentation relativement simple.
 5. Ici, c'est un exemple avec TAL-SH, mais la construction est identique avec ExaTensor.

```
call contraction_444(("p1", "p2", "o3", "o4"/), ("o3", "o4", "o1", "o2"/), &
& ("p1", "p2", "o1", "o2"/), sigma2, 1.0d0, 1.0d0, nrep, LeftTensor=r2, &
& RightTensor=B%W_ijmn, irrep_left=state_sym, irrep_right=1)
```

5.3 Résultats sur Summit

L'article de Pototschnig *et al.* [150] présente différents travaux réalisés avec ExaCorr sur Summit. Le but est de voir le comportement de cette bibliothèque pour des systèmes de taille importante. Nous présenterons ici deux travaux caractéristiques de cet article.

5.3.1 La symétrie

ExaCorr, par choix d'implémentation, ne prend pas en compte la symétrie spatiale des systèmes et les calculs sont obligatoirement en $(n_{occ}^m n_{vir}^n)$. Même s'il y a un gain à la prendre en compte, comme le montrent « historiquement » Stanton *et al.* [189], l'objectif de cette implémentation est son application à de plus gros systèmes pour lesquels il y a généralement absence de symétrie.

Le but n'est pas de reprendre l'étude sur les dimères or-argon de l'article, mais de regarder les temps nécessaires aux calculs *CC* selon le module de *Dirac* utilisé. Nous allons comparer RELCCSD sans et avec symétrie (R19/ R19s)⁶ avec ExaCorr. Rappelons que RELCCSD ne montre pas une haute performance parallèle au-delà de quelques dizaines de processus MPI.

Nous reportons une partie du tableau 5 de l'article sur le tableau (tab.5.1).

Les résultats sont ceux auxquels nous pouvions nous attendre. La symétrie incluse dans le code historique RELCCSD (R19s) est très intéressante pour de « petits » systèmes, mais quand la taille augmente, c'est l'efficacité dans la parallélisation qui prend toute son importance. Il y a « une inflexion » selon la taille des systèmes, où le gain apporté par ExaCorr compense la non prise en compte de la symétrie par rapport à RELCCSD qui, lui, l'intègre.

6. Release 2019 : R19 sans symétrie, R19s : avec symétrie.

	Système	N	R19	R19s	ExaCorr
CCSD	AuAr^+	4	327	7	84
	AuAr^+	8	232	6	71
	AuAr_2^+	8	1129	7	73
	AuAr_3^+	8	2512	849	183
	AuAr_3^+	20	1585	586	143

TABLEAU 5.1 – Durée des calculs en secondes pour la nouvelle implémentation et RELCCSD release 2019 (l’implémentation de référence), avec symétrie (R19s) et sans (R19). Tous les calculs ont été réalisés avec la *Base* dyall-v2z. Il y a 243/42 orbitales atomiques, 32/14 spineurs occupés et 104/32 virtuels pour les atomes or/argon. N est le nombre de nœuds utilisés sur Summit.

5.3.2 Complexe d’uranyl : le trinitrate d’uranyl $[\text{UO}_2(\text{NO}_3)_3]^-$

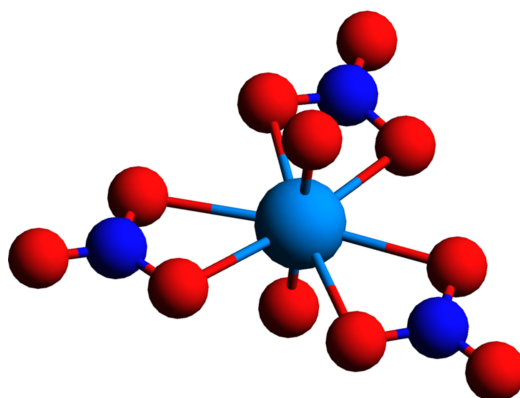


FIGURE 5.1 – Structure du complexe d’uranyl $[\text{UO}_2(\text{NO}_3)_3]^-$ déduite de la structure rayon-X du cristal $\text{RbUO}_2(\text{NO}_3)_3$ [14].

L’uranyl UO_2^{2+} , de par son intérêt et sa dangerosité, fait partie des composés étudiés depuis de nombreuses années [46, 149, 160]. Pour $[\text{UO}_2(\text{NO}_3)_3]^-$, de par sa taille les calculs *CC*, avec des bases décontractées et avec autant de virtuelles, sont inaccessibles sur des clusters « usuels »⁷. Ce composé est donc apparu comme un candidat idéal pour être traité sur Summit. Cependant, le but ici

7. Même en prenant en compte la symétrie.

n'est pas de reprendre l'étude menée dans l'article ni sur la composante q_{zz} de l'*Electric Field Gradient* (EFG) pour $[\text{UO}_2(\text{NO}_3)_3]^-$ ni sur la convergence du critère t_1 , mais de montrer la capacité de calcul sur un *HPC* comme Summit.

Les données obtenues sont reproduites en partie dans le tableau (tab.5.2).

Calcul	O	x_O	V	x_V	N	C	NormC	Durée	
								Totale	T+ Λ
dyall-v2z									
1	106	0,53	534	0,42	80	9	1,00	3h34	2h55
2	156	0,77	534	0,42	80	20	2,17	5h33	5h06
3	156	0,77	818	0,64	180	109	5,30	11h21	10h43
4	202	1	534	0,42	200	33	1,45	8h02	7h29
5	202	1	680	0,53	1025	87	0,75	12h28	11h58
6	202	1	818	0,64	1025	183	1,56	18h23	17h49
7	202	1	896	0,68	1025	263	2,25	23h04	22h21
dyall-v3z									
8	106	0,53	694	0,33	240	26	0,95	6h00	4h08
9	156	0,72	694	0,33	240	57	2,06	10h24	8h02
10	156	0,72	944	0,46	1025	193	1,65	17h48	15h44

TABLEAU 5.2 – Durée des calculs pour obtenir les amplitudes $T + \Lambda$ et Totale (incluant la transformation des intégrales) pour le complexe trinitrate d'uranyl $[\text{UO}_2(\text{NO}_3)_3]^-$ - X2C-AMFI. O : le nombre de spineurs occupés corrélés; V : le nombre de spineurs virtuels corrélés; N : le nombre de nœuds en calcul parallèle; C : le coût estimé pour le calcul (facteur de $1,0 \times 10^{14}$); NormC : valeur de C normalisée (Voir texte).

De prime abord, il faut noter la possibilité offerte de travailler avec plus de 1000 nœuds de calcul ($N=1025$), ce qui nous permet de pouvoir traiter un système tel que celui-ci, c'est-à-dire pouvoir corrélérer $O = 202$ OM occupées et $V = 896$ virtuelles. En considérant $C = O^2V^4$, nous devons donc réaliser $C=263 \times 10^{14}$ opérations, mais c'est un minimum. La seule prérogative étant que la durée du calcul soit inférieure à 24 heures.

Nous travaillons avec la Durée(T+ Λ) qui est le temps des itérations *CCSD* et la Durée(Totale) qui inclut en plus le temps de transformation des intégrales.

Nous voyons que la multiplication par 30 des coûts de calcul entre le premier et le septième calcul est compensé par une multiplication par 13 du nombre de nœuds.

Il n'est pas possible de relier simplement les durées de calcul avec leurs coûts ; tous les nœuds n'ont pas « travaillé » en même temps au calcul mais certains ont servi de mémoire pour pallier l'augmentation de la taille des tenseurs (avec l'augmentation du nombre de virtuelles). Il faut donc trouver un équilibre entre la quantité de calculs, le nombre de noeuds utilisés et la mémoire nécessaire.

Pour quantifier une première idée du problème de la mémoire, nous travaillons avec une « métrique », NormC, qui est obtenue par le rapport du coût C sur le nombre de nœuds N, normalisé par $C[1]/N[1]$; ainsi : $\text{NormC}[n] = C[n]/(N[n] * C[1]/N[1])$, où n est le numéro du calcul. Le fait que NormC soit compris dans l'intervalle $\sim [1; 2]$ à l'exception du calcul numéro 3 nous indique que tous les calculs réalisés valent de une à deux fois le premier, la seule différence se jouant sur la mémoire. Cependant que Summit, a dans ce cas, un scaling non-linéaire, nous pouvons dire « grossièrement » que le code peut gérer des charges de travail élevées, pour lesquelles beaucoup de mémoire est nécessaire, à peu près de la même manière que pour des charges de travail beaucoup plus petites (notion de weak scaling) sur la durée de 24 heures.

5.4 CC2

Au-delà de réimplémenter les équations *CCSD*, *CCSD(T)* du code « historique »⁸, nous avons profité de ce module pour intégrer (en plus des méthodes discutées dans le chapitre précédent (cf.4)) la méthode « second-order approximate coupled cluster singles and doubles » *CC2* [37], qui est très utilisée par la communauté pour traiter des systèmes de plus grande taille. Et cela a été, en même temps, très utile pour nous familiariser avec ces bibliothèques.

Dans cette méthode, une approximation est faite sur l'ordre de développement des amplitudes \hat{T}_2 de *CCSD*. La méthode *CC2* (*CC2*) est énoncée, par Christiansen *et al.* [37] :

8. RELCCSD dans *Dirac*.

The singles equations are retained in their original form, but the doubles equations are approximated to be correct through first order only, with the singles treated as zeroth-order parameters.

Ainsi, en partant de l'hamiltonien *Normally-Ordered* et en utilisant la notation historique $\hat{H}_N = e^{-\hat{T}_1} H_N e^{\hat{T}_1}$, nous obtenons :

$$\hat{H}_N = \hat{F}_N + \hat{V}_N \quad (5.4.1a)$$

$$\langle \Phi_i^a | \hat{H}_N + [\hat{H}_N, T_2] | \text{HF} \rangle = 0 \quad (5.4.1b)$$

$$\langle \Phi_{ij}^{ab} | [F_N, T_2] + \hat{H}_N | \text{HF} \rangle = 0 \quad (5.4.1c)$$

Les équations pour résoudre les amplitudes \hat{T}_1 sont inchangées (eq.5.4.1b), seuls les termes pour le calcul des amplitudes \hat{T}_2 (eq.5.4.1c) sont modifiés. Le problème réside toujours dans la transformation de ces expressions en expressions programmables. Nous donnons les équations en Annexe 1 (cf. A.1).

5.4.1 Coût computationnel

De nouveau, nous revenons à la notion de coût computationnel. Précédemment, nous avons évoqué cette notion de façon générale sous la forme $\mathcal{O}(N^N)$ en parlant des différentes méthodes de chimie quantique. Pour les méthodes approximées, nous avons été plus précis en parlant de l'opération « la plus coûteuse » en évoquant $(n_{occ}^m n_{vir}^n)$. Pour *CC2*, nous avons été amenés à compter le nombre d'opérations totales effectuées lors d'un calcul *CC2* par le fait qu'une première implémentation rendait les calculs presque aussi longs qu'un calcul *CCSD*.

D'un point de vue théorique, le coût d'un calcul tensoriel se « mesure » en comptant le nombre d'*OM* virtuelles et occupées qui interviennent dans chaque produit. Prenons l'exemple suivant :

$$A_{o_1}^{v_1} = B_{v_b v_c o_j}^{v_1 v_a o_i} \otimes C_{o_1 v_a o_i}^{v_b v_c o_j}, \quad (5.4.2)$$

il y a 5 contractions : entre deux occupées $\{o_i; o_j\}$ et entre trois virtuelles $\{v_a; v_b; v_c\}$; nous avons donc respectivement o^2 et v^3 opérations. Pour construire

l'ensemble du tenseur A , il faut renouveler ces contractions v et o fois pour couvrir respectivement l'ensemble des v_1 et o_1 virtuelles et occupées. Nous arrivons à $N = o^3 \cdot v^4$ opérations.

Ainsi, pour réaliser la contraction de deux tenseurs, il faut N opérations, telles que⁹ :

$$N = v^{\#\{v\}} \cdot o^{\#\{o\}} \quad (5.4.3)$$

Pour comparer les coûts computationnels pour les méthodes $CC2$ et $CCSD$, nous avons compté le nombre d'opérations pour chaque équation programmée en fonction de n_{occ} et n_{vir} . Sur la figure (fig.5.2), les courbes vertes représentent les abaques donnant pour chaque n_{occ} le rapport du nombre d'opérations $\frac{N_{CC2}}{N_{CCSD}}$ en fonction de n_{vir} , i.e. v . Nous avons indiqué par des « + » le minimum de chacune de ces courbes et nous remarquons que pour n_{occ} compris entre 10 et 100, le rapport du nombre d'opérations peut être compris entre 2.5 et 50. Évidemment, ce gain computationnel n'est intéressant que dans la mesure où la qualité de cette approximation sera vérifiée pour des éléments lourds avec des systèmes d'intérêts.

5.4.2 Résultats préliminaires

Différents tests ont été réalisés pour vérifier si le code est correct (*test for correctness*) et performant. Tous ces tests ont été réalisés avec TAL-SH.

Énergie

Les énergies ont été comparées avec le code MRCC [99] et le code Dalton [3], dans un cadre non-relativiste. Le premier calcul s'est porté sur l'eau (tab.5.3), en base cc-pv2 ζ ¹⁰. Nous voyons que les résultats en énergie sont bons, les équations pour les amplitudes t_1 et t_2 ont été correctement implémentées.

9. Nombre d'éléments M uniques d'un ensemble $\{m\}$: $M = \#\{m_1, m_1, m_2, \dots, m_i, m_i, m_i\}$ [25].

10. $d_{O-H}=0,975\ 512\ \text{Å}$, $H\hat{O}H=110,565^\circ$.

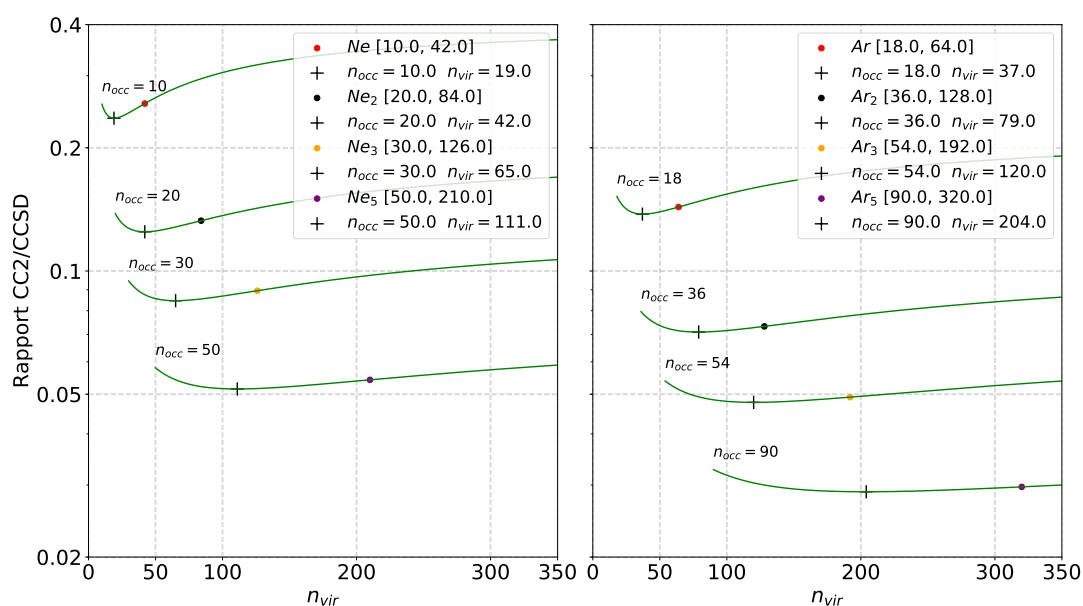


FIGURE 5.2 – Abaque CC2. Pour le nombre OM occupées $n_{occ} = \{10; 20; 30; 50\}$ sur le graphique de gauche et pour $n_{occ} = \{18; 36; 54; 90\}$ sur celui de droite, calcul du rapport du nombre d'opérations $\frac{N(CC2)}{N(CCSD)}$ en fonction de n_{vir} . Les « + » indiquent le minimum de chaque courbe. La description des points et des légendes associées est donnée dans le texte .

	SCF	CC2
Dirac	-76,027 520 738 313 9	-76,287 253 312 919 9
code MRCC	-76,027 520 738 312 1	-76,287 253 312 935 0
Différence	$-1,8 \times 10^{-12}$	$15,1 \times 10^{-12}$
Dalton	-76,027 520 738 31	-76,287 253 312 99

TABLEAU 5.3 – H_2O : calcul non-relativiste, cc-pv2 ζ , comparaison avec le code MRCC et Dalton (en E_h).

Chaînes linéaires d'atomes de néon et d'argon

Nous avons calculé les énergies et mesuré les temps de calcul pour des chaînes linéaires d'atomes de néon et d'argon où les atomes sont à une distance

de $2,4 \text{ \AA}$ les uns des autres (*Base* cc-pv2 ζ , [∞E_h]). Pour comparer nos énergies au code MRCC, certains calculs ont été accomplis en utilisant la décomposition de CHOLESKY pour la transformation des OA en OM ; nous avons donc toujours les mêmes équations CC2.

Nous avons de nouveau utilisé la figure (fig.5.2), par rapport aux nombres n_{occ} et n_{vir} de ces deux systèmes pour calculer les rapports théoriques du nombre d'opérations. Ils sont représentés par des points sur la figure et écrits, dans la légende, sous la forme : Element_{Nombre d'atome} [n_{occ}, n_{vir}]. Nous noterons la faible différence entre les minimums des abaques (représentés par des croix) et les rapports théoriques attendus (représentés par des points).

Les résultats pour le néon et l'argon sont présentés respectivement dans les tableaux (tab.5.4) et (tab.5.5).

Premièrement, les énergies obtenues sont identiques à celles obtenues avec les codes MRCC et Dalton, à la précision demandée de $1 \times 10^{-9} E_h$ et ne sont pas reportées ici. Deuxièmement, il faut remarquer la qualité de la méthode CC2, avec laquelle toutes les valeurs calculées sont supérieures à celles obtenues par CCSD dans la limite de 1,2 eV pour 5Ar. Les erreurs sont croissantes ; il est difficile de parler de linéarité avec peu de valeur mais il n'y a pas d'incohérence. Troisièmement, il faut remarquer dans ces exemples que les calculs CC2 nécessitent moins d'itérations (N_{it}) pour atteindre la convergence. Il faudrait regarder comment les contractions $\hat{T}_2 \otimes \hat{V}$ du calcul CCSD influent initialement et à chaque itération mais aucune étude systématique n'a été réalisée.

Enfin, en ce qui concerne le rapport des temps de calcul pour chaque itération, $R1$, et le rapport théorique du nombre d'opérations, $R2$, la comparaison reste difficile. Dans l'idéal, le rapport est de 1 puisqu'il n'y a pas de raison qu'une opération soit plus longue à réaliser avec une méthode qu'avec l'autre. Nous pouvons dire, même si nous ne sommes que sur un seul nœud avec TAL-SH que ce sont certainement des problèmes de mémoire et il faudra mener une investigation plus approfondie pour en trouver l'origine. Quoiqu'il en soit, ce « comptage » terme à terme des coûts nous a permis de comprendre que c'était une mauvaise implémentation des intermédiaires qui expliquait la lenteur du calcul CC2 lors de la toute première version.

	Ne	2Ne	3Ne	5Ne ^a
$[n_{occ}, n_{vir}]$	[10,42]	[20,84]	[30,126]	[50,210]
SCF	-128,4888	-256,5577	-384,6102	-640,7111
CCSD				
Énergie	-128,7463	-257,0743	-385,3838	-641,9979
t(s)	4	254	2929	60523
N_{it}	10	12	13	13
$t_{CCSD}=t(s)/N_{it}$	0,4	21,2	225	4656
CC2				
Énergie	-128,7449	-257,0718	-385,3796	-641,9900
t(s)	3	66	668	3607
N_{it}	9	11	11	12
$t_{CC2}=t(s)/N_{it}$	0,33	6	61	301
R1 : Rapport t_{CCSD}/t_{CC2}	1,2	3,53	3,71	15,49
R2 : Rapport théorique ^b	4,23	8,02	11,82	18,45
Rapport R2/R1	3,53	2,27	3,19	1,19
$\Delta(CCSD - CC2)$ en(mE_h)	-1,364	-2,4671	-4,194	-7,886

TABLEAU 5.4 – Étude CC2 pour une chaîne linéaire d’atomes de néon - Énergie en (E_h) - cc-pv2 ζ - N_{it} est le nombre d’itérations - les atomes sont à 2,4 Å les uns des autres - ^a La décomposition de CHOLSKY a été utilisée (seuil à $1 \times 10^{-9} E_h$). ^b Valeurs issues de (fig.5.2).

5.4.3 Travaux en cours de réalisation et perspectives vers Exa-Tensor

Pour l’énergie CC2, il convient désormais d’utiliser ce code pour analyser la performance de la méthode sur des éléments lourds dans le cadre relativiste en comparaison avec CCSD.

En ce qui concerne les temps de calcul, un travail sur les intermédiaires est

	Ar	2Ar	3Ar ^a	5Ar ^a
$[n_{occ}, n_{vir}]$	[18, 64]	[36, 128]	[54, 192]	[90, 320]
HF	-526,7999	-1053,5470	-1580,2924	-2633,7829
CCSD				
Énergie	-527,1193	-1054,1900	-1581,2591	-2635,3967
t(s)	46	2705	53119	1182195
N_{it}	11	12	13	13
$t_{CCSD}=t(s)/N_{it}$	4,2	225	4086	90938
CC2				
Énergie	-527,1096	-1054,1719	-1581,2324	-2635,3531
t(s)	21	616	5850	64650
N_{it}	9	11	11	12
$t_{CC2}=t(s)/N_{it}$	2,3	56	532	5387
R1 : Rapport t_{CCSD}/t_{CC2}	1,79	4,03	7,68	16,88
R2 : Rapport théorique ^b	6,98	13,65	20,33	33,69
Rapport R2/R1	3,89	3,39	2,65	2,00
$\Delta(CCSD-CC2)$ en(mE_h)	-9,692	-18,179	-26,663	-43,644

TABLEAU 5.5 – Étude CC2 pour une chaîne linéaire d’atomes d’argon - Énergie en (E_h) - cc-pv2 ζ - N_{it} est le nombre d’itérations - les atomes sont à 2,4 Å les uns des autres - ^a La décomposition de CHOLESKY a été utilisée (seuil à $1 \times 10^{-9} E_h$).
^b Valeurs issues de (fig.5.2).

toujours intéressant, comme cela est apparu précédemment mais également avec les méthodes *EOM-CC* (cf.4.6). Éviter de calculer deux ou trois fois les mêmes produits tensoriels est un avantage certain. Mais les problèmes liés à des contractions sous-optimales apparaissent à chaque étape. Si ExaCorr est prévu pour calculer rapidement, c’est la mémoire qu’il faut éviter de « saturer », i.e. le problème du découpage en lot. Cette complication apparaît constamment ;

imaginons l'opération :

$$t_i^a = - \sum_{me} f_e^m t_i^e t_m^a \quad (5.4.4)$$

Nous avons les deux contractions suivantes :

$$X_e^a = f_e^m \otimes t_m^a \quad (5.4.5)$$

$$t_i^a = -X_e^a \otimes t_i^e, \quad (5.4.6)$$

mais nous préférons réaliser :

$$X_i^m = f_e^m \otimes t_i^e \quad (5.4.7)$$

$$t_i^a = -X_i^m \otimes t_m^a \quad (5.4.8)$$

En effet, le transfert des données, au travers de X_i^m ne concerne alors qu'un tenseur de taille n_{occ}^2 . Là encore, ces éléments doivent être pensés à chaque étape lors de calculs parallélisés. Après l'approche diagrammatique qui nous a permis d'obtenir les équations programmables, l'implémentation doit également être pensée avec soin.

Ensuite, la programmation des équations λ et des matrices densités pour *CC2* a également été réalisée, mais il reste une erreur d'implémentation qui nous empêche de montrer des résultats finalisés. En effet, pour une molécule d'eau en base *cc-pv2 ζ* , avec *CC2* nous obtenons un moment dipolaire de $p = -0,757\,873\,7$ u a . Par comparaison avec le code *MRCC*¹¹ avec lequel nous obtenons une valeur de $p = -0,757\,840\,2$ u a , cela fait une différence de $-3,35 \times 10^{-5}$ u a . Le travail est encore en cours quant à trouver la contraction responsable de cette erreur.

Enfin, une fois *CC2* terminé, nous pourrions également travailler sur *EOM-CC2*¹². Une étude menée avec cette méthode par Dutta *et al.* [53] sur 24 colorants organiques montre que les *IP* de cœur sont obtenus avec un écart-type relativement équivalent à *EOM-CC* par rapport à *EOM-CCSDT* ($\sigma_{EOM-CCSD} = 0,48$ eV et $\sigma_{EOM-CC(2)} = 0,37$ eV). De plus, *CC2* est complémentaire à *P-EOMCC-MBPT(2)*. Nous citerons Goings *et al.* [70] :

11. Avec lequel aucune différence n'est observée avec *CCSD* ni pour l'énergie ni pour le moment dipolaire.

12. Indépendamment de *LR-CC2* [37].

As low-scaling wave function based methods, P-EOM-MBPT2 and CC2 can be considered complementary excited state methods to each other. Both methods have the same expression for the doubles-doubles block[...]. [...], one way to view the difference between CC2 and P-EOM-MBPT2 is that the former favors \hat{T}_1 (or is " \hat{T}_1 heavy") and the latter favors \hat{T}_2 (" \hat{T}_2 heavy").

Il serait donc intéressant de pouvoir comparer les deux méthodes pour nos systèmes d'intérêts.

5.5 Théorie de la réponse linéaire - Application à la polarisabilité

Dans cette section, nous cherchons à obtenir la polarisabilité α , propriété du 2^e ordre, à l'aide de la théorie de la réponse linéaire appliquée à CCSD. Nous utilisons les développements théoriques de Christiansen *et al.* [36, 38]¹³ dans lesquels la décomposition en série de FOURIER nous permet de simplifier les calculs pour des perturbations périodiques.

La polarisabilité est la propriété d'un système à acquérir un moment dipolaire sous l'action d'un champ électrique. Le moment dipolaire induit \vec{p} est proportionnel au champ électrique \vec{E} appliqué selon la relation :

$$\vec{p} = \bar{\alpha}\vec{E}, \quad (5.5.1)$$

$\bar{\alpha}$ est un tenseur de rang 2. C'est une propriété de la matière. D'un point de vue macroscopique, les effets sont, par exemple, la modification de l'indice de réfraction (impliquant la biréfringence). En chimie quantique, la polarisabilité est une grandeur déjà calculée pour de nombreux systèmes [132, 115, 170] et nous pouvons nous servir de cette donnée pour paramétrer les champs de force [158, 159].

Les propriétés statiques peuvent être calculées par des méthodes analytiques, ou par des méthodes numériques avec l'utilisation de *finite fields* [146]. Pour les propriétés dynamiques, nous utiliserons de façon analytique des perturbations

13. Notons les travaux de Koch *et al.* [102] et de Pawłowski *et al.* [147].

dépendantes du temps.

$$\alpha_{XY}(\omega_X; -\omega_Y) = \langle\langle \hat{X}; \hat{Y} \rangle\rangle_{\omega_Y} \quad (5.5.2)$$

Le calcul des propriétés par la théorie *LR-CC* a déjà un long historique [104, 101]. Mais nous avons pour but de programmer les expressions de *LR-CC*¹⁴ dans ExaCorr, prenant ainsi en compte les effets relativistes, et pour prérogative d'utiliser une méthode itérative pour la résolution des équations de réponse, basée sur l'implémentation de la méthode de Davidson.

5.5.1 Polarisation dans le cadre de CC

Nous avons pris comme point de départ le formalisme de la *Quasi-Energy* Q [113, 162] qui est la généralisation de l'énergie dans le cas dépendant du temps. Cela nous permet de travailler avec des méthodes variationnelles ou non (comme c'est le cas ici pour *CC*). Si $H = H_0 + V(t)$ où $V(t)$ est la perturbation, nous pouvons écrire :

$$Q(t) = \left\langle \tilde{0} \left| \left(H - i \frac{\partial}{\partial t} \right) \right| \tilde{0} \right\rangle, \quad (5.5.3)$$

où $|\tilde{0}\rangle = |0\rangle$ dans la limite où la perturbation tend vers 0 et qui peut être déterminée par des méthodes variationnelles.

$V(t)$ est une perturbation périodique, représentée par un opérateur \hat{X} , d'intensité ε_X , décomposable en une série de Fourier :

$$V(t) = \sum_x \varepsilon_X(0)X + \sum_{k=-N}^N e^{-i\omega_k t} \sum_x \varepsilon_X(\omega_k)X \quad (5.5.4)$$

Notons qu'utiliser des exponentielles complexes nous oblige à considérer des fréquences de $-N$ à N , et pour que l'hamiltonien reste hermitien, il faut imposer : $\omega_{-k} = -\omega_k$ et $\varepsilon_X(\omega_k) = \varepsilon_X^*(\omega_k)$.

Dans le cas où \hat{X} serait l'observable d'intérêt et \hat{Y} ou \hat{Z} des perturbations

¹⁴. Plutôt que *Sum Over State* ou l'algorithme *Direct Inversion in the Iterative Subspace (DIIS)* [153].

externes ¹⁵ :

$$\begin{aligned} \langle X \rangle(t) &= \langle \bar{0} | X | \bar{0} \rangle \\ &= \langle X \rangle_0 + \sum_{k_1} e^{-i\omega_{k_1} t} \sum_y \langle \langle X; Y \rangle \rangle_{\omega_{k_1}} \varepsilon_y(\omega_{k_1}) \\ &\quad + \frac{1}{2} \sum_{k_1, k_2} e^{-i(\omega_{k_1} + \omega_{k_2}) t} \sum_{y, z} \langle \langle X; Y, Z \rangle \rangle_{\omega_{k_1}, \omega_{k_2}} \varepsilon_y(\omega_{k_1}) \varepsilon_z(\omega_{k_2}) + \dots \end{aligned} \quad (5.5.5)$$

En définissant également la moyenne sur une période T :

$$\{Q\}_T = \frac{1}{T} \int_{-T/2}^{T/2} Q dt, \quad T | V^t(t+T) = V^t(t) \quad (5.5.6)$$

Nous pouvons exprimer $\{Q\}_T$ en fonction de la perturbation :

$$\begin{aligned} \{Q\}_T &= E_0 + \sum_x \langle X \rangle \varepsilon_x(0) + \frac{1}{2} \sum_{x, y} \sum_{k_1} \langle \langle X; Y \rangle \rangle_{\omega_{k_1}} \varepsilon_x(-\omega_{k_1}) \varepsilon_y(\omega_{k_1}) \\ &\quad + \frac{1}{6} \sum_{x, y, z} \sum_{k_1, k_2} \langle \langle X; Y, Z \rangle \rangle_{\omega_{k_1}, \omega_{k_2}} \varepsilon_x(-\omega_{k_1} - \omega_{k_2}) \varepsilon_y(\omega_{k_1}) \varepsilon_z(\omega_{k_2}) + \dots \end{aligned} \quad (5.5.7)$$

La valeur moyenne temporelle, i.e. l'intégration par rapport au temps sur une période, impose que la somme des pulsations soit nulle.

La polarisabilité s'exprime alors comme la dérivée de $\{Q\}_T$ en fonction des intensités des perturbations :

$$\langle \langle X; Y \rangle \rangle_{\omega_{k_1}} = \frac{d^2 \{Q\}_T}{d\varepsilon_x(\omega_{k_1}) d\varepsilon_y(-\omega_{k_1})} \quad (5.5.8)$$

Nous pouvons remarquer que cette expression est semblable à celle obtenue dans le cas indépendant du temps pour la polarisabilité statique. Nous cherchons maintenant à appliquer le principe variationnel en utilisant un ensemble des paramètres λ , dont dépend la fonction d'onde, en minimisant la moyenne temporelle quasi-énergie $\{Q(t, \lambda)\}$. Cependant, la théorie *Coupled-Cluster* n'est pas variationnelle, les paramètres λ de la fonction d'onde dépendent de cer-

15. La différence s'écrivant mathématiquement par la séparation à l'aide de « ; ».

taines conditions. Pour résoudre ce problème, nous utilisons alors l'expression lagrangienne (sous contrainte) :

$$\mathcal{L}(t, \lambda, \bar{\lambda}) = Q(t, \lambda) + \bar{\lambda}e(t, \lambda) \quad (5.5.9)$$

Ici apparaissent de nouveaux paramètres, les multiplicateurs de Lagrange $\bar{\lambda}$, et $e(t, \lambda)$ de nouvelles fonctions dépendantes du temps.

Nous ne développerons pas les différentes équations du lagrangien selon la théorie CC et l'application des règles (2n+1) pour les paramètres et (2n+2) pour les multiplicateurs [36, 84], nous poursuivons avec les expressions algébriques.

Ainsi, la polarisation s'exprime comme :

$$\langle\langle X, Y \rangle\rangle_{\omega_Y} = \frac{1}{2} C^{\pm\omega} P(X(\omega_X), Y(\omega_Y)) \times \left[\eta^X + \frac{1}{2} t^X F(\omega_X) \right] t^Y(\omega_Y) \quad (5.5.10)$$

t^X et t^Y sont les amplitudes perturbées au premier ordre, et :

$$\eta_\mu^Y = \left\langle \Lambda \left[[Y, \tau_\mu] \right] \middle| \text{CC} \right\rangle \quad (5.5.11)$$

$$F_{\mu\nu} = \left\langle \Lambda \left[[[H_0, \tau_\mu], \tau_\nu] \right] \middle| \text{CC} \right\rangle \quad (5.5.12)$$

$C^{\pm\omega} f(\omega_x, \omega_y) = f(\omega_x, \omega_y) + f^*(-\omega_x, -\omega_y)$ assure que la polarisabilité est réelle. L'opérateur de permutation $P(X(\omega_0), Y(\omega_{k_1}))$ assure que la polarisabilité est symétrique par changement d'observable \hat{X} et de perturbation \hat{Y} associés à leur fréquence respective.

Pour obtenir les amplitudes à l'ordre ⁽¹⁾ en perturbation, i.e. t^X et t^Y , nous devons résoudre :

$$(\mathbf{A} - \omega_Y \mathbf{I}) t^Y(\omega_Y) + \xi^Y = 0. \quad (5.5.13)$$

\mathbf{A} est le jacobien (non-symétrique).

Les expressions algébriques sont les suivantes :

$$\xi_\mu^Y = \langle \mu | Y | \text{CC} \rangle \quad (5.5.14)$$

$$A_{\mu\nu} = \langle \bar{\mu} | [H_0, \tau_\nu] | \text{CC} \rangle \quad (5.5.15)$$

5.5.2 Expressions programmables

Comme pour la partie *CC2*, nous devons traduire ces expressions algébriques en expressions programmables, i.e. en produits tensoriels. Avec de légères modifications de l'implémentation, nous pensons pouvoir utiliser l'algorithme de Davidson. Nous avons donc utilisé l'approche diagrammatique (Annexe E).

Pour parvenir aux amplitudes perturbées au premier ordre, nous devons résoudre l'équation (eq.5.5.13). Sachant que nous utiliserons un *vecteur d'essai* pour obtenir t^Y et que la matrice A , le jacobien qui n'est rien d'autre que \mathcal{H} , a déjà été programmée pour obtenir les équations *EOM-CC*, nous n'avons besoin que de l'expression de ξ (eq.5.5.14) : $\xi_\mu^Y = \langle \mu | Y | CC \rangle$.

ξ_μ^Y

Le moment dipolaire étant un opérateur 1-électron, nous utilisons l'expression $Y_i^a = \langle a | Y | i \rangle$. Nous avons dû développer l'ensemble des diagrammes permettant la projection de $Y | CC \rangle$ sur les déterminants simples et doubles excitations $\mu = \left(\begin{array}{l} \langle \mu_i^a | \\ \langle \mu_{ij}^{ab} | \end{array} \right)$. Nous présentons deux diagrammes en figure (fig.5.3) sur lesquels nous pouvons observer la projection qu'il faudra appliquer sur μ .

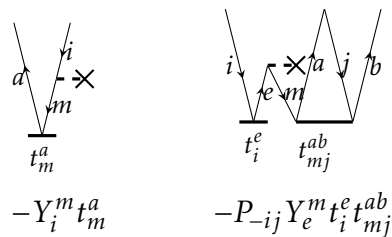


FIGURE 5.3 – Exemple de diagrammes présents dans la construction du tenseur ξ .

Nous arrivons, avec $P_{-uv}f(u, v) = f(u, v) - f(v, u)$, à ¹⁶ :

$$\xi_i^{Ya} = +Y_i^a + Y_e^a t_i^e - Y_i^m t_m^a - Y_e^m t_i^e t_m^a + Y_e^m t_{im}^{ae} \quad (5.5.16)$$

$$\xi_{ij}^{Yab} = +P_{-ab} Y_f^b t_{ij}^{af} - P_{-ij} Y_j^m t_{im}^{ab} - P_{-ij} Y_e^m t_i^e t_{mj}^{ab} - P_{-ab} Y_f^m t_{ij}^{fb} t_m^a, \quad (5.5.17)$$

Nous avons donc la possibilité de résoudre l'équation (eq.5.5.13) : $(\mathbf{A} - \omega_Y \mathbf{1}) t^Y (\omega_Y) + \xi^Y = 0$. En effet, si nous associons t^Y à la matrice des vecteurs d'essais $\mathbf{B}^{(m)}$, et si nous posons $\xi^Y = -\mathbf{A}\mathbf{B}^{(m)} + \omega_Y \mathbf{1}\mathbf{B}^{(m)}$, nous retrouvons la forme $R_\sigma = \mathcal{H}\mathbf{b}$ (cf.1.5.4). De plus, $\mathbf{B}^{H(m)}\xi^Y = -\tilde{\mathbf{A}} + \omega_Y \mathbf{1}^{(m)}$, où $\tilde{\mathbf{A}}$ est la matrice dans l'espace réduit. Nous pouvons ensuite réaliser la diagonalisation de $\tilde{\mathbf{A}}$, i.e. $(\tilde{\mathbf{A}} - (-\omega_Y + \lambda_k^{(m)}))C_k^{(m)} = 0$ et utiliser $t^Y = \mathbf{B}^{(m)}C_k^{(m)}$ où t^Y retrouve la dimension $((t_1 + t_2))$.

η_μ^Y

Ensuite, nous devons trouver l'expression de η_μ^Y (eq.5.5.11) : $\eta_\mu^Y = \langle \Lambda \left[\left[Y, \tau_\mu \right] \right] \text{CC} \rangle$. Il a été plus difficile de déterminer tous les éléments de ce tenseur de par la présence de λ , qui autorise des contractions supplémentaires ¹⁷ :

$$\begin{aligned} \eta_c^{Yk} = & Y_c^k + Y_c^e \lambda_e^k - Y_m^k \lambda_c^m - Y_c^m t_m^e \lambda_e^k - Y_e^k t_m^e \lambda_c^m + t_m^f \lambda_{ce}^{km} Y_f^e - t_n^e \lambda_{ce}^{km} Y_m^n + \lambda_{ce}^{km} Y_m^e \\ & - \frac{1}{2} t_{mn}^{fe} \lambda_{fe}^{mk} Y_c^n - \frac{1}{2} t_{nm}^{fe} \lambda_{fc}^{nm} Y_e^k + t_{mn}^{ef} \lambda_{ce}^{km} Y_f^n - t_n^e t_m^f \lambda_{ce}^{km} Y_f^n \end{aligned} \quad (5.5.18)$$

$$\eta_{cd}^{Ykl} = P_{-kl} P_{-cd} \lambda_c^k Y_d^l + P_{-cd} \lambda_{ce}^{kl} Y_d^e - P_{-kl} \lambda_{cd}^{km} Y_m^l - P_{-cd} t_m^e \lambda_{ce}^{kl} Y_d^m - P_{-ij} t_m^e \lambda_{cd}^{km} Y_e^l \quad (5.5.19)$$

à projeter sur $(|v_k^c\rangle |v_{kl}^{cd}\rangle)$.

t^{XF}

Enfin, il reste l'expression (eq.5.5.12) : $F_{\mu\nu} = \langle \Lambda \left[\left[\left[H_0, \tau_\mu \right], \tau_\nu \right] \right] \text{CC} \rangle$. Pour des raisons de connexion entre les différents tenseurs, il est seulement possible de

16. Les termes colorés correspondent aux diagrammes de la figure (fig.5.3).

17. Pour ne pas alourdir les équations, nous avons utilisé la notation $\lambda = \bar{t}^{(0)}$.

construire $t^X.F$, pour avoir l'ensemble des éléments sous la forme $\left(\begin{smallmatrix} x \\ tF \end{smallmatrix}\right)_c^k$. De plus, l'apparition de $e^{-\hat{T}}\hat{H}_N e^{\hat{T}}$ grâce à la commutation $[\tau_\mu, e^{\hat{T}}]$ nous permet d'utiliser les intermédiaires \mathcal{F} et \mathcal{W} déjà utilisés pour les calculs des amplitudes t et des multiplicateurs $\bar{t}^{(0)} = \lambda$. Nous arrivons donc à deux tenseurs¹⁸ :

$$\begin{aligned} \left(\begin{smallmatrix} x \\ tF \end{smallmatrix}\right)_c^k = & + t_e^x \mathcal{W}_{ca}^{ki} - \lambda_c^i t_i^x \mathcal{F}_a^k - \lambda_a^k t_i^x \mathcal{F}_c^i + \lambda_e^k t_i^x \mathcal{W}_{ca}^{ei} - \lambda_c^m t_i^x \mathcal{W}_{ma}^{ki} - \lambda_a^m t_i^x \mathcal{W}_{mc}^{ik} \\ & + \lambda_e^i t_i^x \mathcal{W}_{ac}^{ek} - \frac{1}{2} \lambda_c^i t_{ij}^x \mathcal{W}_{ab}^{kj} - \frac{1}{2} \lambda_a^k t_{ij}^x \mathcal{W}_{cb}^{ij} + \lambda_a^i t_{ij}^x \mathcal{W}_{bc}^{jk} - \lambda_{ae}^{mk} t_i^x \mathcal{W}_{mc}^{ie} - \lambda_{ec}^{im} t_i^x \mathcal{W}_{am}^{ek} \\ & + \frac{1}{2} \lambda_{ef}^{ki} t_i^x \mathcal{W}_{ca}^{ef} + \frac{1}{2} \lambda_{ca}^{mn} t_i^x \mathcal{W}_{mn}^{ki} - \frac{1}{2} \lambda_{ac}^{ij} t_{ij}^x \mathcal{F}_b^k - \frac{1}{2} \lambda_{ab}^{ik} t_{ij}^x \mathcal{F}_c^j - \lambda_{eb}^{ik} t_{ij}^x \mathcal{W}_{ac}^{ej} \\ & + \lambda_{ac}^{mj} t_{ij}^x \mathcal{W}_{mb}^{ik} - \frac{1}{4} \lambda_{ec}^{ij} t_{ij}^x \mathcal{W}_{ab}^{ek} + \frac{1}{4} \lambda_{ab}^{mk} t_{ij}^x \mathcal{W}_{mc}^{ij} + \frac{1}{2} \lambda_{ae}^{ij} t_{ij}^x \mathcal{W}_{bc}^{ek} \\ & - \frac{1}{2} \lambda_{ab}^{im} t_{ij}^x \mathcal{W}_{mc}^{jk} \end{aligned} \quad (5.5.20)$$

$$\begin{aligned} \left(\begin{smallmatrix} x \\ tF \end{smallmatrix}\right)_{cd}^{kl} = & + P_{-cd} P_{-kl} \lambda_c^k t_i^x \mathcal{W}_{da}^{li} - P_{-kl} \lambda_a^k t_i^x \mathcal{W}_{cd}^{il} - P_{-cd} \lambda_c^i t_i^x \mathcal{W}_{ad}^{kl} - P_{-kl} \lambda_{cd}^{ki} t_i^x \mathcal{F}_a^l - P_{-cd} \lambda_{ca}^{kl} t_i^x \mathcal{F}_d^i \\ & + P_{-kd} \lambda_{ce}^{kl} t_i^x \mathcal{W}_{da}^{ei} - P_{-kl} \lambda_{cd}^{km} t_i^x \mathcal{W}_{ma}^{li} + P_{-cd} P_{-kl} \lambda_{ca}^{ml} t_i^x \mathcal{W}_{md}^{ki} + \lambda_{cd}^{mi} t_i^x \mathcal{W}_{ma}^{kl} \\ & - P_{-cd} P_{-kl} \lambda_{ed}^{ki} t_i^x \mathcal{W}_{ca}^{el} - \lambda_{ea}^{kl} t_i^x \mathcal{W}_{cd}^{ei} - \frac{1}{2} P_{-cd} \lambda_{ca}^{kl} t_{ij}^x \mathcal{W}_{db}^{ij} - \frac{1}{2} P_{-kl} \lambda_{cd}^{ki} t_{ij}^x \mathcal{W}_{ab}^{lj} \\ & + \frac{1}{4} \lambda_{cd}^{ij} t_{ij}^x \mathcal{W}_{ab}^{kl} + \frac{1}{4} \lambda_{ab}^{kl} t_{ij}^x \mathcal{W}_{cd}^{ij} + P_{-kl} P_{-cd} \lambda_{cb}^{kj} t_{ij}^x \mathcal{W}_{da}^{li} - \frac{1}{2} P_{-kl} \lambda_{ab}^{ik} t_{ij}^x \mathcal{W}_{cd}^{jl} \\ & - \frac{1}{2} P_{-cd} \lambda_{ac}^{ij} t_{ij}^x \mathcal{W}_{bd}^{kl} \end{aligned} \quad (5.5.21)$$

à projeter sur $(|v_k^c\rangle |v_{kl}^{cd}\rangle)$.

Nous avons programmé l'ensemble de ces expressions. Mais, nous ne sommes pas en mesure de donner nos premières valeurs. La construction des tenseurs, notamment tF , a présenté de nombreux défis. Cependant, le seul élément manquant étant la résolution de l'équation des amplitudes perturbées $(\mathbf{A} - \omega_Y \mathbf{I}) t^Y (\omega_Y) + \xi^Y = 0$, nous espérons obtenir rapidement nos premiers résultats.

18. Pour des raisons de lecture, nous avons préféré écrire $\overset{x}{t}$ plutôt que t^X .

5.6 Conclusion

Dirac possède donc désormais un nouveau module de calcul tensoriel ExaCorr permettant la gestion de calcul hautement parallélisé. Malgré les réglages à réaliser pour l'utiliser de façon optimale, la portabilité est excellente et cela facilite le débogage, notamment avec TAL-SH.

Plusieurs travaux sont en cours ; nous avons évoqué la matrice simple densité pour le moment dipolaire avec *CC2*, mais en fait c'est l'ensemble des matrices simples et doubles densité *CCSD* qui ont été programmées et nous travaillons à leur finalisation (cf. Annexe E). C'est un vrai challenge de les implémenter de façon optimale au vu du nombre de diagrammes qui entrent en jeu.

Toujours pour *CC2*, nous espérons disposer de EOM-*CC2* et pouvoir comparer avec les méthodes approximées déjà implémentées.

De plus, nous possédons désormais les équations programmables nous permettant d'obtenir la polarisabilité dans le cadre *LR-CC* dans un formalisme relativiste, les équations ne sont pas encore implémentées de façon optimale et il conviendra d'y parvenir avant de pouvoir commencer l'étude de plus gros systèmes. Bien évidemment, il sera nécessaire de commencer par vérifier le cas historique qu'est LiH avant toute autre chose [104, 31]. Le jacobien de ce système étant relativement petit, l'inversion pourrait être réalisée prochainement ; nous pourrions alors obtenir les expressions des amplitudes perturbées. Nous espérons également disposer de l'algorithme de Davidson dans un avenir relativement proche.

Enfin, les perspectives sont nombreuses ; celle qui semble cependant le plus à notre portée, une fois le formalisme *LR-CC* terminé, est la réponse quadratique $\langle\langle X; Y, Z \rangle\rangle$ (l'hyperpolarisabilité [35]), ce qui nécessitera alors la prise en compte des multiplicateurs de LAGRANGE à l'ordre 1.

Conclusion

Dans cette thèse, nous avons travaillé à comprendre le comportement de méthodes basées sur *Equation of Motion Coupled-Cluster*, dans un cadre relativiste, pour obtenir des propriétés de valence et de cœur.

Nous avons montré l'intérêt d'utiliser la méthode *Core-Valence Separation-EOM* (chapitre 3) pour obtenir *les potentiels d'ionisation et les énergies d'excitation* des électrons de cœur. Certaines combinaisons d'approximations associées à CVS fonctionnent correctement (*Frozen Core, No Double*). Ces méthodes sont intéressantes pour des systèmes contenant des éléments lourds possédant beaucoup d'électrons internes; nous pourrions y gagner en temps de calculs. Enfin, avec les systèmes étudiés, les niveaux d'énergie atteints nous imposent d'utiliser des hamiltoniens relativistes; nous avons dans ce sens montré que la méthode à 2-composantes ${}^2\text{DC}^{\text{M}}$ donne des résultats identiques à ceux de méthodes à 4-composantes jusqu'à la période de l'iode. Pour les périodes suivantes, les erreurs dues au *picture-change* deviennent plus importantes. L'ajout du terme de Gaunt et des effets dus à l'électrodynamique quantique sont également à prendre en compte pour atteindre la précision expérimentale.

Nous avons mis en évidence que certaines approximations, *Partitioned* et *Many-Body Perturbation Theory 2d order* (chapitre 4), pouvaient être appliquées à la matrice *EOM* pour limiter les coûts de calculs. La meilleure approximation à utiliser n'est pas toujours la même selon la propriété désirée. Cependant, la performance de *Partitioned-EOMCC-MBPT(2)* dans les cas étudiés pose des questions intéressantes. Il faudra étudier le comportement de toutes ces approximations pour des systèmes de plus grande taille. La mise en œuvre a été faite avec la bibliothèque RELCCSD, il faudra également vérifier leurs performances sur ExaCorr.

Comme les besoins informatiques sont de plus en plus importants, nous avons commencé à étudier l'intérêt de ExaCorr (chapitre 5), nouvelle bibliothèque, permettant dans *Dirac* de paralléliser les méthodes post-*HF* de façon performante. Nous avons profité de ce nouveau code pour implémenter, entre autres, *CC2*. Enfin, nous avons fait des avancées vers l'implémentation des méthodes de réponse linéaire *LR-CC*. Nous nous sommes focalisés dans cette thèse sur des perturbations électriques. Ayant obtenu les expressions programmables pour la polarisabilité, nous sommes en train de travailler à leur mise en œuvre.

Il convient de remarquer que l'approche diagrammatique a été utilisée à chaque étape. Dessiner des diagrammes, pour effectuer le passage entre les expressions algébriques et les expressions programmables, a représenté l'essentiel de notre travail sur le formalisme des méthodes *CCSD*, *EOM-CC* et *LR-CC*.

Enfin, comme nous l'avons indiqué tout au long de cette thèse, les perspectives sont nombreuses. Il reste un grand travail pour mieux prendre en compte la corrélation électronique dans les méthodes approximées. Pour améliorer la précision de *EOM*, les amplitudes triples T_3 pourront être implémentées de façon explicite ou de façon perturbative. Pour le coût de calcul, un soin particulier doit être porté à l'implémentation et il reste du travail sur le comportement des bibliothèques parallélisant les calculs. À l'avenir, les problèmes de mémoire pourront être réduits par l'application de la décomposition de *CHOLESKY* à la matrice *EOM*, mais nous pourrions également envisager *similarity-transformed-EOM* ou encore *EOM-CC2*.

Bibliographie

- [1] DIRAC, a relativistic ab initio electronic structure program, Release DIRAC19 (2019), written by A. S. P. Gomes, T. Saue, L. Visscher, H. J. Aa. Jensen, and R. Bast, with contributions from I. A. Aucar, V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, E. Faßhauer, T. Fleig, O. Fossgaard, L. Halbert, E. D. Hedegård, B. Heimlich–Paris, T. Helgaker, J. Henriksson, M. Iliaš, Ch. R. Jacob, S. Knecht, S. Komorovský, O. Kullie, J. K. Lærdahl, C. V. Larsen, Y. S. Lee, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, J. M. H. Olsen, Y. C. Park, J. K. Pedersen, M. Pernpointner, R. di Remigio, K. Ruud, P. Sałek, B. Schimmelpfennig, B. Senjean, A. Shee, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, M. L. Vidal, S. Villaume, O. Visser, T. Winther, and S. Yamamoto (available at <http://dx.doi.org/10.5281/zenodo.3572669>, see also <http://www.diracprogram.org>).
- [2] Rüdiger ACHILLES et Andrea BONFIGLIOLI. « The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin ». In : *Archive for History of Exact Sciences* 66.3 (avr. 2012), p. 295-358. DOI : 10.1007/s00407-012-0095-8. URL : <https://doi.org/10.1007/s00407-012-0095-8>.
- [3] Kestutis AIDAS et al. « The Dalton quantum chemistry program system : The Dalton program ». en. In : *Wiley Interdisciplinary Reviews : Computational Molecular Science* 4.3 (mai 2014), p. 269-284. ISSN : 17590876. DOI : 10.1002/wcms.1172. URL : <http://doi.wiley.com/10.1002/wcms.1172> (visité le 12/06/2021).
- [4] Jan ALMLÖF, Trygve HELGAKER et Peter R. TAYLOR. « Gaussian basis sets for high-quality ab initio calculations ». In : *The Journal of Physical Chemistry* 92.11 (juin 1988), p. 3029-3033. DOI : 10.1021/j100322a003. URL : <https://doi.org/10.1021/j100322a003>.
- [5] Adel ALMOUKHALALATI et al. « Electron correlation within the relativistic no-pair approximation ». In : *The Journal of Chemical Physics* 145.7 (août 2016), p. 074104. DOI : 10.1063/1.4959452. URL : <https://doi.org/10.1063/1.4959452>.

- [6] Carl D. ANDERSON. « The Positive Electron ». In : *Physical Review* 43.6 (mar. 1933), p. 491-494. DOI : 10.1103/physrev.43.491. URL : <https://doi.org/10.1103/physrev.43.491>.
- [7] Dirk ANDRAE. « Finite nuclear charge density distributions in electronic structure calculations for atoms and molecules ». In : *Physics Reports* 336.6 (oct. 2000), p. 413-525. DOI : 10.1016/s0370-1573(00)00007-7. URL : [https://doi.org/10.1016/s0370-1573\(00\)00007-7](https://doi.org/10.1016/s0370-1573(00)00007-7).
- [8] *Aurora - Argonne Leadership Computing Facility*. 2021. URL : <https://www.alcf.anl.gov/aurora>.
- [9] Evert J. BAERENDS et al. « Relativistic atomic orbital contractions and expansions : magnitudes and explanations ». In : *Journal of Physics B : Atomic, Molecular and Optical Physics* 23.19 (oct. 1990), p. 3225-3240. DOI : 10.1088/0953-4075/23/19/010. URL : <https://doi.org/10.1088/0953-4075/23/19/010>.
- [10] Paul S. BAGUS. « Self-Consistent-Field Wave Functions for Hole States of Some Ne-Like and Ar-Like Ions ». In : *Physical Review* 139.3A (août 1965), A619-A634. DOI : 10.1103/physrev.139.a619. URL : <https://doi.org/10.1103/physrev.139.a619>.
- [11] Paul S. BAGUS, Eugene S. ILTON et Connie J. NELIN. « The interpretation of XPS spectra : Insights into materials properties ». In : *Surface Science Reports* 68.2 (juin 2013), p. 273-304. DOI : 10.1016/j.surfrep.2013.03.001. URL : <https://doi.org/10.1016/j.surfrep.2013.03.001>.
- [12] Paul S. BAGUS et al. « Mechanisms responsible for chemical shifts of core-level binding energies and their relationship to chemical bonding ». In : *Journal of Electron Spectroscopy and Related Phenomena* 100.1-3 (oct. 1999), p. 215-236. DOI : 10.1016/s0368-2048(99)00048-1. URL : [https://doi.org/10.1016/s0368-2048\(99\)00048-1](https://doi.org/10.1016/s0368-2048(99)00048-1).
- [13] Albert A. BAKKE, Hsiang-Wen CHEN et William L. JOLLY. « A table of absolute core-electron binding-energies for gaseous atoms and molecules ». In : *Journal of Electron Spectroscopy and Related Phenomena* 20.3 (jan. 1980), p. 333-366. DOI : 10.1016/0368-2048(80)85030-4. URL : [https://doi.org/10.1016/0368-2048\(80\)85030-4](https://doi.org/10.1016/0368-2048(80)85030-4).
- [14] G. A. BARCLAY, T. M. SABINE et J. C. TAYLOR. « The crystal structure of rubidium uranyl nitrate : A neutron-diffraction study ». In : *Acta Crystallographica* 19.2 (août 1965), p. 205-209. DOI : 10.1107/s0365110x65003110. URL : <https://doi.org/10.1107/s0365110x65003110>.

- [15] Rodney J. BARTLETT et al. « Index of multi-determinantal and multi-reference character in coupled-cluster theory ». In : *The Journal of Chemical Physics* 153.23 (déc. 2020), p. 234103. DOI : 10.1063/5.0029339. URL : <https://doi.org/10.1063/5.0029339>.
- [16] Nelson H. F. BEEBE et Jan LINDERBERG. « Simplifications in the generation and transformation of two-electron integrals in molecular calculations ». en. In : *International Journal of Quantum Chemistry* 12.4 (oct. 1977), p. 683-705. ISSN : 0020-7608, 1097-461X. DOI : 10.1002/qua.560120408. URL : <http://doi.wiley.com/10.1002/qua.560120408> (visité le 11/06/2021).
- [17] BELENOS. URL : <http://www.meteofrance.fr/nous-connaitre/activites-et-metiers/les-supercalculateurs-de-meteo-france>.
- [18] Nicholas A. BESLEY. « Density Functional Theory Based Methods for the Calculation of X-Ray Spectroscopy ». In : *Accounts of Chemical Research* 53.7 (juil. 2020), p. 1306-1315. DOI : 10.1021/acs.accounts.0c00171. URL : <https://doi.org/10.1021/acs.accounts.0c00171>.
- [19] Nicholas A. BESLEY. « Modeling of the spectroscopy of core electrons with density functional theory ». In : *WIREs Computational Molecular Science* (mar. 2021). DOI : 10.1002/wcms.1527. URL : <https://doi.org/10.1002/wcms.1527>.
- [20] Nicholas A. BESLEY, Andrew T. B. GILBERT et Peter M. W. GILL. « Self-consistent-field calculations of core excited states ». In : *The Journal of Chemical Physics* 130.12 (mar. 2009), p. 124308. DOI : 10.1063/1.3092928. URL : <https://doi.org/10.1063/1.3092928>.
- [21] Francesco BORGATTI et al. « Revisiting the origin of satellites in core-level photoemission of transparent conducting oxides : The case of n-doped SnO₂ ». In : *Physical Review B* 97.15 (avr. 2018). DOI : 10.1103/physrevb.97.155102. URL : <https://doi.org/10.1103/physrevb.97.155102>.
- [22] Max BORN et Robert OPPENHEIMER. « Zur Quantentheorie der Molekeln ». In : *Annalen der Physik* 389.20 (1927), p. 457-484. DOI : 10.1002/andp.19273892002. URL : <https://doi.org/10.1002/andp.19273892002>.
- [23] Yassine BOUCHAFRA et al. « Predictive Simulations of Ionization Energies of Solvated Halide Ions with Relativistic Embedded Equation of Motion Coupled Cluster Theory ». In : *Physical Review Letters* 121.26 (déc. 2018). DOI : 10.1103/physrevlett.121.266001. URL : <https://doi.org/10.1103/physrevlett.121.266001>.

- [24] Nacer BOUDJEMIA et al. « Deep core photoionization of iodine in CH₃I and CF₃I molecules : how deep down does the chemical shift reach? » In : *Physical Chemistry Chemical Physics* 21.10 (2019), p. 5448-5454. DOI : 10.1039/c8cp07307d. URL : <https://doi.org/10.1039/c8cp07307d>.
- [25] Alain BOUVIER, Michel GEORGE et François LE LIONNAIS. *Dictionnaire des mathématiques*. Paris : Presses universitaires de France, 1996. ISBN : 978-2-13047821-8.
- [26] Barbara BRENA et Yi LUO. « Time-dependent DFT calculations of core electron shake-up states of metal-(free)-phthalocyanines ». In : *Radiation Physics and Chemistry* 75.11 (nov. 2006), p. 1578-1581. DOI : 10.1016/j.radphyschem.2005.07.017. URL : <https://doi.org/10.1016/j.radphyschem.2005.07.017>.
- [27] David BRIGGS et Graham BEAMSON. « XPS studies of the oxygen 1s and 2s levels in a wide range of functional polymers ». en. In : *Analytical Chemistry* 65.11 (juin 1993), p. 1517-1523. ISSN : 0003-2700, 1520-6882. DOI : 10.1021/ac00059a006. URL : <https://pubs.acs.org/doi/abs/10.1021/ac00059a006> (visité le 16/06/2021).
- [28] Gerald E. BROWN et David G. RAVENHALL. « On the interaction of two electrons ». In : *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 208.1095 (sept. 1951), p. 552-559. DOI : 10.1098/rspa.1951.0181. URL : <https://doi.org/10.1098/rspa.1951.0181>.
- [29] James B. BURKHOLDER, R. A. COX et A. R. RAVISHANKARA. « Atmospheric Degradation of Ozone Depleting Substances, Their Substitutes, and Related Species ». In : *Chemical Reviews* 115.10 (avr. 2015), p. 3704-3759. DOI : 10.1021/cr5006759. URL : <https://doi.org/10.1021/cr5006759>.
- [30] Laurent CANTREL et al. « Research Works on Iodine and Ruthenium Behavior in Severe Accident Conditions ». In : *Journal of Nuclear Engineering and Radiation Science* 4.2 (mar. 2018). DOI : 10.1115/1.4038223. URL : <https://doi.org/10.1115/1.4038223>.
- [31] Marco CARICATO, Gary W. TRUCKS et Michael J. FRISCH. « On the difference between the transition properties calculated with linear response- and equation of motion-CCSD approaches ». In : *The Journal of Chemical Physics* 131.17 (nov. 2009), p. 174104. DOI : 10.1063/1.3255990. URL : <https://doi.org/10.1063/1.3255990>.
- [32] Kristoffer CARLSSON et Fredrik EKRE. « Tensors.jl — Tensor Computations in Julia ». In : *Journal of Open Research Software* 7 (2019). DOI : 10.5334/jors.182. URL : <https://doi.org/10.5334/jors.182>.

- [33] Lorenz S. CEDERBAUM, Wolfgang DOMCKE et Jochen SCHIRMER. « Many-body theory of core holes ». en. In : *Physical Review A* 22.1 (juil. 1980), p. 206-222. ISSN : 0556-2791. DOI : 10.1103/PhysRevA.22.206. URL : <https://link.aps.org/doi/10.1103/PhysRevA.22.206> (visité le 14/06/2021).
- [34] Hyeon CHOI et al. « Photodissociation dynamics of the triiodide anion (I_3^-) ». In : *The Journal of Chemical Physics* 113.6 (août 2000), p. 2255-2262. DOI : 10.1063/1.482040. URL : <https://doi.org/10.1063/1.482040>.
- [35] Ove CHRISTIANSEN, Jürgen GAUSS et John F. STANTON. « Frequency-dependent polarizabilities and first hyperpolarizabilities of CO and H₂O from coupled cluster calculations ». In : *Chemical Physics Letters* 305.1-2 (mai 1999), p. 147-155. DOI : 10.1016/S0009-2614(99)00358-9. URL : [https://doi.org/10.1016/S0009-2614\(99\)00358-9](https://doi.org/10.1016/S0009-2614(99)00358-9).
- [36] Ove CHRISTIANSEN, Poul JØRGENSEN et Christof HÄTTIG. « Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy ». In : *International Journal of Quantum Chemistry* 68.1 (1998), p. 1-52. DOI : 10.1002/(sici)1097-461x(1998)68:1<1::aid-qua1>3.0.co;2-z. URL : [https://doi.org/10.1002/\(sici\)1097-461x\(1998\)68:1%3C1::aid-qua1%3E3.0.co;2-z](https://doi.org/10.1002/(sici)1097-461x(1998)68:1%3C1::aid-qua1%3E3.0.co;2-z).
- [37] Ove CHRISTIANSEN, Henrik KOCH et Poul JØRGENSEN. « The second-order approximate coupled cluster singles and doubles model CC2 ». In : *Chemical Physics Letters* 243.5-6 (sept. 1995), p. 409-418. DOI : 10.1016/0009-2614(95)00841-q. URL : [https://doi.org/10.1016/0009-2614\(95\)00841-q](https://doi.org/10.1016/0009-2614(95)00841-q).
- [38] Ove CHRISTIANSEN et al. « Integral-direct coupled cluster calculations of frequency-dependent polarizabilities, transition probabilities and excited-state properties ». In : *The Journal of Chemical Physics* 108.7 (fév. 1998), p. 2801-2816. DOI : 10.1063/1.475671. URL : <https://doi.org/10.1063/1.475671>.
- [39] CINES : Centre Informatique National de l'Enseignement Supérieur. URL : <https://www.cines.fr>.
- [40] Daniel B. COOK. *Handbook of Computational Quantum Chemistry*. Dover Books on Chemistry. Dover Publications, 2005. ISBN : 9780486443072. URL : <https://books.google.fr/books?id=Aa0qAwAAQBAJ>.
- [41] Sonia CORIANI et Henrik KOCH. « Communication : X-Ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework ». In : *The Journal of Chemical Physics* 143.18

- (nov. 2015), p. 181103. DOI : 10.1063/1.4935712. URL : <https://doi.org/10.1063/1.4935712>.
- [42] Marie-Emmanuelle COUPRIE. « New generation of light sources : Present and future ». In : *Journal of Electron Spectroscopy and Related Phenomena* 196 (oct. 2014), p. 3-13. DOI : 10.1016/j.eispec.2013.12.007. URL : <https://doi.org/10.1016/j.eispec.2013.12.007>.
- [43] T. Daniel CRAWFORD et Henry F. SCHAEFER III. « An Introduction to Coupled Cluster Theory for Computational Chemists ». In : *Reviews in Computational Chemistry*. John Wiley & Sons, Inc., jan. 2007, p. 33-136. DOI : 10.1002/9780470125915.ch2. URL : <https://doi.org/10.1002/9780470125915.ch2>.
- [44] Dieter CREMER, Wenli ZOU et Michael FILATOV. « Dirac-exact relativistic methods : the normalized elimination of the small component method ». In : *WIREs Computational Molecular Science* 4.5 (fév. 2014), p. 436-467. DOI : 10.1002/wcms.1181. URL : <https://doi.org/10.1002/wcms.1181>.
- [45] Ernest R. DAVIDSON. « The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices ». In : *Journal of Computational Physics* 17.1 (jan. 1975), p. 87-94. DOI : 10.1016/0021-9991(75)90065-0. URL : [https://doi.org/10.1016/0021-9991\(75\)90065-0](https://doi.org/10.1016/0021-9991(75)90065-0).
- [46] Robert G. DENNING. « Electronic Structure and Bonding in Actinyl Ions and their Analogs ». In : *The Journal of Physical Chemistry A* 111.20 (avr. 2007), p. 4125-4143. DOI : 10.1021/jp071061n. URL : <https://doi.org/10.1021/jp071061n>.
- [47] Paul A. M. DIRAC. « Quantised singularities in the electromagnetic field, » in : *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 133.821 (sept. 1931), p. 60-72. DOI : 10.1098/rspa.1931.0130. URL : <https://doi.org/10.1098/rspa.1931.0130>.
- [48] Paul A. M. DIRAC. « The quantum theory of the electron ». In : *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (fév. 1928), p. 610-624. DOI : 10.1098/rspa.1928.0023. URL : <https://doi.org/10.1098/rspa.1928.0023>.
- [49] Jean DOUCET. « Rayonnement synchrotron et applications ». In : (2011), p. 35.

- [50] Andreas DREUW et Michael WORMIT. « The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states ». In : *Wiley Interdisciplinary Reviews : Computational Molecular Science* 5.1 (sept. 2014), p. 82-95. DOI : 10.1002/wcms.1206. URL : <https://doi.org/10.1002/wcms.1206>.
- [51] Thom H. DUNNING. « Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen ». In : *The Journal of Chemical Physics* 90.2 (jan. 1989), p. 1007-1023. DOI : 10.1063/1.456153. URL : <https://doi.org/10.1063/1.456153>.
- [52] Rémi DUPUY et al. « Core level photoelectron spectroscopy of heterogeneous reactions at liquid–vapor interfaces : Current status, challenges, and prospects ». In : *The Journal of Chemical Physics* 154.6 (fév. 2021), p. 060901. DOI : 10.1063/5.0036178. URL : <https://doi.org/10.1063/5.0036178>.
- [53] Achintya Kumar DUTTA, Nayana VAVALA et Sourav PAL. *Assessment of Low Scaling Approximations to EOM-CCSD Method for Ionization Potential*. 2017. arXiv : 1708.01293 [physics.chem-ph].
- [54] Achintya Kumar DUTTA et al. « Partitioned EOMEA-MBPT(2) : An Efficient N5 Scaling Method for Calculation of Electron Affinities ». In : *Journal of Chemical Theory and Computation* 10.5 (avr. 2014), p. 1923-1933. DOI : 10.1021/ct4009409. URL : <https://doi.org/10.1021/ct4009409>.
- [55] Kenneth G. DYALL. « Relativistic Quadruple-Zeta and Revised Triple-Zeta and Double-Zeta Basis Sets for the 4p, 5p, and 6p Elements ». In : *Theoretical Chemistry Accounts* 115.5 (mar. 2006), p. 441-447. DOI : 10.1007/s00214-006-0126-0. URL : <https://doi.org/10.1007/s00214-006-0126-0>.
- [56] Kenneth G. DYALL et Knut FAEGRI Jr. *Introduction to Relativistic Quantum Chemistry*. Oxford University Press, 2007. ISBN : 978-0-19-514086-6. URL : <https://oxford.universitypressscholarship.com/view/10.1093/oso/9780195140866.001.0001/isbn-9780195140866>.
- [57] John H. D. ELAND et Raimund FEIFEL. « Triple ionization of HCl via states with a 2p core hole ». In : *The Journal of Chemical Physics* 151.11 (sept. 2019), p. 114301. DOI : 10.1063/1.5115552. URL : <https://doi.org/10.1063/1.5115552>.
- [58] Evgeny EPIFANOVSKY et al. « General implementation of the resolution-of-the-identity and Cholesky representations of electron repulsion integrals within coupled-cluster and equation-of-motion methods : Theory and benchmarks ». en. In : *The Journal of Chemical Physics* 139.13 (oct. 2013),

- p. 134105. ISSN : 0021-9606, 1089-7690. DOI : 10.1063/1.4820484. URL : <http://aip.scitation.org/doi/10.1063/1.4820484> (visité le 12/06/2021).
- [59] Susanna K. ERIKSSON et al. « Solvent Dependence of the Electronic Structure of I⁻ and I₃⁻ ». In : *The Journal of Physical Chemistry B* 118.11 (mar. 2014), p. 3164-3174. DOI : 10.1021/jp500533n. URL : <https://doi.org/10.1021/jp500533n>.
- [60] Charles S. FADLEY. « X-Ray photoelectron spectroscopy : Progress and perspectives ». In : *Journal of Electron Spectroscopy and Related Phenomena* 178-179 (mai 2010), p. 2-32. DOI : 10.1016/j.eispec.2010.01.006. URL : <https://doi.org/10.1016/j.eispec.2010.01.006>.
- [61] Xiaolong FAN et al. « Atmospheric gaseous hydrochloric and hydrobromic acid in urban Beijing, China : detection, source identification and potential atmospheric impacts ». In : *Atmospheric Chemistry and Physics* 21.14 (2021), p. 11437-11452. DOI : 10.5194/acp-21-11437-2021. URL : <https://acp.copernicus.org/articles/21/11437/2021/>.
- [62] Michael J. FLYNN. « Some Computer Organizations and Their Effectiveness ». In : *IEEE Transactions on Computers* C-21.9 (sept. 1972), p. 948-960. DOI : 10.1109/tc.1972.5009071. URL : <https://doi.org/10.1109/tc.1972.5009071>.
- [63] Camille FORTIN. « Etudes par simulations numériques et moléculaires de la réactivité atmosphérique de l'iode ». 2019LILUR053. Thèse de doct. 2019. URL : <http://www.theses.fr/2019LILUR053/document>.
- [64] Adam E. A. FOU DA et Nicholas A. BESLEY. « Assessment of basis sets for density functional theory-based calculations of core-electron spectroscopies ». In : *Theoretical Chemistry Accounts* 137.1 (déc. 2017). DOI : 10.1007/s00214-017-2181-0. URL : <https://doi.org/10.1007/s00214-017-2181-0>.
- [65] W. H. FURRY. « On Bound States and Scattering in Positron Theory ». In : *Physical Review* 81.1 (jan. 1951), p. 115-124. DOI : 10.1103/physrev.81.115. URL : <https://doi.org/10.1103/physrev.81.115>.
- [66] Jürgen GAUSS, John F. STANTON et Rodney J. BARTLETT. « Coupled-cluster open-shell analytic gradients : Implementation of the direct product decomposition approach in energy gradient calculations ». en. In : *The Journal of Chemical Physics* 95.4 (août 1991), p. 2623-2638. ISSN : 0021-9606, 1089-7690. DOI : 10.1063/1.460915. URL : <http://aip.scitation.org/doi/10.1063/1.460915> (visité le 06/06/2018).

- [67] Jan GEERTSEN, Magnus RITBY et Rodney J. BARTLETT. « The equation-of-motion coupled-cluster method : Excitation energies of Be and CO ». In : *Chemical Physics Letters* 164.1 (déc. 1989), p. 57-62. DOI : 10.1016/0009-2614(89)85202-9. URL : [https://doi.org/10.1016/0009-2614\(89\)85202-9](https://doi.org/10.1016/0009-2614(89)85202-9).
- [68] GENCI : *Grand Équipement National de Calcul Intensif*. URL : <https://www.genci.fr/fr>.
- [69] Mary K. GILLES, Mark L. POLAK et W. C. LINEBERGER. « Photoelectron spectroscopy of the halogen oxide anions FO^- , ClO^- , BrO^- , IO^- , OClO^- , and OIO^- ». In : *The Journal of Chemical Physics* 96.11 (juin 1992), p. 8012-8020. DOI : 10.1063/1.462352. URL : <https://doi.org/10.1063/1.462352>.
- [70] Joshua J. GOINGS et al. « Assessment of low-scaling approximations to the equation of motion coupled-cluster singles and doubles equations ». In : *The Journal of Chemical Physics* 141.16 (oct. 2014), p. 164116. DOI : 10.1063/1.4898709. URL : <https://doi.org/10.1063/1.4898709>.
- [71] Victor GOLD, éd. *The IUPAC Compendium of Chemical Terminology*. International Union of Pure et Applied Chemistry (IUPAC), 2019. DOI : 10.1351/goldbook. URL : <https://doi.org/10.1351/goldbook>.
- [72] André Severo Pereira GOMES et al. « The electronic structure of the triiodide ion from relativistic correlated calculations : A comparison of different methodologies ». In : *The Journal of Chemical Physics* 133.6 (août 2010), p. 064305. DOI : 10.1063/1.3474571. URL : <https://doi.org/10.1063/1.3474571>.
- [73] André Severo Pereira GOMES et al. « Towards systematically improvable models for actinides in condensed phase : the electronic spectrum of uranyl in $\text{Cs}_2\text{UO}_2\text{Cl}_4$ as a test case ». In : *Physical Chemistry Chemical Physics* 15.36 (2013), p. 15153. DOI : 10.1039/c3cp52090k. URL : <https://doi.org/10.1039/c3cp52090k>.
- [74] Grzegorz GRECZYNSKI et Lars HULTMAN. « X-Ray photoelectron spectroscopy : Towards reliable binding energy referencing ». In : *Progress in Materials Science* 107 (jan. 2020), p. 100591. DOI : 10.1016/j.pmatsci.2019.100591. URL : <https://doi.org/10.1016/j.pmatsci.2019.100591>.
- [75] Gilbert GRELL et al. « Multi-reference approach to the calculation of photoelectron spectra including spin-orbit coupling ». In : *The Journal of Chemical Physics* 143.7 (août 2015), p. 074104. DOI : 10.1063/1.4928511. URL : <https://doi.org/10.1063/1.4928511>.

- [76] Steven R. GWALTNEY, Marcel NOOIJEN et Rodney J. BARTLETT. « Simplified methods for equation-of-motion coupled-cluster excited state calculations ». In : *Chemical Physics Letters* 248.3-4 (jan. 1996), p. 189-198. DOI : 10.1016/0009-2614(95)01329-6. URL : [https://doi.org/10.1016/0009-2614\(95\)01329-6](https://doi.org/10.1016/0009-2614(95)01329-6).
- [77] Matthieu HAEFELE. *FPGA : un nouveau type d'accélérateur pour le calcul?* Meeting : Café Calcul. Juin 2021. URL : https://lma-umr5142.univ-pau.fr/fr/organisation/membres/cv_-mhaefele-fr.html.
- [78] Loïc HALBERT et al. « Investigating solvent effects on the magnetic properties of molybdate ions (MoO_4^{2-}) with relativistic embedding ». In : *International Journal of Quantum Chemistry* 120.21 (mar. 2020). DOI : 10.1002/qua.26207. URL : <https://doi.org/10.1002/qua.26207>.
- [79] Loïc HALBERT et al. « Relativistic EOM-CCSD for Core-Excited and Core-Ionized State Energies Based on the Four-Component Dirac-Coulomb(-Gaunt) Hamiltonian ». In : *Journal of Chemical Theory and Computation* 17.6 (mai 2021), p. 3583-3598. DOI : 10.1021/acs.jctc.0c01203. URL : <https://doi.org/10.1021/acs.jctc.0c01203>.
- [80] Michael HANRATH. « A simple and operational test for external connectivity of tensors in many-body methods ». In : *Theoretical Chemistry Accounts* 137.6 (juin 2018). DOI : 10.1007/s00214-018-2265-5. URL : <https://doi.org/10.1007/s00214-018-2265-5>.
- [81] William HAYES et Frederick C. BROWN. « Absorption by Some Molecular Gases in the Extreme Ultraviolet ». In : *Physical Review A* 6.1 (juil. 1972), p. 21-30. DOI : 10.1103/physreva.6.21. URL : <https://doi.org/10.1103/physreva.6.21>.
- [82] Martin HEAD-GORDON, Manabu OUMI et David MAURICE. « Quasidegenerate second-order perturbation corrections to single-excitation configuration interaction ». In : *Molecular Physics* 96.4 (fév. 1999), p. 593-602. DOI : 10.1080/00268979909482996. URL : <https://doi.org/10.1080/00268979909482996>.
- [83] Martin HEAD-GORDON et al. « A doubles correction to electronic excited states from configuration interaction in the space of single substitutions ». In : *Chemical Physics Letters* 219.1-2 (mar. 1994), p. 21-29. DOI : 10.1016/0009-2614(94)00070-0. URL : [https://doi.org/10.1016/0009-2614\(94\)00070-0](https://doi.org/10.1016/0009-2614(94)00070-0).

- [84] Trygve HELGAKER, Poul JØRGENSEN et Jeppe OLSEN. *Molecular Electronic-Structure Theory*. English. OCLC : 904835967. 2014. ISBN : 9781119019558 9781119019565. URL : <https://www.wiley.com/en-us/Molecular+Electronic+Structure+Theory-p-9780471967552> (visité le 27/08/2021).
- [85] Bernd A. HESS et al. « A mean-field spin-orbit method applicable to correlated wavefunctions ». In : *Chemical Physics Letters* 251.5-6 (mar. 1996), p. 365-371. DOI : 10.1016/0009-2614(96)00119-4. URL : [https://doi.org/10.1016/0009-2614\(96\)00119-4](https://doi.org/10.1016/0009-2614(96)00119-4).
- [86] Kimihiko HIRAO et Hiroshi NAKATSUJI. « A generalization of the Davidson's method to large nonsymmetric eigenvalue problems ». In : *Journal of Computational Physics* 45.2 (fév. 1982), p. 246-254. DOI : 10.1016/0021-9991(82)90119-x. URL : [https://doi.org/10.1016/0021-9991\(82\)90119-x](https://doi.org/10.1016/0021-9991(82)90119-x).
- [87] Sebastian HÖFENER et Wim KLOPPER. « Natural transition orbitals for the calculation of correlation and excitation energies ». In : *Chemical Physics Letters* 679 (juil. 2017), p. 52-59. DOI : 10.1016/j.cpllett.2017.04.083. URL : <https://doi.org/10.1016/j.cpllett.2017.04.083>.
- [88] *HPE SGI : Hewlett Packard Entreprise*. Liquid cooled, tray-based, high-density clustered computer system. URL : <https://buy.hpe.com/fr/fr/servers/sgi-system/sgi-8600-system/sgi-8600-system/hpe-sgi-8600-system/p/1010032504>.
- [89] Stefan HÜFNER. *Photoelectron Spectroscopy*. Springer Berlin Heidelberg, 2003. DOI : 10.1007/978-3-662-09280-4. URL : <https://doi.org/10.1007/978-3-662-09280-4>.
- [90] *IDRIS : Institut du développement et des ressources en informatique scientifique - http://www.idris.fr*. URL : <http://www.idris.fr>.
- [91] Miroslav ILIAŠ et Trond SAUE. « An infinite-order two-component relativistic Hamiltonian by a simple one-step transformation ». In : *The Journal of Chemical Physics* 126.6 (fév. 2007), p. 064102. DOI : 10.1063/1.2436882. URL : <https://doi.org/10.1063/1.2436882>.
- [92] *Intel Xeon W-3265M : Central Processing Unit*. URL : <https://www.intel.fr/content/www/fr/fr/products/sku/193746/intel-xeon-w3265-processor-33m-cache-2-70-ghz/specifications.html>.
- [93] Róbert IZSÁK. « Single-reference coupled cluster methods for computing excitation energies in large molecules : The efficiency and accuracy of approximations ». In : *WIREs Computational Molecular Science* 10.3 (sept. 2019). DOI : 10.1002/wcms.1445. URL : <https://doi.org/10.1002/wcms.1445>.

- [94] James F. JANAK. « Proof that $\partial E/\partial n_i = \epsilon_i$ in density-functional theory ». In : *Physical Review B* 18.12 (déc. 1978), p. 7165-7168. DOI : 10.1103/physrevb.18.7165. URL : <https://doi.org/10.1103/physrevb.18.7165>.
- [95] Jean Zay HPC. URL : <http://www.idris.fr/jean-zay/jean-zay-presentation.html>.
- [96] Frank JENSEN. *Introduction to computational chemistry*. 2nd ed. John Wiley & Sons, 2007. ISBN : 978-0-470-01186-7.
- [97] James Prescott JOULE. « XXXVIII. On the heat evolved by metallic conductors of electricity, and in the cells of a battery during electrolysis ». In : *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 19.124 (oct. 1841), p. 260-277. DOI : 10.1080/14786444108650416. URL : <https://doi.org/10.1080/14786444108650416>.
- [98] JUTHO et al. *Jutho/TensorOperations.jl : v1.1.0*. 2019. DOI : 10.5281/ZENODO.3245497. URL : <https://zenodo.org/record/3245497>.
- [99] Mihály KÁLLAY et al. « The MRCC program system : Accurate quantum chemistry from water to proteins ». In : *The Journal of Chemical Physics* 152.7 (fév. 2020), p. 074107. DOI : 10.1063/1.5142048. URL : <https://doi.org/10.1063/1.5142048>.
- [100] Rick A. KENDALL, Thom H. DUNNING et Robert J. HARRISON. « Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions ». In : *The Journal of Chemical Physics* 96.9 (mai 1992), p. 6796-6806. DOI : 10.1063/1.462569. URL : <https://doi.org/10.1063/1.462569>.
- [101] Rika KOBAYASHI, Henrik KOCH et Poul JØRGENSEN. « Calculation of frequency-dependent polarizabilities using coupled-cluster response theory ». In : *Chemical Physics Letters* 219.1-2 (mar. 1994), p. 30-35. DOI : 10.1016/0009-2614(94)00051-4. URL : [https://doi.org/10.1016/0009-2614\(94\)00051-4](https://doi.org/10.1016/0009-2614(94)00051-4).
- [102] Henrik KOCH et Poul JØRGENSEN. « Coupled cluster response functions ». In : *The Journal of Chemical Physics* 93.5 (sept. 1990), p. 3333-3344. DOI : 10.1063/1.458814. URL : <https://doi.org/10.1063/1.458814>.
- [103] Henrik KOCH, Alfredo SÁNCHEZ DE MERÁS et Thomas B. PEDERSEN. « Reduced scaling in electronic structure calculations using Cholesky decompositions ». en. In : *The Journal of Chemical Physics* 118.21 (juin 2003), p. 9481-9484. ISSN : 0021-9606, 1089-7690. DOI : 10.1063/1.1578621. URL : <http://aip.scitation.org/doi/10.1063/1.1578621> (visité le 11/06/2021).

- [104] Henrik KOCH et al. « Calculation of size-intensive transition moments from the coupled cluster singles and doubles linear response function ». In : *The Journal of Chemical Physics* 100.6 (mar. 1994), p. 4393-4400. DOI : 10.1063/1.466321. URL : <https://doi.org/10.1063/1.466321>.
- [105] Henrik KOCH et al. « Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, CH⁺, CO, and H₂O ». In : *The Journal of Chemical Physics* 93.5 (sept. 1990), p. 3345-3350. DOI : 10.1063/1.458815. URL : <https://doi.org/10.1063/1.458815>.
- [106] Xiangrui KONG et al. « Coexistence of Physisorbed and Solvated HCl at Warm Ice Surfaces ». In : *The Journal of Physical Chemistry Letters* 8.19 (sept. 2017), p. 4757-4762. DOI : 10.1021/acs.jpcllett.7b01573. URL : <https://doi.org/10.1021/acs.jpcllett.7b01573>.
- [107] Karol KOZIOŁ et Gustavo A. AUCAR. « QED effects on individual atomic orbital energies ». In : *The Journal of Chemical Physics* 148.13 (avr. 2018), p. 134101. DOI : 10.1063/1.5026193. URL : <https://doi.org/10.1063/1.5026193>.
- [108] Hendrik A. KRAMERS. In : *Proc. Acad. Amsterdam* 33 (1930).
- [109] Anna I. KRYLOV. « Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species : The Hitchhiker's Guide to Fock Space ». en. In : *Annual Review of Physical Chemistry* 59.1 (mai 2008), p. 433-462. ISSN : 0066-426X, 1545-1593. DOI : 10.1146/annurev.physchem.59.032607.093602. URL : <http://www.annualreviews.org/doi/10.1146/annurev.physchem.59.032607.093602> (visité le 13/06/2021).
- [110] Jon K. LAERDAHL, Trond SAUE et Knut FAEGRI Jr. « Direct relativistic MP2 : properties of ground state CuF, AgF and AuF ». In : *Theoretical Chemistry Accounts : Theory, Computation, and Modeling (Theoretica Chimica Acta)* 97.1-4 (oct. 1997), p. 177-184. DOI : 10.1007/s002140050251. URL : <https://doi.org/10.1007/s002140050251>.
- [111] Cornelius LANZOS. « An iteration method for the solution of the eigenvalue problem of linear differential and integral operators ». In : *Journal of Research of the National Bureau of Standards* 45.4 (oct. 1950), p. 255. DOI : 10.6028/jres.045.026. URL : <https://doi.org/10.6028/jres.045.026>.

- [112] Arie LANDAU et al. « Frozen natural orbitals for ionized states within equation-of-motion coupled-cluster formalism ». en. In : *The Journal of Chemical Physics* 132.1 (jan. 2010), p. 014109. ISSN : 0021-9606, 1089-7690. DOI : 10.1063/1.3276630. URL : <http://aip.scitation.org/doi/10.1063/1.3276630> (visité le 12/06/2021).
- [113] Peter W. LANGHOFF, S. T. EPSTEIN et M. KARPLUS. « Aspects of Time-Dependent Perturbation Theory ». In : *Reviews of Modern Physics* 44.3 (juil. 1972), p. 602-644. ISSN : 0034-6861. DOI : 10.1103/RevModPhys.44.602.
- [114] Timothy J. LEE et al. « Theoretical investigations of molecules composed only of fluorine, oxygen and nitrogen : determination of the equilibrium structures of FOOF, (NO)₂ and FNNF and the transition state structure for FNNF cis-trans isomerization ». In : *Theoretica Chimica Acta* 75.2 (1989), p. 81-98. DOI : 10.1007/bf00527711. URL : <https://doi.org/10.1007/bf00527711>.
- [115] Vito LIBRANDO et Alparone ANDREA. « Electronic Polarizability as a Predictor of Biodegradation Rates of Dimethylnaphthalenes. An Ab Initio and Density Functional Theory Study ». In : *Environmental Science & Technology* 41.5 (2007). PMID : 17396655, p. 1646-1652. DOI : 10.1021/es061632+. eprint : <https://doi.org/10.1021/es061632+>. URL : <https://doi.org/10.1021/es061632+>.
- [116] Hans LISCHKA et al. « Multireference Approaches for Excited States of Molecules ». In : *Chemical Reviews* 118.15 (juil. 2018), p. 7293-7361. DOI : 10.1021/acs.chemrev.8b00244. URL : <https://doi.org/10.1021/acs.chemrev.8b00244>.
- [117] Ya-Jun LIU et al. « Spin-Orbit Ab Initio Investigation of the Ultraviolet Photolysis of Diiodomethane ». In : *ChemPhysChem* 8.6 (avr. 2007), p. 890-898. DOI : 10.1002/cphc.200600737. URL : <https://doi.org/10.1002/cphc.200600737>.
- [118] Lu LIU. « Exploration of astatine chemistry in solution : focus on the Pourbaix diagram in noncomplexing medium and characterization of astatine-mediated halogen bonds ». Theses. Ecole nationale supérieure Mines-Télécom Atlantique, oct. 2020. URL : <https://tel.archives-ouvertes.fr/tel-03123005>.
- [119] Wenjian LIU. « Big picture of relativistic molecular quantum mechanics ». In : *National Science Review* 3.2 (déc. 2015), p. 204-221. DOI : 10.1093/nsr/nwv081. URL : <https://doi.org/10.1093/nsr/nwv081>.

- [120] Wenjian LIU. « Essentials of relativistic quantum chemistry ». In : *The Journal of Chemical Physics* 152.18 (mai 2020), p. 180901. DOI : 10.1063/5.0008432. URL : <https://doi.org/10.1063/5.0008432>.
- [121] Wenjian LIU et Ingvar LINDGREN. « Going beyond “no-pair relativistic quantum chemistry” ». In : *The Journal of Chemical Physics* 139.1 (juil. 2013), p. 014108. DOI : 10.1063/1.4811795. URL : <https://doi.org/10.1063/1.4811795>.
- [122] Wenjian LIU et Daoling PENG. « Exact two-component Hamiltonians revisited ». In : *The Journal of Chemical Physics* 131.3 (juil. 2009), p. 031104. DOI : 10.1063/1.3159445. URL : <https://doi.org/10.1063/1.3159445>.
- [123] Pierre-François LOOS, Anthony SCEMAMA et Denis JACQUEMIN. « The Quest for Highly Accurate Excitation Energies : A Computational Perspective ». In : *The Journal of Physical Chemistry Letters* 11.6 (mar. 2020), p. 2374-2383. DOI : 10.1021/acs.jpcllett.0c00014. URL : <https://doi.org/10.1021/acs.jpcllett.0c00014>.
- [124] Per-Olov LÖWDIN. « Correlation Problem in Many-Electron Quantum Mechanics I. Review of Different Approaches and Discussion of Some Current Ideas ». In : *Advances in Chemical Physics*. John Wiley & Sons, Inc., mar. 2007, p. 207-322. DOI : 10.1002/9780470143483.ch7. URL : <https://doi.org/10.1002/9780470143483.ch7>.
- [125] Per-Olov LÖWDIN. « Studies in perturbation theory ». In : *Journal of Molecular Spectroscopy* 10.1-6 (jan. 1963), p. 12-33. DOI : 10.1016/0022-2852(63)90151-6. URL : [https://doi.org/10.1016/0022-2852\(63\)90151-6](https://doi.org/10.1016/0022-2852(63)90151-6).
- [126] Marcus LUNDBERG et Mickaël G. DELCEY. « Multiconfigurational Approach to X-Ray Spectroscopy of Transition Metal Complexes ». In : *Transition Metals in Coordination Environments*. Springer International Publishing, 2019, p. 185-217. DOI : 10.1007/978-3-030-11714-6_7. URL : https://doi.org/10.1007/978-3-030-11714-6_7.
- [127] Lixiang LUO et al. « Pre-exascale accelerated application development : The ORNL Summit experience ». In : *IBM Journal of Research and Development* 64.3/4 (mai 2020), 11 :1-11 :21. DOI : 10.1147/jrd.2020.2965881. URL : <https://doi.org/10.1147/jrd.2020.2965881>.
- [128] Dmitry I. LYAKH. « Algebraic connectivity analysis in molecular electronic structure theory I : coulomb potential, tensor connectivity, ε -approximation ». In : *Molecular Physics* 110.14 (juil. 2012), p. 1469-1492. DOI : 10.1080/00268976.2012.663509. URL : <https://doi.org/10.1080/00268976.2012.663509>.

- [129] Dmitry I. LYAKH. « Domain-specific virtual processors as a portable programming and execution model for parallel computational workloads on modern heterogeneous high-performance computing architectures ». In : *International Journal of Quantum Chemistry* 119.12 (mar. 2019), e25926. DOI : 10.1002/qua.25926. URL : <https://doi.org/10.1002/qua.25926>.
- [130] Dmitry I. LYAKH. *TAL-SH : Numerical tensor algebra library for shared-memory heterogeneous nodes equipped with multicore CPU and NVIDIA GPU*. URL : https://github.com/DmitryLyakh/TAL_SH.
- [131] Dimitrios MAGANAS et al. « Comparison of multireference ab initio wavefunction methodologies for X-Ray absorption edges : A case study on $[\text{Fe(II/III)Cl}_4]^{2-/1-}$ molecules ». In : *The Journal of Chemical Physics* 150.10 (mar. 2019), p. 104106. DOI : 10.1063/1.5051613. URL : <https://doi.org/10.1063/1.5051613>.
- [132] George MAROULIS, éd. *Atoms, molecules and clusters in electric fields : theoretical approaches to the calculation of electric polarizability*. Series in computational, numerical and mathematical methods in sciences and engineering v. 1. OCLC : ocm75423356. London : Hackensack, NJ : Imperial College Press ; distributed by World Scientific Publishing, 2006. ISBN : 9781860946769.
- [133] Larry E. McMURCHIE et Ernest R. DAVIDSON. « One- and two-electron integrals over cartesian gaussian functions ». In : *Journal of Computational Physics* 26.2 (fév. 1978), p. 218-231. DOI : 10.1016/0021-9991(78)90092-x. URL : [https://doi.org/10.1016/0021-9991\(78\)90092-x](https://doi.org/10.1016/0021-9991(78)90092-x).
- [134] Alberto MODELLI et al. « Electron attachment to the fluoro-, bromo-, and iodomethanes studied by means of electron transmission spectroscopy and $X\alpha$ calculations ». In : *The Journal of Chemical Physics* 96.3 (fév. 1992), p. 2061-2070. DOI : 10.1063/1.462058. URL : <https://doi.org/10.1063/1.462058>.
- [135] Christian MØLLER et Milton S. PLESSET. « Note on an Approximation Treatment for Many-Electron Systems ». In : *Physical Review* 46.7 (oct. 1934), p. 618-622. DOI : 10.1103/physrev.46.618. URL : <https://doi.org/10.1103/physrev.46.618>.
- [136] John F. MOULDER et Jill CHASTAIN. *Handbook of X-Ray Photoelectron Spectroscopy : A Reference Book of Standard Spectra for Identification and Interpretation of XPS Data*. Physical Electronics Division, Perkin-Elmer Corporation, 1992. ISBN : 9780962702624. URL : https://books.google.fr/books?id=A%5C_XGQgAACAAJ.

- [137] Arnaldo NAVES DE BRITO et al. « A theoretical study of X-Ray photoelectron spectra of model molecules for polymethylmethacrylate ». In : *The Journal of Chemical Physics* 95.4 (août 1991), p. 2965-2974. DOI : 10.1063/1.460898. URL : <https://doi.org/10.1063/1.460898>.
- [138] Simon P. NEVILLE et al. « Beyond structure : ultrafast X-Ray absorption spectroscopy as a probe of non-adiabatic wavepacket dynamics ». In : *Faraday Discussions* 194 (2016), p. 117-145. DOI : 10.1039/c6fd00117c. URL : <https://doi.org/10.1039/c6fd00117c>.
- [139] Marcel NOOIJEN, Kodagenahalli R. SHAMASUNDAR et Debashis MUKHERJEE. « Reflections on size-extensivity, size-consistency and generalized extensivity in many-body theory ». In : *Molecular Physics* 103.15-16 (août 2005), p. 2277-2298. DOI : 10.1080/00268970500083952. URL : <https://doi.org/10.1080/00268970500083952>.
- [140] *NVidia Pascal V100 SXM2 : Graphics Processing Unit*. URL : <https://developer.nvidia.com/hpc-application-performance>.
- [141] Michael K. ODLING-SMEE et al. « Observation of post-collision interaction (PCI) in HBr using two-dimensional photoelectron spectroscopy ». In : *Journal of Physics B : Atomic, Molecular and Optical Physics* 32.11 (jan. 1999), p. 2529-2537. DOI : 10.1088/0953-4075/32/11/305. URL : <https://doi.org/10.1088/0953-4075/32/11/305>.
- [142] *OLCF : Oak Ridge Leadership Computing Facility*. URL : <https://www.olcf.ornl.gov>.
- [143] Małgorzata OLEJNICZAK, Radovan BAST et André Severo Pereira GOMES. « On the calculation of second-order magnetic properties using subsystem approaches in a relativistic framework ». In : *Physical Chemistry Chemical Physics* 19.12 (2017), p. 8400-8415. DOI : 10.1039/c6cp08561j. URL : <https://doi.org/10.1039/c6cp08561j>.
- [144] Jeppe OLSEN et al. « Surprising cases of divergent behavior in Møller-Plesset perturbation theory ». In : *The Journal of Chemical Physics* 105.12 (sept. 1996), p. 5082-5090. DOI : 10.1063/1.472352. URL : <https://doi.org/10.1063/1.472352>.
- [145] Masaki OURA et al. « Hard X-Ray photoelectron spectroscopy on heavy atoms and heavy-element containing molecules using synchrotron radiation up to 35 keV at SPring-8 undulator beamlines ». In : *New Journal of Physics* 21.4 (avr. 2019), p. 043015. DOI : 10.1088/1367-2630/ab09a3. URL : <https://doi.org/10.1088/1367-2630/ab09a3>.

- [146] Josef PALDUS et Xiangzhu LI. « Calculation of static molecular properties in the framework of the unitary group based coupled cluster approach ». In : *Canadian Journal of Chemistry* 74.6 (juin 1996), p. 918-930. DOI : 10.1139/v96-101. URL : <https://doi.org/10.1139/v96-101>.
- [147] Filip PAWŁOWSKI, Jeppe OLSEN et Poul JØRGENSEN. « Molecular response properties from a Hermitian eigenvalue equation for a time-periodic Hamiltonian ». In : *The Journal of Chemical Physics* 142.11 (mar. 2015), p. 114109. DOI : 10.1063/1.4913364. URL : <https://doi.org/10.1063/1.4913364>.
- [148] Eetu PELIMANNI et al. « Core and Valence Level Photoelectron Spectroscopy of Nanosolvated KCl ». In : *The Journal of Physical Chemistry A* (mai 2021). DOI : 10.1021/acs.jpca.1c01539. URL : <https://doi.org/10.1021/acs.jpca.1c01539>.
- [149] Kristine PIERLOOT et al. « Electronic spectrum of UO_2^{2+} and $[\text{UO}_2\text{Cl}_4]^{2-}$ calculated with time-dependent density functional theory ». In : *The Journal of Chemical Physics* 126.19 (mai 2007), p. 194311. DOI : 10.1063/1.2735297. URL : <https://doi.org/10.1063/1.2735297>.
- [150] Johann V. POTOTSCHNIG et al. « Implementation of Relativistic Coupled Cluster Theory for Massively Parallel GPU-Accelerated Computing Architectures ». In : *Journal of Chemical Theory and Computation* (août 2021). DOI : 10.1021/acs.jctc.1c00260. URL : <https://doi.org/10.1021/acs.jctc.1c00260>.
- [151] Christos PSARRAS et al. « The landscape of software for tensor computations ». en. In : *arXiv :2103.13756 [cs]* (mai 2021). arXiv : 2103.13756. URL : <http://arxiv.org/abs/2103.13756> (visité le 21/06/2021).
- [152] Noèlia PUEYO BELLAFONT, Paul S. BAGUS et Francesc ILLAS. « Prediction of core level binding energies in density functional theory : Rigorous definition of initial and final state contributions and implications on the physical meaning of Kohn-Sham energies ». In : *The Journal of Chemical Physics* 142.21 (juin 2015), p. 214102. DOI : 10.1063/1.4921823. URL : <https://doi.org/10.1063/1.4921823>.
- [153] Péter PULAY. « Convergence acceleration of iterative sequences. the case of scf iteration ». In : *Chemical Physics Letters* 73.2 (juil. 1980), p. 393-398. DOI : 10.1016/0009-2614(80)80396-4. URL : [https://doi.org/10.1016/0009-2614\(80\)80396-4](https://doi.org/10.1016/0009-2614(80)80396-4).

- [154] Pekka PYYKKÖ. « Relativistic effects in structural chemistry ». en. In : *Chemical Reviews* 88.3 (mai 1988), p. 563-594. ISSN : 0009-2665, 1520-6890. DOI : 10.1021/cr00085a006. URL : <https://pubs.acs.org/doi/abs/10.1021/cr00085a006> (visité le 06/05/2021).
- [155] Pekka PYYKKÖ. « The Physics behind Chemistry and the Periodic Table ». In : *Chemical Reviews* 112.1 (juil. 2011), p. 371-384. DOI : 10.1021/cr200042e. URL : <https://doi.org/10.1021/cr200042e>.
- [156] Pekka PYYKKÖ et Jean Paul DESCLAUX. « Relativity and the periodic system of elements ». In : *Accounts of Chemical Research* 12.8 (août 1979), p. 276-281. DOI : 10.1021/ar50140a002. URL : <https://doi.org/10.1021/ar50140a002>.
- [157] Krishnan RAGHAVACHARI et al. « Reprint of : A fifth-order perturbation comparison of electron correlation theories ». In : *Chemical Physics Letters* 589 (déc. 2013), p. 37-40. DOI : 10.1016/j.cplett.2013.08.064. URL : <https://doi.org/10.1016/j.cplett.2013.08.064>.
- [158] Florent RÉAL et al. « Quantum Chemical and Molecular Dynamics Study of the Coordination of Th(IV) in Aqueous Solvent ». In : *The Journal of Physical Chemistry B* 114.48 (nov. 2010), p. 15913-15924. DOI : 10.1021/jp108061s. URL : <https://doi.org/10.1021/jp108061s>.
- [159] Florent RÉAL et al. « 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⁻? » In : *The Journal of Chemical Physics* 144.12 (mar. 2016), p. 124513. DOI : 10.1063/1.4944613. URL : <https://doi.org/10.1063/1.4944613>.
- [160] Florent RÉAL et al. « Theoretical investigation of the energies and geometries of photoexcited uranyl(VI) ion : A comparison between wavefunction theory and density functional theory ». In : *The Journal of Chemical Physics* 127.21 (déc. 2007), p. 214302. DOI : 10.1063/1.2814157. URL : <https://doi.org/10.1063/1.2814157>.
- [161] Antoine ROOSE et al. « Classical Molecular Dynamics Study of Small-Chain Carboxylic Acid Aerosol Particles ». In : *ACS Earth and Space Chemistry* 3.3 (fév. 2019), p. 380-389. DOI : 10.1021/acsearthspacechem.8b00172. URL : <https://doi.org/10.1021/acsearthspacechem.8b00172>.
- [162] Trond SAUE. « Chapter 7 Post Dirac-Hartree-Fock methods—properties ». In : *Theoretical and Computational Chemistry*. Elsevier, 2002, p. 332-400. DOI : 10.1016/s1380-7323(02)80033-4. URL : [https://doi.org/10.1016/s1380-7323\(02\)80033-4](https://doi.org/10.1016/s1380-7323(02)80033-4).

- [163] Trond SAUE. « Relativistic Hamiltonians for Chemistry : A Primer ». In : *ChemPhysChem* 12.17 (nov. 2011), p. 3077-3094. DOI : 10.1002/cphc.201100682. URL : <https://doi.org/10.1002/cphc.201100682>.
- [164] Trond SAUE et al. « Principles of direct 4-component relativistic SCF : application to caesium auride ». In : *Molecular Physics* 91.5 (1997), p. 937-950. DOI : 10.1080/002689797171058. URL : <https://www.tandfonline.com/doi/abs/10.1080/002689797171058>.
- [165] Trond SAUE et al. « The DIRAC code for relativistic molecular calculations ». In : *The Journal of Chemical Physics* 152.20 (mai 2020), p. 204104. DOI : 10.1063/5.0004844. URL : <https://doi.org/10.1063/5.0004844>.
- [166] Bernd SCHIMMELPFENNIG. *AMFI, an atomic mean-field spin-orbit integral program*. 1996.
- [167] Jochen SCHIRMER. « Beyond the random-phase approximation : A new approximation scheme for the polarization propagator ». en. In : *Physical Review A* 26.5 (nov. 1982), p. 2395-2416. ISSN : 0556-2791. DOI : 10.1103/PhysRevA.26.2395. URL : <https://link.aps.org/doi/10.1103/PhysRevA.26.2395> (visité le 11/06/2021).
- [168] Jochen SCHIRMER et Alexander THIEL. « An intermediate state representation approach to K-shell ionization in molecules. I. Theory ». In : *The Journal of Chemical Physics* 115.23 (déc. 2001), p. 10621-10635. DOI : 10.1063/1.1418437. URL : <https://doi.org/10.1063/1.1418437>.
- [169] Peter SCHWERDTFEGER. « Relativistic effects in properties of gold ». In : *Heteroatom Chemistry* 13.6 (2002), p. 578-584. DOI : 10.1002/hc.10093. URL : <https://doi.org/10.1002/hc.10093>.
- [170] Peter SCHWERDTFEGER et Jeffrey K. NAGLE. « 2018 Table of static dipole polarizabilities of the neutral elements in the periodic table ». In : *Molecular Physics* 117.9-12 (2019), p. 1200-1225. DOI : 10.1080/00268976.2018.1535143. eprint : <https://doi.org/10.1080/00268976.2018.1535143>. URL : <https://doi.org/10.1080/00268976.2018.1535143>.
- [171] *Scipy Representation of a kernel density estimate using Gaussian kernels*. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.gaussian_kde.html. Accessed : 2025-08-15.
- [172] Martin P. SEAH. « Summary of ISO/TC 201 Standard : VII ISO 15472 : 2001, Surface chemical analysis, X-Ray photoelectron spectrometers, Calibration of energy scales ». In : *Surface and Interface Analysis* 31.8 (2001), p. 721-723. DOI : 10.1002/sia.1076. URL : <https://doi.org/10.1002/sia.1076>.

- [173] Martin P. SEAH et W. A. DENCH. « Quantitative electron spectroscopy of surfaces : A standard data base for electron inelastic mean free paths in solids ». In : *Surface and Interface Analysis* 1.1 (fév. 1979), p. 2-11. DOI : 10.1002/sia.740010103. URL : <https://doi.org/10.1002/sia.740010103>.
- [174] Martin P. SEAH, I. S. GILMORE et G. BEAMSON. « XPS : binding energy calibration of electron spectrometers 5-re-evaluation of the reference energies ». In : *Surface and Interface Analysis* 26.9 (août 1998), p. 642-649. DOI : 10.1002/(sici)1096-9918(199808)26:9<642::aid-sia408>3.0.co;2-3. URL : [https://doi.org/10.1002/\(sici\)1096-9918\(199808\)26:9%3C642::aid-sia408%3E3.0.co;2-3](https://doi.org/10.1002/(sici)1096-9918(199808)26:9%3C642::aid-sia408%3E3.0.co;2-3).
- [175] Yihan SHAO et al. « Advances in molecular quantum chemistry contained in the Q-Chem 4 program package ». In : *Molecular Physics* 113.2 (sept. 2014), p. 184-215. DOI : 10.1080/00268976.2014.952696. URL : <https://doi.org/10.1080/00268976.2014.952696>.
- [176] Alexander G. SHARD. « Practical guides for X-Ray photoelectron spectroscopy : Quantitative XPS ». In : *Journal of Vacuum Science & Technology A* 38.4 (juil. 2020), p. 041201. DOI : 10.1116/1.5141395. URL : <https://doi.org/10.1116/1.5141395>.
- [177] Isaiah SHAVITT. « The History and Evolution of Gaussian Basis Sets ». In : *Israel Journal of Chemistry* 33.4 (1993), p. 357-367. DOI : 10.1002/ijch.199300044. URL : <https://doi.org/10.1002/ijch.199300044>.
- [178] Isaiah SHAVITT et Rodney J. BARTLETT. *Many-body methods in chemistry and physics : MBPT and coupled-cluster theory*. Cambridge molecular science. Cambridge University Press, 2009. ISBN : 978-0-521-81832-2.
- [179] Avijit SHEE et al. « Equation-of-motion coupled-cluster theory based on the 4-component Dirac–Coulomb(–Gaunt) Hamiltonian. Energies for single electron detachment, attachment, and electronically excited states ». In : *The Journal of Chemical Physics* 149.17 (nov. 2018), p. 174113. DOI : 10.1063/1.5053846. URL : <https://doi.org/10.1063/1.5053846>.
- [180] C. David SHERRILL et Henry F. SCHAEFER. « The Configuration Interaction Method : Advances in Highly Correlated Approaches ». In : *Advances in Quantum Chemistry*. Elsevier, 1999, p. 143-269. DOI : 10.1016/s0065-3276(08)60532-8. URL : [https://doi.org/10.1016/s0065-3276\(08\)60532-8](https://doi.org/10.1016/s0065-3276(08)60532-8).

- [181] Jocelyne SHIM et al. « Calibration and applications of the Δ MP2 method for calculating core electron binding energies ». In : *Physical Chemistry Chemical Physics* 13.13 (2011), p. 5703. DOI : 10.1039/c0cp01591a. URL : <https://doi.org/10.1039/c0cp01591a>.
- [182] Per E. M. SIEGBAHN et al. « The complete active space SCF (CASSCF) method in a Newton–Raphson formulation with application to the HNO molecule ». In : *The Journal of Chemical Physics* 74.4 (fév. 1981), p. 2384-2396. DOI : 10.1063/1.441359. URL : <https://doi.org/10.1063/1.441359>.
- [183] Jetze SIKKEMA et al. « The molecular mean-field approach for correlated relativistic calculations ». In : *The Journal of Chemical Physics* 131.12 (sept. 2009), p. 124116. DOI : 10.1063/1.3239505. URL : <https://doi.org/10.1063/1.3239505>.
- [184] Christopher SOUTH et al. « 4-Component relativistic calculations of L3ionization and excitations for the isoelectronic species UO_2^{2+} , OUN^+ and UN_2 ». In : *Physical Chemistry Chemical Physics* 18.31 (2016), p. 21010-21023. DOI : 10.1039/c6cp00262e. URL : <https://doi.org/10.1039/c6cp00262e>.
- [185] Stephen H. SOUTHWORTH et al. « Observing pre-edge K-shell resonances in Kr, Xe, and XeF_2 ». In : *Physical Review A* 100.2 (août 2019). DOI : 10.1103/physreva.100.022507. URL : <https://doi.org/10.1103/physreva.100.022507>.
- [186] Alfred SPECTOR et David GIFFORD. « The space shuttle primary computer system ». In : *Communications of the ACM* 27.9 (sept. 1984), p. 872-900. DOI : 10.1145/358234.358246. URL : <https://doi.org/10.1145/358234.358246>.
- [187] John F. STANTON et Rodney J. BARTLETT. « The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties ». en. In : *The Journal of Chemical Physics* 98.9 (mai 1993), p. 7029-7039. ISSN : 0021-9606, 1089-7690. DOI : 10.1063/1.464746. URL : <http://aip.scitation.org/doi/10.1063/1.464746> (visité le 06/06/2018).
- [188] John F. STANTON et Jürgen GAUSS. « Perturbative treatment of the similarity transformed Hamiltonian in equation-of-motion coupled-cluster approximations ». In : *The Journal of Chemical Physics* 103.3 (juil. 1995), p. 1064-1076. DOI : 10.1063/1.469817. URL : <https://doi.org/10.1063/1.469817>.

- [189] John F. STANTON et al. « A direct product decomposition approach for symmetry exploitation in many-body methods. I. Energy calculations ». In : *The Journal of Chemical Physics* 94.6 (mar. 1991), p. 4334-4345. DOI : 10.1063/1.460620. URL : <https://doi.org/10.1063/1.460620>.
- [190] Fred A. STEVIE et Carrie L. DONLEY. « Introduction to X-Ray photoelectron spectroscopy ». In : *Journal of Vacuum Science & Technology A* 38.6 (déc. 2020), p. 063204. DOI : 10.1116/6.0000412. URL : <https://doi.org/10.1116/6.0000412>.
- [191] SUMMIT HPC - Oak Ridge Leadership Computing Facility. URL : <https://www.olcf.ornl.gov/summit/>.
- [192] Ayaki SUNAGA et Trond SAUE. *Towards highly accurate calculations of parity violation in chiral molecules : relativistic coupled-cluster theory including QED-effects*. 2021. arXiv : 2107.00682 [physics.chem-ph].
- [193] Attila SZABO et Neil S. OSTLUND. *Modern Quantum Chemistry : Introduction to Advanced Electronic Structure Theory (Dover Books on Chemistry)*. Dover Publications, 2012. ISBN : 0486691861.
- [194] Attila TAJTI et Péter G. SZALAY. « Investigation of the Impact of Different Terms in the Second Order Hamiltonian on Excitation Energies of Valence and Rydberg States ». In : *Journal of Chemical Theory and Computation* 12.11 (oct. 2016), p. 5477-5482. DOI : 10.1021/acs.jctc.6b00723. URL : <https://doi.org/10.1021/acs.jctc.6b00723>.
- [195] Attila TAJTI, Levente TULIPÁN et Péter G. SZALAY. « Accuracy of Spin-Component Scaled ADC(2) Excitation Energies and Potential Energy Surfaces ». In : *Journal of Chemical Theory and Computation* 16.1 (déc. 2019), p. 468-474. DOI : 10.1021/acs.jctc.9b01065. URL : <https://doi.org/10.1021/acs.jctc.9b01065>.
- [196] Yuji TAKAHATA et Delano P. CHONG. « DFT calculation of core- and valence-shell electron excitation and ionization energies of 2, 1, 3-benzothiadiazole C₆H₄SN₂, 1, 3, 2, 4-benzodithiadiazine C₆H₄S₂N₂, and 1, 3, 5, 2, 4-benzotrithiadiazepine C₆H₄S₃N₂ ». In : *Journal of Electron Spectroscopy and Related Phenomena* 185.11 (nov. 2012), p. 475-485. DOI : 10.1016/j.elspec.2012.09.015. URL : <https://doi.org/10.1016/j.elspec.2012.09.015>.
- [197] Andrew G. TAUBE et Rodney J. BARTLETT. « Frozen Natural Orbitals : Systematic Basis Set Truncation for Coupled-Cluster Theory ». en. In : *Collection of Czechoslovak Chemical Communications* 70.6 (2005), p. 837-850. ISSN : 0010-0765, 1212-6950. DOI : 10.1135/cccc20050837. URL : <http://cccc.uochb.cas.cz/70/6/0837/> (visité le 13/06/2021).

- [198] Debbie F. TUAN et Oktay SINANOĞLU. « Many-Electron Theory of Atoms and Molecules. IV. Be Atom and Its Ions ». In : *The Journal of Chemical Physics* 41.9 (nov. 1964), p. 2677-2688. DOI : 10.1063/1.1726337. URL : <https://doi.org/10.1063/1.1726337>.
- [199] Steven VANCOILLIE et al. « Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Studies on Mononuclear First-Row Transition-Metal Systems ». In : *Journal of Chemical Theory and Computation* 7.12 (nov. 2011), p. 3961-3977. DOI : 10.1021/ct200597h. URL : <https://doi.org/10.1021/ct200597h>.
- [200] Marta L. VIDAL et al. « New and Efficient Equation-of-Motion Coupled-Cluster Framework for Core-Excited and Core-Ionized States ». In : *Journal of Chemical Theory and Computation* 15.5 (avr. 2019), p. 3117-3133. DOI : 10.1021/acs.jctc.9b00039. URL : <https://doi.org/10.1021/acs.jctc.9b00039>.
- [201] Lucas VISSCHER. « Approximate molecular relativistic Dirac-Coulomb calculations using a simple Coulombic correction ». In : *Theoretical Chemistry Accounts : Theory, Computation, and Modeling (Theoretica Chimica Acta)* 98.2-3 (déc. 1997), p. 68-70. DOI : 10.1007/s002140050280. URL : <https://doi.org/10.1007/s002140050280>.
- [202] Lucas VISSCHER et Kenneth G. DYALL. « DIRAC-FOCK atomic electronic structure calculations using different nuclear charge distributions ». In : *Atomic Data and Nuclear Data Tables* 67.2 (nov. 1997), p. 207-224. DOI : 10.1006/adnd.1997.0751. URL : <https://doi.org/10.1006/adnd.1997.0751>.
- [203] Lucas VISSCHER, Jacek STYSZYŃSKI et W. C. NIEUWPOORT. « Relativistic and correlation effects on molecular properties. II. The hydrogen halides HF, HCl, HBr, HI, and HAt ». In : *The Journal of Chemical Physics* 105.5 (août 1996), p. 1987-1994. DOI : 10.1063/1.472066. URL : <https://doi.org/10.1063/1.472066>.
- [204] Lucas VISSCHER et al. « Kinetic balance in contracted basis sets for relativistic calculations ». In : *International Journal of Quantum Chemistry* 40.S25 (1991), p. 131-139. DOI : 10.1002/qua.560400816. URL : <https://doi.org/10.1002/qua.560400816>.
- [205] Zhifan WANG, Zheyang TU et Fan WANG. « Equation-of-Motion Coupled-Cluster Theory for Excitation Energies of Closed-Shell Systems with Spin-Orbit Coupling ». In : *Journal of Chemical Theory and Computation* 10.12 (nov. 2014), p. 5567-5576. DOI : 10.1021/ct500854m. URL : <https://doi.org/10.1021/ct500854m>.

- [206] John D. WATTS et Rodney J. BARTLETT. « The coupled-cluster single, double, and triple excitation model for open-shell single reference functions ». In : *The Journal of Chemical Physics* 93.8 (oct. 1990), p. 6104-6105. DOI : 10.1063/1.459002. URL : <https://doi.org/10.1063/1.459002>.
- [207] John D. WATTS, Jürgen GAUSS et Rodney J. BARTLETT. « Coupled-cluster methods with noniterative triple excitations for restricted open-shell Hartree-Fock and other general single determinant reference functions. Energies and analytical gradients ». In : *The Journal of Chemical Physics* 98.11 (juin 1993), p. 8718-8733. DOI : 10.1063/1.464480. URL : <https://doi.org/10.1063/1.464480>.
- [208] Jan WENZEL, Michael WORMIT et Andreas DREUW. « Calculating core-level excitations and X-Ray absorption spectra of medium-sized closed-shell molecules with the algebraic-diagrammatic construction scheme for the polarization propagator ». In : *Journal of Computational Chemistry* 35.26 (août 2014), p. 1900-1915. DOI : 10.1002/jcc.23703. URL : <https://doi.org/10.1002/jcc.23703>.
- [209] Bernd WINTER et Manfred FAUBEL. « Photoemission from Liquid Aqueous Solutions ». In : *Chemical Reviews* 106.4 (mar. 2006), p. 1176-1211. DOI : 10.1021/cr040381p. URL : <https://doi.org/10.1021/cr040381p>.
- [210] *Xilinx XCVU9P : field-programmable gate array*. URL : <https://www.xilinx.com/products/boards-and-kits.html>.
- [211] Xiang YUAN, Lucas VISSCHER et André SEVERO PEREIRA GOMES. - In preparation.
- [212] Xuechen ZHENG et Lan CHENG. « Performance of Delta-Coupled-Cluster Methods for Calculations of Core-Ionization Energies of First-Row Elements ». In : *Journal of Chemical Theory and Computation* 15.9 (juil. 2019), p. 4945-4955. DOI : 10.1021/acs.jctc.9b00568. URL : <https://doi.org/10.1021/acs.jctc.9b00568>.
- [213] Lei ZHU et al. « Photodissociation of gas-phase I_3^- : product branching in the visible and UV regions ». In : *Chemical Physics Letters* 350.3-4 (déc. 2001), p. 233-239. DOI : 10.1016/S0009-2614(01)01288-X. URL : [https://doi.org/10.1016/S0009-2614\(01\)01288-X](https://doi.org/10.1016/S0009-2614(01)01288-X).

Équations

A.1 Équations CC2

Pour $t_i^a D_i^a$, avec $D_i^a = f_i^i - f_a^a$, il n'y a pas de changement [66, 189] :

$$t_i^a D_i^a = + f_i^a + t_i^e \mathcal{F}_e^a - t_m^a \mathcal{F}_i^m + t_{im}^{ae} \mathcal{F}_e^m - t_n^f V_{if}^{na} - \frac{1}{2} t_{im}^{ef} V_{ef}^{ma} - \frac{1}{2} t_{mn}^{ae} V_{ei}^{nm}, \quad (\text{A.1.1})$$

Pour $t_{ij}^{ab} D_{ij}^{ab}$, avec $D_{ij}^{ab} = f_i^i + f_j^j - f_a^a - f_b^b$ et l'opérateur de permutation $P_-(pq)f(\dots, p, q, \dots) = f(\dots, p, q, \dots) - f(\dots, q, p, \dots)$, nous pouvons écrire :

$$\begin{aligned} \text{CC2 } t_{ij}^{ab} D_{ij}^{ab} = & + V_{ij}^{ab} + P_-(ij)t_i^e \mathcal{W}_{ej}^{ab} \text{CC2} - P_-(ab)t_m^a \mathcal{W}_{ij}^{mb} + t_m^a t_n^b \mathcal{W}_{ij}^{mn} \text{CC2} \\ & + t_i^e t_j^f \mathcal{W}_{ef}^{ab} \text{CC2} + P_-(ab)f_e^b t_{ij}^{ae} - P_-(ij)f_j^m t_{im}^{ab} \end{aligned} \quad (\text{A.1.2})$$

avec :

$$\mathcal{W}_{ij}^{mn} \text{CC2} = + V_{ij}^{mn} + P_-(ij)t_j^e V_{ie}^{mn} + \frac{1}{2} P_-(ij)t_i^e t_j^f V_{ef}^{mn} \quad (\text{A.1.3})$$

$$\mathcal{W}_{ef}^{ab} \text{CC2} = + V_{ef}^{ab} - P_-(ab)t_m^b V_{ef}^{am} + \frac{1}{2} P_-(ab)t_m^a t_n^b V_{ef}^{mn} \quad (\text{A.1.4})$$

$$\mathcal{W}_{ej}^{ab} \text{CC2} = + V_{ej}^{ab} - \frac{1}{2} P_-(ab)t_m^a V_{ej}^{mb} \quad (\text{A.1.5})$$

$$\mathcal{W}_{ij}^{mb} \text{CC2} = + V_{ij}^{mb} + \frac{1}{2} P_-(ij)t_i^e V_{ej}^{mb} \quad (\text{A.1.6})$$

A.2 Méthodes approximées

A.2.1 EOM-MBPT(2)

Les modifications sont indiquées par un code couleur :

- en vert les dépendances en τ ou en $t_1 \otimes t_1$;
- en bleu les dépendances internes aux intermédiaires ;
- en rouge : \mathcal{W}_{vv}^{vo} et \mathcal{W}_{ov}^{oo} dépendent de t_1 , la multiplication par t_2 est impossible pour respecter le second ordre en perturbation.

Pour EE

$$\begin{aligned} R\sigma_i^a = & + \bar{\mathcal{F}}_e^a r_i^e - \bar{\mathcal{F}}_i^m r_m^a + \mathcal{W}_{ei}^{ma} r_m^e \\ & + \bar{\mathcal{F}}_e^m r_{mi}^{ea} + \frac{1}{2} \mathcal{W}_{ef}^{am} r_{im}^{ef} - \frac{1}{2} \mathcal{W}_{ie}^{mn} r_{mn}^{ea} \end{aligned} \quad (\text{A.2.1})$$

$$\begin{aligned} R\sigma_{ij}^{ab} = & - P_-(ab) \mathcal{W}_{ij}^{mb} r_m^a + P_-(ij) \mathcal{W}_{ej}^{ab} r_i^e + P_-(ab) \left(\mathbf{V}_{fe}^{bm} r_n^e \right) t_{ij}^{af} - P_-(ij) \left(\mathbf{V}_{je}^{nm} r_m^e \right) t_{in}^{ab} \\ & + P_-(ab) \bar{\mathcal{F}}_e^b r_{ij}^{ae} - P_-(ij) \bar{\mathcal{F}}_j^m r_{im}^{ab} + \frac{1}{2} \mathcal{W}_{ij}^{mn} r_{mn}^{ab} + P_-(ab) P_-(ij) \mathcal{W}_{ej}^{mb} r_{im}^{ae} \\ & - \frac{1}{2} P_-(ab) \left(V_{fe}^{nm} r_{mn}^{ea} \right) t_{ij}^{fb} - \frac{1}{2} P_-(ij) \left(V_{fe}^{mn} r_{im}^{fe} \right) t_{jn}^{ba} + \frac{1}{2} \mathcal{W}_{ef}^{ab} r_{ij}^{ef} \end{aligned} \quad (\text{A.2.2})$$

pour IP , en prenant soin de séparer les lignes pour r_o et r_{oo}^v :

$$\begin{aligned} R\sigma_i = & - \bar{\mathcal{F}}_i^m r_m \\ & + \bar{\mathcal{F}}_e^m r_{im}^e - \frac{1}{2} \mathcal{W}_{ie}^{mn} r_{mn}^e \end{aligned} \quad (\text{A.2.3})$$

$$\begin{aligned} R\sigma_{ij}^a = & - \mathcal{W}_{ij}^{am} r_m \\ & - P_-(ij) \bar{\mathcal{F}}_j^m r_{im}^a + \bar{\mathcal{F}}_e^a r_{ij}^e + \mathcal{W}_{ij}^{mn} r_{mn}^a + P_-(ij) \mathcal{W}_{ie}^{am} r_{jm}^e - \frac{1}{2} \left(V_{ef}^{mn} r_{mn}^f \right) t_{ij}^{ae} \end{aligned} \quad (\text{A.2.4})$$

pour EA :

$$\begin{aligned} R\sigma^a = & + \bar{\mathcal{F}}_e^a r^e \\ & + \bar{\mathcal{F}}_e^m r_m^{ae} + \frac{1}{2} \mathcal{W}_{ef}^{am} r_m^{ef} \end{aligned} \quad (\text{A.2.5})$$

$$\begin{aligned} R\sigma_i^{ab} = & + \mathcal{W}_{ie}^{ab} r^e \\ & + P_-(ab) \bar{\mathcal{F}}_e^a r_i^{eb} - \bar{\mathcal{F}}_i^m r_m^{ab} + P_-(ab) \mathcal{W}_{ie}^{am} r_m^{eb} + \frac{1}{2} \mathcal{W}_{ef}^{ab} r_i^{ef} - \frac{1}{2} \left(V_{fe}^{mn} r_n^{fe} \right) t_{im}^{ab} \end{aligned} \quad (\text{A.2.6})$$

Avec les intermédiaires suivants :

$$\text{MBPT}(2)\bar{\mathcal{F}}_v^v = \bar{\mathcal{F}}_e^a = +f_e^a - f_e^m t_m^a + t_m^f V_{ef}^{am} - \frac{1}{2} t_{mn}^{af} V_{ef}^{mn} \quad (\text{A.2.7})$$

$$\text{MBPT}(2)\bar{\mathcal{F}}_o^o = \bar{\mathcal{F}}_i^m = +f_i^m + f_e^m t_i^e + t_n^e V_{ie}^{mn} + \frac{1}{2} t_{in}^{ef} V_{ef}^{mn} \quad (\text{A.2.8})$$

$$\text{MBPT}(2)\mathcal{W}_{oo}^{oo} = \mathcal{W}_{ij}^{mn} = +V_{ij}^{mn} + P_-(ij) t_i^e V_{ej}^{mn} + \frac{1}{2} t_{ij}^{ef} V_{ef}^{mn} \quad (\text{A.2.9})$$

$$\text{MBPT}(2)\mathcal{W}_{vv}^{vv} = \mathcal{W}_{ef}^{ab} = +V_{ef}^{ab} - P_-(ab) t_m^b V_{ef}^{am} + \frac{1}{2} t_{mn}^{ab} V_{ef}^{mn} \quad (\text{A.2.10})$$

$$\text{MBPT}(2)\mathcal{W}_{vo}^{ov} = \mathcal{W}_{ej}^{mb} = +V_{ej}^{mb} + V_{ef}^{mb} t_j^f - V_{ej}^{mn} t_n^b - V_{ef}^{mn} \left(t_{jn}^{fb} \right) \quad (\text{A.2.11})$$

$$\begin{aligned} \text{MBPT}(2)\mathcal{W}_{vo}^{vv} = \mathcal{W}_{ei}^{ab} = & +V_{ei}^{ab} - f_e^m t_{mi}^{ab} + t_i^f V_{ef}^{ab} + \frac{1}{2} V_{ei}^{mn} t_{mn}^{ab} \\ & - P_-(ab) V_{ef}^{mb} t_{mi}^{af} - P_-(ab) t_m^a V_{ei}^{mb} \end{aligned} \quad (\text{A.2.12})$$

$$\begin{aligned} \text{MBPT}(2)\mathcal{W}_{oo}^{ov} = \mathcal{W}_{ij}^{mb} = & +V_{ij}^{mb} + f_e^m t_{ij}^{be} - t_n^b V_{ij}^{mn} + \frac{1}{2} V_{ef}^{mb} t_{ij}^{ef} \\ & + P_-(ij) V_{ie}^{mn} t_{jn}^{be} + P_-(ij) t_i^e V_{ej}^{mb} \end{aligned} \quad (\text{A.2.13})$$

A.2.2 Partitioned-EOM

Seul le bloc \mathcal{H}_{DD} est modifié.

Pour EE :

$$[\bar{H}_{DD}^{[0]} \mathcal{R}]_{ij}^{ab} = +P_-(ab) f_e^b r_{ij}^{ae} - P_-(ij) f_j^m r_{im}^{ab} \quad (\text{A.2.14})$$

Pour IP , en prenant en compte la partie r_{oo}^v :

$$[\bar{H}_{DD}^{[0]} \mathcal{R}]_{ij}^a = -P_-(ij) f_j^m r_{im}^a + f_e^a r_{ij}^e \quad (\text{A.2.15})$$

Pour EA , en prenant en compte la partie r_o^{vv} :

$$[\bar{H}_{DD}^{[0]} \mathcal{R}]_i^{ab} = +P_-(ab) f_e^a r_i^{eb} - f_i^m r_m^{ab} \quad (\text{A.2.16})$$

Annexe **B**

Annexes du Chapitre 3

B.1 Courbes d'énergie potentielle

B.1.1 Courbes

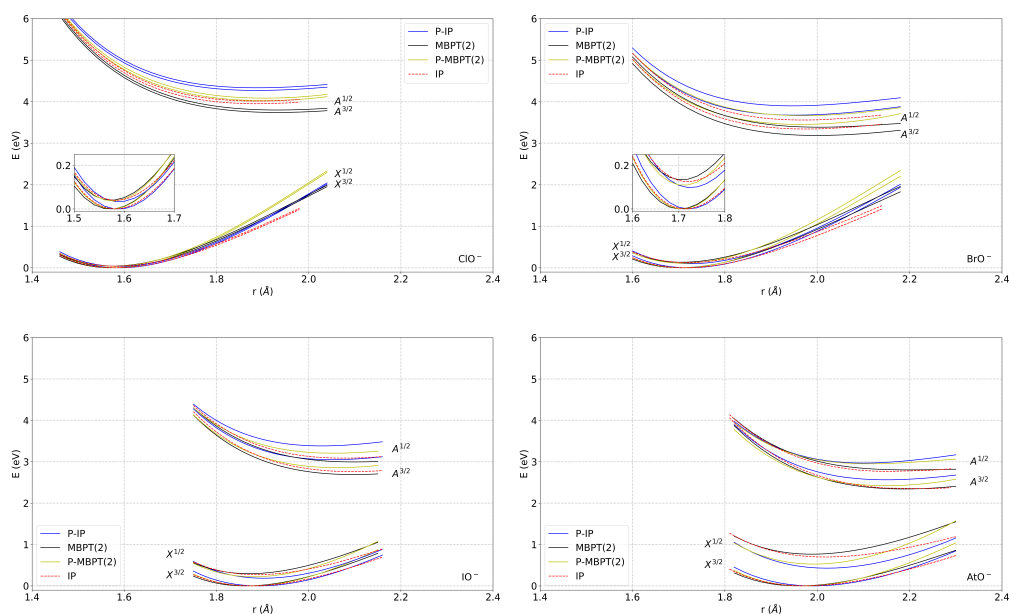


FIGURE B.1 – Courbes d'énergie potentielle - XO^- $X \in \{Cl; Br; I; At\}$ pour toutes les méthodes : IP, P-IP, MBPT(2)-IP et P-MBPT(2)-IP - Énergie : E en eV et la distance X-O : r en Å.

B.1.2 Tables : Position des minimums - Énergies de transition

TABLEAU B.1 – ClO^- : Minimums des *Courbes d'énergie potentielle* - E_{MO} énergie absolue (eV), et r la distance Cl-O la position des minimums - Δ_r la différence entre les distances minimales (Å. Δ_E la différence entre les minimums des énergies (eV)).

T	r	IP E_{MO} +14.58e3	P-IP Δ_r	Δ_E	PT(2) Δ_r	Δ_E	P-PT(2) Δ_r	Δ_E
$X^{3/2}$	1.5818	-10.9061	0.0065	-0.5095	-0.0092	0.0465	-0.0054	-0.3601
$X^{1/2}$	1.5828	-10.8665	0.0063	-0.5148	-0.0094	0.0477	-0.0057	-0.3634
$A^{3/2}$	1.8889	-6.9541	-0.0006	-0.1888	0.0429	-0.1638	0.0145	-0.2859
$A^{1/2}$	1.8875	-6.8909	0.0004	-0.1883	0.0451	-0.1678	0.0175	-0.2893

TABLEAU B.2 – ClO^- : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).

Δ_T	exp [69]	Δ_{IP}	Δ_{P-IP}	$\Delta_{PT(2)}$	$\Delta_{P-PT(2)}$	$\Delta_{P-IP} - \Delta_{IP}$	$\Delta_{PT(2)} - \Delta_{IP}$	$\Delta_{P-PT(2)} - \Delta_{IP}$
$X^{3/2} - X^{3/2}$		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$X^{1/2} - X^{3/2}$	0.0397	0.0396	0.0343	0.0408	0.0363	-0.0054	0.0012	-0.0034
$A^{3/2} - X^{3/2}$		3.9520	4.2727	3.7418	4.0262	0.3207	-0.2103	0.0741
$A^{1/2} - X^{3/2}$		4.0153	4.3364	3.8009	4.0860	0.3212	-0.2144	0.0708

TABLEAU B.3 – BrO^- : Minimums des *Courbes d'énergie potentielle* - E_{MO} énergie absolue (eV), et r la distance Br-O la position des minimums - Δ_r la différence entre les distances minimales (Å. Δ_E la différence entre les minimums des énergies (eV)).

T	r	IP E_{MO} +72.90e3	P-IP Δ_r	Δ_E	PT(2) Δ_r	Δ_E	P-PT(2) Δ_r	Δ_E
$X^{3/2}$	1.7141	-7.2208	0.0073	-0.5370	-0.0112	0.0878	-0.0070	-0.3341
$X^{1/2}$	1.7192	-7.0952	0.0064	-0.5649	-0.0125	0.0958	-0.0088	-0.3505
$A^{3/2}$	1.9751	-3.8787	-0.0234	-0.2027	0.0242	-0.0673	-0.0108	-0.2226
$A^{1/2}$	1.9692	-3.6622	-0.0229	-0.1957	0.0360	-0.0848	0.0001	-0.2277

TABLEAU B.4 – BrO^- : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).

Δ_T	exp [69]	Δ_{IP}	Δ_{P-IP}	$\Delta_{PT(2)}$	$\Delta_{P-PT(2)}$	$\Delta_{P-IP} - \Delta_{IP}$	$\Delta_{PT(2)} - \Delta_{IP}$	$\Delta_{P-PT(2)} - \Delta_{IP}$
$X^{3/2} - X^{3/2}$		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$X^{1/2} - X^{3/2}$	0.1270	0.1256	0.0978	0.1336	0.1092	-0.0279	0.0080	-0.0164
$A^{3/2} - X^{3/2}$		3.3421	3.6764	3.1871	3.4536	0.3344	-0.1550	0.1115
$A^{1/2} - X^{3/2}$		3.5586	3.8999	3.3860	3.6650	0.3414	-0.1726	0.1065

TABLEAU B.5 – IO^- : Minimums des *Courbes d'énergie potentielle* - E_{MO} énergie absolue (eV), et r la distance I-O la position des minimums - Δ_r la différence entre les distances minimales (Å. Δ_E la différence entre les minimums des énergies (eV)).

T	r	IP E_{MO} +1.955e5	P-IP Δ_r	Δ_E	PT(2) Δ_r	Δ_E	P-PT(2) Δ_r	Δ_E
$X^{3/2}$	1.8758	-26.7779	0.0164	-0.5414	-0.0124	0.0974	-0.0026	-0.3049
$X^{1/2}$	1.8874	-26.5063	0.0149	-0.6274	-0.0165	0.1214	-0.0063	-0.3554
$A^{3/2}$	2.0950	-24.0159	-0.0427	-0.2527	0.0020	0.0263	-0.0334	-0.2087
$A^{1/2}$	2.0767	-23.6909	-0.0484	-0.2448	0.0234	0.0064	-0.0227	-0.1845

TABLEAU B.6 – IO^- : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).

Δ_T	exp [69]	Δ_{IP}	Δ_{P-IP}	$\Delta_{PT(2)}$	$\Delta_{P-PT(2)}$	$\Delta_{P-IP} - \Delta_{IP}$	$\Delta_{PT(2)} - \Delta_{IP}$	$\Delta_{P-PT(2)} - \Delta_{IP}$
$X^{3/2}-X^{3/2}$		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$X^{1/2}-X^{3/2}$	0.2593	0.2717	0.1856	0.2956	0.2211	-0.0860	0.0240	-0.0505
$A^{3/2}-X^{3/2}$		2.7620	3.0507	2.6909	2.8583	0.2887	-0.0711	0.0962
$A^{1/2}-X^{3/2}$		3.0870	3.3835	2.9960	3.2074	0.2966	-0.0910	0.1204

TABLEAU B.7 – AtO^- : Minimums des *Courbes d'énergie potentielle* - E_{MO} énergie absolue (eV), et r la distance At-O la position des minimums - Δ_r la différence entre les distances minimales (Å. Δ_E la différence entre les minimums des énergies (eV)).

T	r	IP E_{MO} +6.247e5	P-IP Δ_r	Δ_E	PT(2) Δ_r	Δ_E	P-PT(2) Δ_r	Δ_E
$X^{3/2}$	1.9734	-20.3352	0.0181	-0.5979	-0.0131	0.2357	-0.0037	-0.2263
$X^{1/2}$	2.0180	-19.6367	0.0031	-0.8677	-0.0298	0.3031	-0.0248	-0.3974
$A^{3/2}$	2.2090	-17.9910	-0.0594	-0.3774	-0.0201	0.2337	-0.0597	-0.1468
$A^{1/2}$	2.1645	-17.5705	-0.0642	-0.3925	0.0339	0.2746	-0.0517	-0.0394

TABLEAU B.8 – AtO^- : Énergie des transitions Δ_E et différence avec Δ_{IP} en (eV).

Δ_T	Δ_{IP}	Δ_{P-IP}	$\Delta_{PT(2)}$	$\Delta_{P-PT(2)}$	$\Delta_{P-IP} - \Delta_{IP}$	$\Delta_{PT(2)} - \Delta_{IP}$	$\Delta_{P-PT(2)} - \Delta_{IP}$
$X^{3/2}-X^{3/2}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$X^{1/2}-X^{3/2}$	0.6985	0.4287	0.7659	0.5274	-0.2698	0.0674	-0.1711
$A^{3/2}-X^{3/2}$	2.3442	2.5647	2.3422	2.4237	0.2205	-0.0020	0.0795
$A^{1/2}-X^{3/2}$	2.7647	2.9701	2.8036	2.9516	0.2055	0.0389	0.1869

B.2 Géométrie des molécules

$$d_{I_3^- \text{ linéaire}} = 2,9300 \text{ \AA}$$

$$d_{H-Cl} = 1,2759 \text{ \AA};$$

$$d_{H-Br} = 1,4107 \text{ \AA};$$

	ClO ⁻	BrO ⁻	IO ⁻	AtO ⁻	TsO ⁻
d(Å)	1.582000	1.714000	1.875000	1.973000	2.137000

TABLEAU B.9 – XO⁻ - Distances internucléaires en Å.

Atome	x	y	z
C	0.000000	0.000000	0.216500
H	0.000000	-0.905400	0.823100
H	0.000000	0.905400	0.823100
I	1.824300	0.000000	-0.931000
I	-1.824300	0.000000	-0.931000

Atome	x	y	z
C	0.09467453024	0.00000000000	0.17555798960
H	0.10344624410	-0.90692908870	0.77962478910
H	0.10344624410	0.90692908870	0.77962478910
Br	1.69838949800	-0.00000000000	-0.92135799380
I	-1.75939712800	-0.00000000000	-0.93249672430

TABLEAU B.10 – CH₂I₂ & CH₂IBr - Coordonnées cartésiennes en Å.

Annexe **C**

Relativistic EOM-CCSD for
Core-Excited and Core-Ionized
State Energies Based on the
Four-Component
Dirac-Coulomb(-Gaunt) Hamiltonian

Relativistic EOM-CCSD for Core-Excited and Core-Ionized State Energies Based on the Four-Component Dirac–Coulomb(–Gaunt) Hamiltonian

Loïc Halbert,* Marta L. Vidal,* Avijit Shee,* Sonia Coriani,* and André Severo Pereira Gomes*

Cite This: *J. Chem. Theory Comput.* 2021, 17, 3583–3598

Read Online

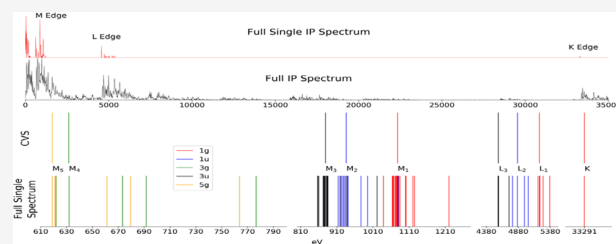
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: We report an implementation of the core–valence separation approach to the four-component relativistic Hamiltonian-based equation-of-motion coupled-cluster with singles and doubles theory (CVS-EOM-CCSD) for the calculation of relativistic core-ionization potentials and core-excitation energies. With this implementation, which is capable of exploiting double group symmetry, we investigate the effects of the different CVS-EOM-CCSD variants and the use of different Hamiltonians based on the exact two-component (X2C) framework on the energies of different core-ionized and -excited states in halogen- (CH_3I , HX , and X^- , $\text{X} = \text{Cl–At}$) and xenon-containing (Xe , XeF_2) species. Our results show that the X2C molecular mean-field approach [Sikkema, J.; et al. *J. Chem. Phys.* 2009, 131, 124116], based on four-component Dirac–Coulomb mean-field calculations ($^2\text{DC}^M$), is capable of providing core excitations and ionization energies that are nearly indistinguishable from the reference four-component energies for up to and including fifth-row elements. We observe that two-electron integrals over the small-component basis sets lead to non-negligible contributions to core binding energies for the K and L edges for atoms such as iodine or astatine and that the approach based on Dirac–Coulomb–Gaunt mean-field calculations ($^2\text{DCG}^M$) are significantly more accurate than X2C calculations for which screened two-electron spin–orbit interactions are included via atomic mean-field integrals.



1. INTRODUCTION

X-ray spectroscopies, which typically probe core electrons through electronic excitation or ionization, are particularly suitable techniques to study the local environment of atoms, molecules, and materials, as the localized nature of the core orbitals makes them very selective and sensitive.^{1,2}

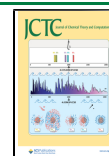
Over the last few years, new X-ray free-electron lasers (XFEL) and last-generation synchrotrons have started operating. These advanced light sources have opened the door to a variety of new X-ray-based spectroscopies,^{3–5} including those operating in time-resolved and nonlinear regimes.^{2–6} As stated, for instance, by Milne, Penfold, and Chergui in their review:⁷ “[...] The progress of experimental techniques for core level spectroscopies is unraveling subtle spectral features implying that high-level theoretical approaches are required to interpret them [...]”. The experimental X-ray spectra need to be compared to highly accurate theoretical calculations to assign spectral features and to relate experimental measurements to the structure and dynamic properties of the probed molecular system.

Among the methods, one can use to calculate core binding energies ΔSCF ,^{8,9} which stands out due to the combination of low-computational cost and good results. By performing separate self-consistent field (SCF) calculations on the original (generally neutral and closed-shell) and ionized (generally open-shell) species, it introduces the important orbital

relaxation that accompanies the creation of the core hole, though at the expense of requiring several calculations for the edges of interest and with the potential complication that the required open-shell calculations may be difficult to converge. The same approach can be applied in a straightforward manner to density functional theory (DFT),¹⁰ with the additional complication that it will show a certain dependence on the chosen density functional. The same idea of performing separate calculations for the initial and the target states can be used to devise approaches based on correlated wave functions such as MP2¹¹ or coupled-cluster (CC).^{12,13} Analogous methods for core excitations have also been formulated,^{14–22} though this has proven somewhat more cumbersome than for ionizations, since one has to select orbital pairs that should represent the transitions and converge such excited configurations. In a number of these approaches, the orbitals for the ground and excited/ionized states are not orthogonal, making, for instance, the calculation of transition moments more involved.

Received: November 18, 2020

Published: May 4, 2021



The above difficulties are circumvented with approaches that diagonalize the same Hamiltonian for both the original and the ionized or excited states. As such, the excited-state approaches based on wave function theory, like, for instance, (multi-reference) configuration interaction (MRCI),^{23–26} equation-of-motion CC (EOM-CC)^{27–35} and multireference (EOM)-CC,^{26,36} and linear response complete active space SCF (LR-CASSCF),³⁷ Green's functions such as the algebraic diagrammatic construction (ADC),^{38–41} or time-dependent density functional theory (TDDFT),^{10,42,43} can all be used to target core-excited or ionized states. One downside is, however, that now relaxation has to be accounted for by the correlated electronic structure approach. Among these, single-reference CC methods and in particular those based on EOM and the closely related linear response (LRCC) formalisms^{44–50} are capable of treating electron correlation accurately and in a balanced way between the ground and excited states, including a good deal of relaxation effects, and have an appealing “black box” nature (though they are less adept than MRCI/multireference coupled cluster (MRCC) at treating cases for which the ground state shows a strong multireference nature).

Irrespective of the correlated method, a particular difficulty for determining core states by diagonalizing a many-body Hamiltonian is that such states appear at high energies, where spectra are potentially very dense (see for example Figure 2). Thus, a naive application of procedures, which target exterior roots (such as the Davidson algorithm), would require for solving a very large number of roots. One way to overcome this issue is through the use of methods capable of solving interior roots without requiring that all of them are targeted, as those recently devised and applied to multiconfigurational,³⁷ EOM-CC,^{32,51} and mean-field^{52,53} wave functions.

A more widespread strategy,^{23,28,29,31,34,35,37–40,54–57} which we follow in this paper, is to employ physically motivated approximations to the original Hamiltonian, in particular within the core–valence separation (CVS) approximation.⁵⁸ The core–valence separation is justified by the large spatial and energetic separation between the valence and core orbitals/electrons. On the one hand, CVS allows to straightforwardly address the highly energetic core states with minor modifications to pre-existing eigenvalue solvers. On the other hand, the CVS helps alleviating convergence issues due to a large number of valence ionized states in the frequency region of the core excitations, which is particularly afflicting methods that explicitly include double excitations in their parametrization.^{59,60} One can also regard the core states as metastable Feshbach resonances embedded in the (valence) ionization continuum, and the application of CVS helps stabilizing them. Further details on the CVS approach will be given in Section 2.3.

A second issue to be addressed for core states is the importance of relativistic effects,^{61–63} which significantly alter the energies of inner electrons, and consequently, the core spectra of atoms and molecules: due to the high velocities of their electrons (a significant fraction of the speed of light), core s and p orbitals contract (with an associated lowering of their orbital energies), while due to increased screening, d and f orbitals expand (with an associated increase in their orbital energies). At the same time, orbitals that are degenerate in a nonrelativistic framework (p, d, f) are (strongly) split due to spin–orbit coupling, and particularly so for the innermost orbitals.

Relativistic effects are so pronounced in the core region that, even for molecules containing only elements of the first and

second rows, they are very important for the accurate determination of K and L edge spectra. While rather approximate treatments of relativistic effects can yield accurate results for light elements,^{12,64–67} a more consistent way of treating these effects is through an electronic Hamiltonian based on the four-component Dirac operator, as it is accurate all the way down to the heaviest elements (5d and 6d transition metals, lanthanides, and actinides). As such, four-component-based methods are ideally suited to treat the core spectra across the periodic table and to probe K, L, and M edges of heavy elements,^{11,68,69} which are much more complex to interpret than those of lighter elements. Though the availability of efficient implementations^{70–72} has in recent years enabled mean-field four-component calculation for large-scale applications, the additional basis sets required to represent the small-component part of the molecular spinors still place a considerable burden on such calculations compared to the nonrelativistic case.

An important development to overcome such issues has been the introduction of the so-called exact two-component (X2C) methods,^{73–79} in which a transformation to decouple the positive and negative-energy states of the Dirac Hamiltonian is available in the matrix form and, unlike its more approximate counterparts, yields exactly the same positive energy spectrum as the original four-component Hamiltonian. We refer the reader to a recent comprehensive review on the relativistic electronic structure for more details.⁶³

We note that among the different X2C variants, those that construct the transformation matrix on the basis of converged four-component atomic or molecular mean-field calculations^{75,77,79} are particularly interesting to correlated calculations^{79–83} since they avoid the handling of two-electron integrals over small-component basis functions in the transformation into molecular spinors, while largely avoiding errors with respect to the original four-component Hamiltonian on the description of two-electron spin–orbit coupling contributions, compared to approaches in which the decoupling is performed based on the one-electron Dirac Hamiltonian prior to the mean-field step.⁷⁶

In this study, we present an approach to study core ionization (and excitation) through the use of projectors, inspired by the work of Coriani and Koch^{28,29} but extended to the four-component EOM-CCSD approach developed by Shee et al.⁸⁴ in the DIRAC relativistic electronic structure package.⁸⁵ Because of the availability of different Hamiltonians in DIRAC, we also investigate the performance of two-component approaches⁷⁹ with respect to the four-component one. We refer to, e.g., refs 86–101 for other examples of relativistic EOM-CC implementations, and to refs 12, 31, 34, 35, 55, 59, 60, 65, 102, 103 for other examples of CVS-EOM-CC implementations. We also wish to underline that the literature on the theoretical approaches for core spectroscopy is vast and rapidly increasing, and the studies cited in the previous paragraphs cannot be considered exhaustive. We refer the reader to a number of recent review papers for more thorough accounts.^{2,24,43,104}

The paper is organized as follows. In Section 2, we give a brief outline of EOM-CCSD (Section 2.1) and of CVS (Section 2.2), as well as the details of the current implementation (Section 2.3). In Section 3, we provide the computational details of the calculations. In Section 4, we present the results obtained with the newly implemented method, where we discuss the accuracy of the CVS approximation (Section 4.1), the performance of the different CVS variants (Section 4.2), the influence of the Hamiltonian (Section 4.3), and the comparison to experiment

for ionization (Section 4.4) and excitation (Section 4.5) energies. Finally, Section 5 summarizes our conclusions.

2. METHODS AND IMPLEMENTATION

In what follows, indices (i, j, k, l) , (a, b, c, d) , and (p, q, r, s) refer to occupied, virtual, and general orbitals, respectively.

2.1. EOM-CC. In EOM-CC, the ground state is treated at the coupled-cluster level. Its eigenfunction is hence given by the exponential ansatz

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}}|\Phi_0\rangle; \quad \hat{T} = \sum_{\mu} t_{\mu}\hat{\tau}_{\mu} \quad (1)$$

where Φ_0 is the reference (typically Hartree–Fock) determinant and the operator \hat{T} is the cluster operator. Truncating \hat{T} to single (S) and double (D) excitations yields the coupled-cluster singles-and-doubles (CCSD) model

$$\hat{T} = \hat{T}_1 + \hat{T}_2; \quad \hat{T}_1 = \sum_{ia} t_i^a a_a^\dagger a_i; \quad \hat{T}_2 = \frac{1}{4} \sum_{ijab} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i \quad (2)$$

The energy and the cluster amplitudes are found from the CC equations

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = E \quad (3)$$

$$\langle \Phi_{\mu} | \hat{H} | \Phi_0 \rangle = 0; \quad |\Phi_{\mu}\rangle = \hat{\tau}_{\mu} |\Phi_0\rangle \quad (4)$$

where the similarity-transformed Hamiltonian \hat{H} has been defined as

$$\hat{H} \equiv e^{-\hat{T}} \hat{H} e^{\hat{T}} \quad (5)$$

In the equation-of-motion coupled-cluster (EOM-CC) method,^{44,47} the target states are obtained by the diagonalization of the non-Hermitian similarity-transformed Hamiltonian. This non-Hermiticity gives rise to right (R) and left (L) eigenvectors that are not adjoints of each other, obtained solving two different eigenvalue equations

$$\hat{H} |R_{\mu}\rangle = E_{\mu} |R_{\mu}\rangle \quad (6)$$

$$\langle L_{\mu} | \hat{H} = E_{\mu} \langle L_{\mu} | \quad (7)$$

for a given excited state μ with energy E_{μ} , where the eigenstates are chosen to satisfy the biorthogonality conditions

$$\langle L_{\mu} | R_{\nu} \rangle = \delta_{\mu\nu} \quad (8)$$

Right and left wave functions of the target states have been thus obtained from a linear parametrization of the reference state through the \hat{R} or \hat{L} operators

$$|\Psi_{\mu}\rangle = e^{\hat{T}} \hat{R}_{\mu} |\Phi_0\rangle \quad (9)$$

and

$$\langle \bar{\Psi}_{\mu} | = \langle \Phi_0 | \hat{L}_{\mu} e^{-\hat{T}} \quad (10)$$

Therefore, the choice of the \hat{R} and \hat{L} operators defines which target states to study, yielding different EOM-CC models. In the present work, we focus on the models for electronically excited (EE) states and ionization potentials (IP)

- EOM-CCSD-EE

$$\hat{R}^{\text{EE}} = r_0 + \sum_{ia} r_i^a \{a_a^\dagger a_i\} + \sum_{i>j, a>b} r_{ij}^{ab} \{a_a^\dagger a_b^\dagger a_j a_i\} \quad (11)$$

$$\hat{L}^{\text{EE}} = l_0 + \sum_{ia} l_i^a \{a_i^\dagger a_a\} + \sum_{i>j, a>b} l_{ij}^{ab} \{a_i^\dagger a_j^\dagger a_b a_a\} \quad (12)$$

- EOM-CCSD-IP

$$\hat{R}^{\text{IP}} = \sum_i r_i \{a_i\} + \sum_{i>j, a} r_{ij}^a \{a_a^\dagger a_j a_i\} \quad (13)$$

$$\hat{L}^{\text{IP}} = \sum_i l_i \{a_i^\dagger\} + \sum_{i>j, a} l_{ij}^a \{a_j^\dagger a_i^\dagger a_a\} \quad (14)$$

where curly brackets refer to normal ordering with respect to the Fermi vacuum defined by the reference Φ_0 and the sets $\{r\}$, $\{l\}$ to the amplitudes of the corresponding operators.

We have here truncated our \hat{R} and \hat{L} operators at the singles-doubles level since the same truncation is used for the \hat{T} operators.

2.2. Core–Valence Separation Approximation. The essence of the core–valence separation (CVS) approximation⁵⁸ is to decouple valence and core electrons based on their difference in energy and spatial extension. This allows solving the regular, in this case, EOM-CC equations only in the space of the relevant orbitals. However, different flavors of CVS exist, which introduce different levels of approximation. Coriani and Koch first introduced the CVS approximation within coupled-cluster theory^{28,29} by applying a projector that zeroes out the amplitudes of all excitations that do not involve at least one core electron. Recently, the frozen-core (FC) CVS-EOM-CCSD approach has been proposed that introduces a further approximation, in this case at the ground-state level: the core orbitals are frozen when solving the CC equations for the ground state, whereas they are the only active ones when solving the EOM equations. Both schemes retain the contribution from excitations simultaneously involving two core orbitals. However, these excitations are located in a much higher range of energy in the spectrum and might therefore also be decoupled. This is the strategy adopted, for instance, by Dreuw and co-workers in their implementation of the CVS within the ADC family of methods.^{39,40,105}

In this work, the CVS implementation has been carried out following the recipe proposed by Coriani and Koch:²⁸ a projector P is applied at each iteration of the Davidson procedure during the resolution of the EOM-CC equations. This projector P selectively zeroes out the unwanted contributions to the EOM trial/solution vectors according to the approximation used. In the CVS approach, unwanted contributions generally correspond to excited determinants involving only occupied valence (v) spinors. In the case of EOM-EE, the application of a projector $P \equiv P^{\text{CVS}}$ to a target electronic state $|\Psi_{\mu}^{\text{EE}}\rangle$ corresponds to

$$P^{\text{CVS}} |\Psi_{\mu}^{\text{EE}}\rangle \Rightarrow \mu_r^a = \mu_{rj}^{ab} = 0 \text{ if } i, j \in v \quad (15)$$

though one can also devise variants in which additional contributions are zeroed out, see Section 2.3 for details. With this definition, the eigenvalue equation to be solved to get the energy of the target states becomes

$$P^{\text{CVS}} (\hat{H} P^{\text{CVS}} |R_{\mu}\rangle) = E_{\mu} P^{\text{CVS}} |R_{\mu}\rangle \quad (16)$$

It should be noted that the projection scheme does not yield any savings, in terms of memory usage or operations. Some of us³¹ have recently gone beyond the use of projection operators and introduced an analytical formulation of CVS, in which only the sub-blocks of the similarity-transformed Hamiltonian that are relevant to the calculation of a given core level are actually formed, thus resulting in a more efficient implementation. This formulation is naturally combined with the frozen-core approximation separating valence and core spaces. However, one may argue that it introduces at the same time an additional error from neglecting core correlation.^{31,103}

2.3. Implementation Details. In this work, we focus on exploring the definition of different CVS variants and assessing their performance following the projection-based scheme, due to its ease of implementation, and with it in view of later implementing the analytical CVS³¹ formulation in DIRAC.

We have based our implementation on a flexible scheme to define projection operators.

- A mapping is defined between the excited determinants (and the virtual and occupied spinors' energies associated with each) and the position of each excited determinant in the storage of the ground-state CC amplitudes and EOM-CC coefficients (in the RELCC module of DIRAC, these are stored in triangular forms, and blocked by symmetry, for details see refs 106–108).
- Information is gathered on how the excited determinant space $\{v\}$ will be treated, i.e., whether retained ($P^v = 1$) or projected out ($P^v = 0$). We have made two options available to the user, based on spinor energies ϵ : (i) restricted excitation windows (REWs) for the occupied and virtual, which are defined by setting respective lower (ϵ_L) and upper (ϵ_H) bounds for the spinor's energies and (ii) CVS, via a single energy that acts as the threshold (ϵ_{CV}) for the separation between core and valence.
- Using the information from the previous steps, suitable one-particle (singly ionized or excited) unit trial vectors for the core states are then generated, using the values of the diagonal of \bar{H} within the subspaces defined in (b).

It is important to note that in our implementation, tracking the core-excited or -ionized states requires the use of a root homing procedure, in which new trial vectors are created by maximizing their overlap with the preceding ones.

In addition to CVS and REW for the excited states, we also used the projection setup to implement the frozen-core approximation, such that we can project out the amplitudes corresponding to core (c) orbitals at each iteration during the resolution of the ground-state amplitude equations

$$P^{\text{core}}|\Psi_{\text{CC}}\rangle \Rightarrow t_i^a = t_{ij}^{ab} = 0 \text{ if } i, j \in c \quad (17)$$

As for the excited states, this projector is defined in terms of spinor energies, via a single threshold or upper and lower bounds defining a window. Finally, it should be noted that the thresholds defining the ground-state and excited-state projectors are independent.

3. COMPUTATIONAL DETAILS

All coupled-cluster calculations were carried out with the DIRAC electronic structure code⁸⁵ (with the DIRAC19¹⁰⁹ release and revisions dbbf6a6a, 0757608, 323ab67, 2628039, 1e798e5, b9f45bd). The Dyall basis sets^{110–112} of triple-zeta quality (dyall.acv3z) were employed for all species. For selected calculations, on heavy atoms only, we performed

extrapolations to the complete basis set limit by also considering quadruple-zeta quality Dyall (dyall.acv4z) basis sets. In addition to the Dyall basis sets, the ANO-RCC basis¹¹³ were employed for Xe and XeF₂. The basis sets were kept uncontracted in all calculations.

Unless otherwise noted, all occupied and virtual spinors were considered in the correlation treatment.

Apart from the Dirac–Coulomb (⁴DC) Hamiltonian, for selected calculations, we investigated: (a) the molecular mean-field⁷⁹ approximation to DC (²DC^M) and the Dirac–Coulomb–Gaunt (²DCG^M) Hamiltonians. For the latter, the Gaunt-type integrals are explicitly taken into account only during the four-component SCF step due to the fact that the transformation of these to MO basis is currently not implemented; (b) the DC Hamiltonian with projecting out all negative-energy solutions of the bare-nucleus one-electron Dirac Hamiltonian from the molecular spinor space, which corresponds to the Furry basis for quantum electrodynamics (QED) and no-pair Hamiltonians¹¹⁴ (⁴DC^{PF}); and (c) the exact two-component (X2C)⁷⁶ Hamiltonian, in which we include two-electron spin–orbit contributions to the untransformed two-electron potential via atomic mean-field contributions calculated with the AMFI code,^{115–117} namely, spin-same-orbit (X2C-AMFI) and spin-same-orbit and spin-other-orbit (X2C-G-AMFI).

Unless otherwise noted, we employed the usual approximation of the energy contribution from (SSISS)-type two-electron integrals by a point-charge model.¹¹⁸ As will be discussed below, this approximation introduces negligible errors (<0.01 eV) for lower-energy edges. However, for the higher-energy edges of the heavier systems, its error becomes important.

The XeF₂ structure was taken from ref 119. It corresponds to an optimized structure obtained at the SFX2C-1e/coupled-cluster single double triple (CCSD(T))/unc-atomic natural orbital (ANO)-RCC level ($r_{\text{Xe-F}} = 1.9736 \text{ \AA}$). The coordinates of CH₃I come from the experimental data in ref 120. Finally, the coordinates for HX were taken from ref 121.

The dataset can be retrieved at the Zenodo repository.¹²²

3.1. Approximations Used in CVS. Besides the original CC-CVS approach,²⁸ we have also investigated other approximations:

- Further restricting the definition of the projectors to also drop excited configurations with two core indices (ND, for “No Doubly core hole determinants”); this corresponds to the method suggested, e.g., in ref 40, obtaining

$$\mu_{ij}^{ab} = 0 \text{ if } i, j \in c \quad (18)$$

- Freezing the core in the ground-state calculation, thus retaining only the valence orbitals, typically defined as the orbitals with the highest principal quantum number n (FC-V); this essentially corresponds to using eq 17 while setting ϵ_{CV} to a fairly high value (for example, in the atom of Xe, $\epsilon_{4d} < \epsilon_{CV} < \epsilon_{5s}$).
- Freezing all core orbitals, except for those that are to be targeted in the EOM step (FC-V-except); for example, this means treating a core spinor k , that should make up the most important core-excited configurations for state μ , different from other core spinors in the ground-state calculation. This corresponds to setting projectors for ground and excited states such that

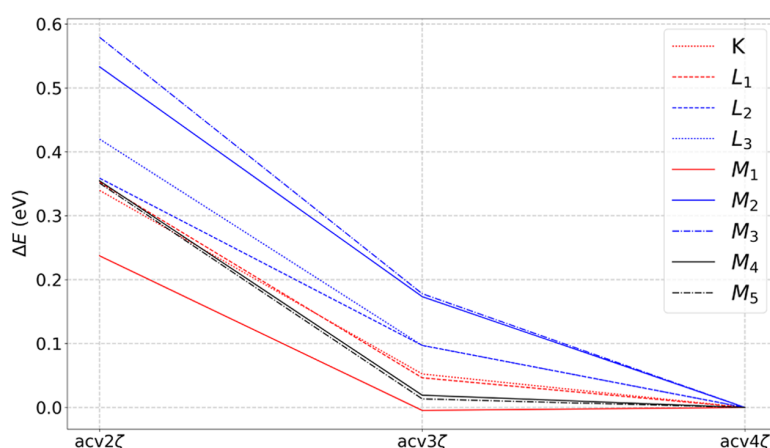


Figure 1. Basis set influence (Dyall basis sets) on the core binding energies of I^- using the CVS approach. All calculations performed with the ${}^2DC^M$ Hamiltonian, and basis sets are kept uncontracted.

$$t_i^a = t_{ij}^{ab} = 0 \text{ if } \begin{cases} i, j \neq k \\ i, j \in c \end{cases} \quad (19)$$

$$\mu_{r_i}^a = \mu_{r_{ij}}^{ab} = 0 \text{ if } i, j \in v \quad (20)$$

This approach is equivalent to the one of Sorensen et al.¹⁰³

- The definition of the core/vale space follows that of the CVS definition/threshold, that is, only the core orbitals with the same or lower energy than the ones belonging to the edge under investigation are frozen (FC-f). This is the same approach adopted in ref 31, and it corresponds to setting projectors for ground and excited states such that

$$\text{if } i, j \in v \begin{cases} t_i^a = t_{ij}^{ab} = 0 \\ \mu_{r_i}^a = \mu_{r_{ij}}^{ab} = 0 \end{cases} \quad (21)$$

- Further restricting the number of frozen-core spinors to only those below the ones of the edge of interest (FC-fpMO frozen core—follow previous MO). For example, in the atom of Xe, if we are interested in the M_5 edge ($3d_{5/2}$), this would correspond to a frozen core in the ground-state calculation, including the $1s_{1/2}$ to $3d_{3/2}$ spinors, and the $4s_{1/2}$ and higher spinors make up the valence space in the EOM calculations.

The variants ND and FC-V were also employed with REWs. Since the performance of REW variants was always found to be inferior to their CVS equivalents, they will not be discussed here. The results are, however, available as the [Supporting Information](#).

4. RESULTS

As outlined above, apart from the different Hamiltonians at our disposal, the original CVS formulation itself can be modified through a number of different approximations. To draw a clearer picture of their interactions, we proceed in a stepwise manner: first, using the ${}^2DC^M$ Hamiltonian and focusing on core ionizations, we compare (A) the performance of CVS-EOM-CCSD to the original EOM-CCSD; and (B) the impact of the different approximations that can be introduced in the CVS

method itself (see [Section 3.1](#)). Then, having established the relative accuracy of CVS, we investigate (C) the effect of the Hamiltonians on the core-ionization energies, and finally (D) proceed to a comparison to experiment of the most accurate setup, focusing mostly on core ionizations and also discussing selected core excitations (E).

In what follows, we will not discuss basis set effects. Nonetheless, in [Figure 1](#), we show for I^- a typical behavior: first, there is very little difference between triple- and quadruple-zeta results (with differences between 0.05 and 0.2 eV), making the corrections, due to basis set incompleteness, so small that we consider these to be unnecessary. Second, if there are non-negligible differences between the convergence of s, p, and d shells as the basis sets are improved, the double-zeta basis sets also yield results that are quite good, with differences in binding energies with respect to the triple-zeta results no larger than around 0.4 eV, a behavior that may make these smaller basis sets interesting for calculations on larger molecules.

These conclusions are in line with earlier basis set analyses on light elements.^{64,123} In the basis set convergence study by Sarangi et al.,¹²³ in particular, it was shown that, for systems containing only light elements, uncontracted basis sets, even of relatively modest quality for the valence, can provide quite reliable core binding energies.

4.1. Comparison of CVS-EOM-IP-CCSD and EOM-IP-CCSD. As mentioned above, the CVS approach provides an efficient and robust way of targeting the low-lying core states. Quantifying the errors introduced by it is therefore an important issue, and we note proposals for methods to estimate it. Coriani and Koch²⁸ evaluated the errors of their CVS variant by comparing to full-space Lanczos results, finding deviations of at most a few hundredths of eV. The authors also proposed a perturbative correction obtained using a Löwdin partition of the Jacobian and the eigenvalue equation. A recent study by Herbst and Fransson¹²⁴ shows that the CVS error is small and stable across multiple systems within the algebraic diagrammatic method (ADC). Thus, we expect to find the same trend for CC. This article also reports the implementation of a postprocessing step, which removes the error to assess its significance.

Here, we can harness the ability of DIRAC to exploit linear symmetry, and directly compare, for selected species, the CVS-EOM-IP-CCSD states with those obtained by the full diagonalization of the different symmetry blocks of the similarity-transformed Hamiltonian.

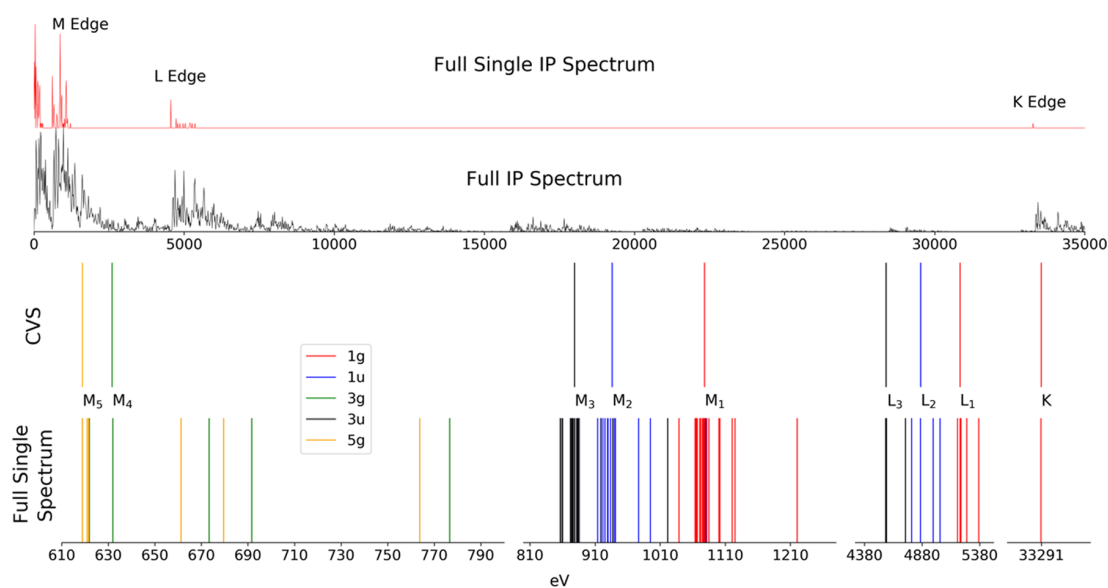


Figure 2. Core-ionization energies of iodide (I^-) from full diagonalization of \bar{H} . (Top) Representation of the eigenspectrum up to 35 keV, separated between all states (black) and those with non-negligible singly ionized contributions (red). Values in ordinate correspond to the number of states within a 10 eV interval, with the largest value rescaled to one. (Bottom) Singly ionized states obtained by full diagonalization of \bar{H} and by the core–valence separation (CVS) approach. The labels ng/nu indicate the absolute value for the projection of the angular momentum ($|m_z| = n/2$, $n = 1, 3, 5$) and for the components of spinors of gerade (g) or ungerade (u) symmetry. All calculations are performed with the ${}^2DC^M$ Hamiltonian.

The difference between the two approaches is shown graphically for I^- in Figure 2. Due to the very dense spectra of \bar{H} , in the top half of the figure, we do not display individual states but rather the number of states per 10 eV intervals. We further discriminate between all states (black) and those containing contributions from singly ionized configurations contributing to at least 1% of the total wave function (red).

Upon having a closer look at the singly ionized states at the bottom half of the figure, we can more clearly see how the singly ionized states obtained with the CVS approximation (labeled “CVS”) closely match a subset of those obtained by the full diagonalization (labeled “full”) of \bar{H} .

We note that the additional states presented for the full calculation correspond not only to states that contain significant singly ionized characters but also small but nonzero $2h1p$ contributions. The presence of such states hints at potential pitfalls when targeting only a subset of such highly excited states for some of the L and M edges—not so much in terms of convergence (we have not encountered particular difficulties in converging our calculations) but rather in terms of assignment and comparison to other theoretical or experimental results.

4.2. Performance of the CVS-EOM-CCSD Variants.

Next, we turn our attention to the performance of the different approximations that can be employed on top of CVS (see Section 3.1 for their description). Our results for the X^- systems are shown in Figure 3. As we see similar trends for Xe, the corresponding figure is shown in the Supporting Information (Figure S1), along with results for the REW approach for the Cl^- to I^- (Figures S2–S4) and HCl to HI (Figures S5–S7).

4.2.1. Freezing Core Orbitals in the Ground State. Freezing the core orbitals in the ground-state CCSD calculation results in a significant lowering of the core-ionization energies. This lowering is considerable if the frozen core is taken to represent all subvalence spinors (FC-V), and it remains non-negligible even if the same threshold is used to define both the frozen orbitals and the core–valence separation region (FC-f).

However, if one does not freeze the aimed edge (FC-fpMO), the ionization energy obtained is much closer to the corresponding CVS one.

Freezing all subvalence spinors but the ones in the edge we are interested in (FC-V-ex), an approximation that could provide potentially lower computational cost for systems with a large number of core electrons, such as the elements in the fourth row and beyond, does not perform better than (FC-f)—in fact, the opposite appears to be the case for the inner core orbitals.

4.2.2. Projecting Out Double-Core-Ionized Configurations. Projecting out doubly excited core-excited determinants (ND) tends to result in an increase of the core-ionization energies, which can be significant though always smaller than the underestimation produced by freezing all of the core spinors.

As a result, the combination of these two approximations leads to an approximate scheme that much better reproduces the CVS energies, though still with non-negligible discrepancies for the heavier elements. Particularly good results are shown by the (FC-ND-f) combination, for which error compensation yields results very similar to the regular CVS approximation. This can be rationalized because core correlation is not taken into account when using (FC-f), whereas the inclusion of the double occupied core orbitals in the EOM step does. Therefore, the combination of both approximations is in fact more consistent.

4.2.3. Efficiency Considerations. Since by using projection operators we do not save in memory or operation counts, the approximations discussed above are not of strong interest by themselves, and will not be used further since, even in the best combination, they will invariably degrade the performance of the original CVS scheme.

We can nevertheless use our findings to discuss the potential trade-offs between cost and accuracy, as a guide to efficient implementations such as those proposed in ref 31 for heavy elements. As discussed then, combining the elimination of doubly core-excited determinants with the use of a frozen core (FC-ND-f, which is equivalent to the scheme in ref 31) seems to

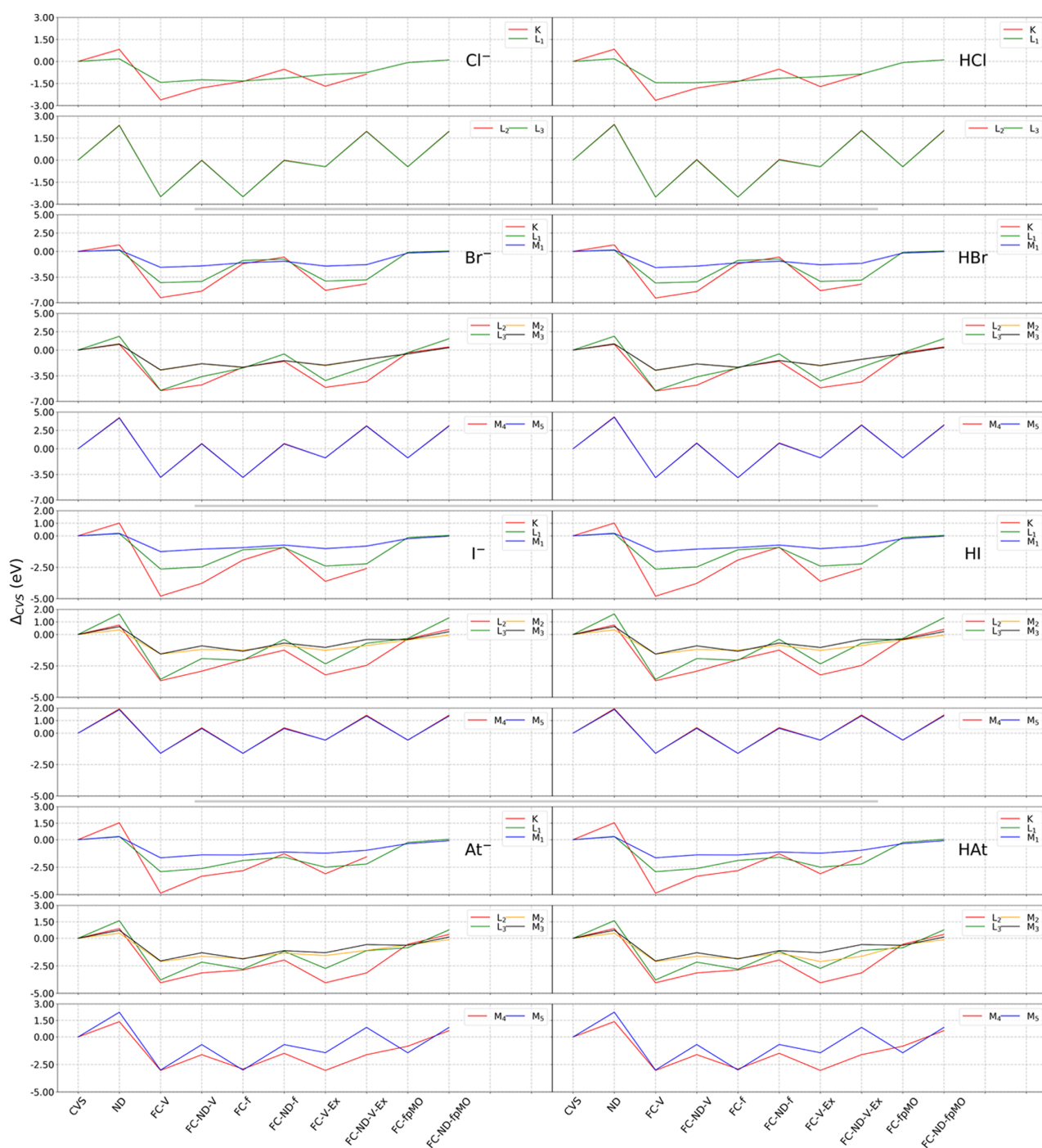


Figure 3. Effect of the CVS variants on the binding energies of X^- and HX systems ($X = \text{Cl}, \text{Br}, \text{I},$ and At). Values (in eV) are relative to the original CVS approach. All calculations are performed with the ${}^2\text{DC}^M$ Hamiltonian.

result in a computationally efficient approach with a good error cancellation balance, and this remains the case as one goes down the periodic table.

In the case of heavy elements, it may also be interesting to keep a large frozen core (CVS-FC-ND-V or CVS-FC-V-Ex) since these introduce non-negligible but seemingly systematic errors with the advantage that a large core would translate in potentially large computational savings.

This is illustrated in [Table S1](#), where we provide, for the representative systems Cl^- and At^- , the operation counts (without reductions due to point group symmetry) for the construction of EOM-EE and EOM-IP σ vectors (based on the expressions from Shee et al.,⁸⁴ see expressions in the [Supporting Information](#)), in the case of an analytical implementation of the major approximations suggested in this work (FC-f and ND). As seen, the FC-f approximation results in the biggest reductions in operation counts, in particular for EOM-IP. It is also important

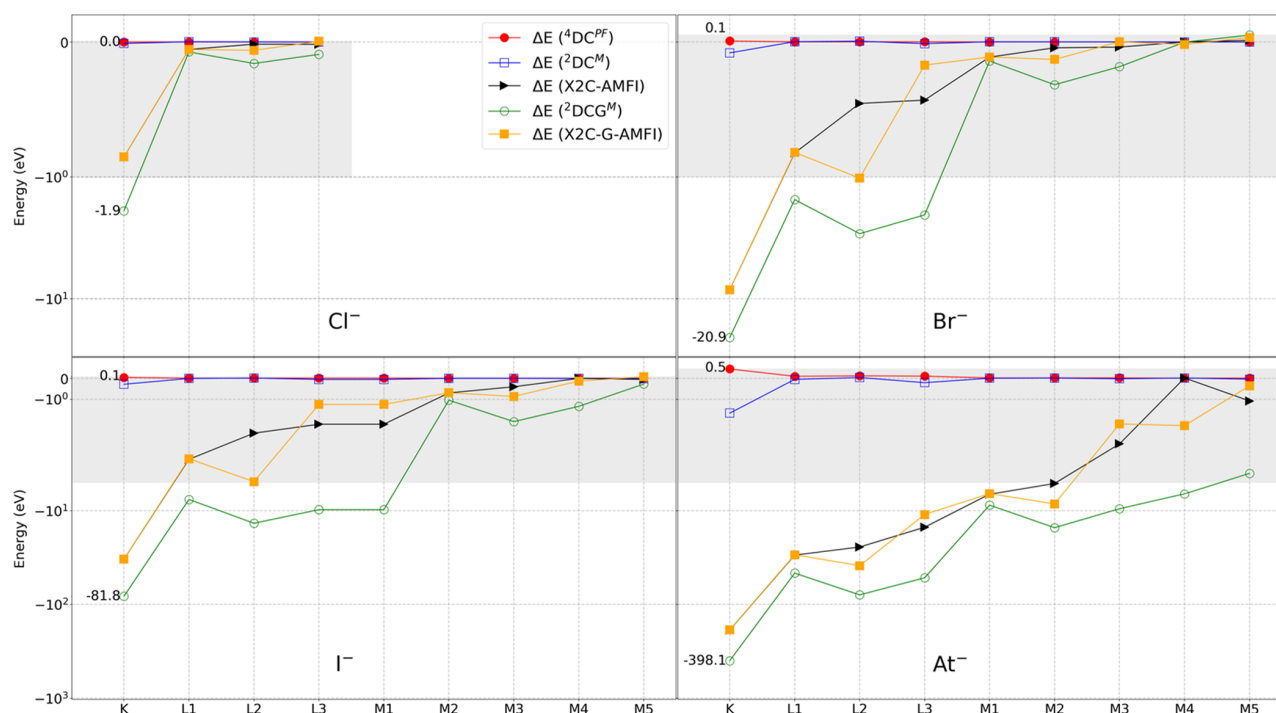


Figure 4. Comparison of the influence of the Hamiltonian to the core binding energies for different edges, for the halide ions (Cl^- to At^-). Instead of binding energies themselves, we present the difference between the binding energies obtained for each Hamiltonian X and ${}^4\text{DC}$ ($\Delta E_m(X) = E_m(X) - E_m({}^4\text{DC})$, in eV, m a particular edge). The scales are logarithmic, except for the areas in gray for which the scales are linear.

Table 1. MP2 and CCSD Correlation Energies as well as Differences in Correlation Energy between ${}^4\text{DC}$ (E_c and ΔE_c , Respectively, in eV), for Different Hamiltonians^a

		${}^4\text{DC}^{PF}$	${}^2\text{DC}^M$	${}^2\text{DCG}^M$	X2C(a)	X2C(b)
Cl^-	$E_c(\text{MP2})$	-16.12	-16.16	-16.16	-16.16	-16.16
	ΔE_c	$<2 \times 10^{-6}$	-0.04	-0.04	-0.04	-0.04
	$E_c(\text{CCSD})$	-16.43	-16.47	-16.47	-16.47	-16.47
	ΔE_c	$<2 \times 10^{-6}$	-0.04	-0.04	-0.04	-0.04
Br^-	$E_c(\text{MP2})$	-41.44	-41.65	-41.64	-41.65	-41.65
	ΔE_c	$<5 \times 10^{-5}$	-0.21	-0.20	-0.21	-0.21
	$E_c(\text{CCSD})$	-39.65	-39.86	-39.85	-39.86	-39.86
	ΔE_c	$<5 \times 10^{-5}$	-0.21	-0.20	-0.21	-0.21
I^-	$E_c(\text{MP2})$	-48.54	-49.07	-49.05	-49.06	-49.06
	ΔE_c	$<3 \times 10^{-4}$	-0.53	-0.51	-0.52	-0.52
	$E_c(\text{CCSD})$	-46.13	-46.65	-46.64	-46.65	-46.65
	ΔE_c	$<3 \times 10^{-4}$	-0.52	-0.51	-0.52	-0.52
At^-	$E_c(\text{MP2})$	-90.95	-92.90	-92.83	-92.86	-92.87
	ΔE_c	$<3 \times 10^{-3}$	-1.95	-1.88	-1.91	-1.92
	$E_c(\text{CCSD})$	-84.30	-86.24	-86.17	-86.20	-86.20
	ΔE_c	$<3 \times 10^{-3}$	-1.94	-1.87	-1.90	-1.91

^aX2C(a) refers to X2C-AMFI and X2C(b) to X2C-G-AMFI.

to note that for a given edge, the heavier the element, the bigger these reductions are.

4.3. Influence of the Hamiltonian on the Core Ionizations. We now turn our attention to the impact of the Hamiltonian on ionization energies. Our results can be found in Figure 4, where we take the ${}^4\text{DC}$ Hamiltonian as a reference and plot the difference in binding energies between it and the other Hamiltonians considered (see the figure caption for details), as ${}^4\text{DC}$ is the only four-component Hamiltonian without approximations (apart from the treatment of the (SSISS)

integrals, which we address below), we can employ in correlated calculations with DIRAC.

First, we observe a very good match between the ${}^4\text{DC}$ and ${}^2\text{DC}^M$ Hamiltonians across the halogen series—that is, differences in absolute values generally fall below 0.001 eV for the M edges, between around 0.01 and 0.1 eV for the L edges, and are of the order of 0.1 eV for the K edge of all species, except for astatine, for which the difference is 1.67 eV. This latter discrepancy is still small compared to the K edge binding energy of 96 keV.

To better understand how these discrepancies arise, we should recall that in the ${}^2\text{DC}^M$ (${}^2\text{DCG}^M$) approximation, a calculation with the ${}^4\text{DC}$ (${}^4\text{DCG}$) Hamiltonian is first carried out for whatever system we are interested in, in general, a molecule, hence, the denomination “molecular mean-field”, though in Figure 4, we are in effect dealing with atoms. Upon convergence, the transformation into two-component is carried out on the Fock matrix itself (so that the two-component spinor energies correspond exactly to the four-component ones) but the two-electron operator is left untransformed^{79,85} and thus introduces a picture-change error in the electron–electron interaction. In valence only calculations, where core–core and core–valence electron correlation is not accounted for, this yields negligible errors for excitation, ionization, and electron attachment processes.⁸⁴ Our results show that this is still the case for core states of light atoms, such as chloride, and for the M and N edges down to the sixth row.

We can numerically assess the effect of using an untransformed two-electron operator by comparing how the correlation energy (E_c) differs between ${}^2\text{DC}^M$ and ${}^4\text{DC}$ (ΔE_c), as shown in Table 1. The ΔE_c for both MP2 and CCSD yields the same trends; therefore, we shall focus on the simpler MP2 model; as canonical orbitals are employed and their energies are the same for ${}^2\text{DC}^M$ and ${}^4\text{DC}$, any differences between the two come from the picture-change error in the two-electron integrals. We see that for chloride, ΔE_c is still relatively small, then it increases 5-fold going from chloride to bromide, doubles from bromide to iodide, and then increases nearly fourfold from iodide to astatide.

The effect of the two-electron picture-change error can also be seen through a comparison of ${}^4\text{DC}$ to a calculation in which negative-energy solutions of the bare-nucleus one-electron Dirac Hamiltonian are projected out from the molecular spinor space (${}^4\text{DC}^{PF}$), whereby rotations between the negative and positive energy states are effectively eliminated. With this, ${}^4\text{DC}^{PF}$ spinor energies are slightly different from the ${}^4\text{DC}$ ones (see dataset¹²²), but the picture-change error in the two-electron integrals is eliminated, and we would thus expect rather similar correlation energies. This picture is consistent with our numerical results, as correlation energies for ${}^4\text{DC}^{PF}$ are closer to the ${}^4\text{DC}$ ones by at least three orders of magnitude than the ${}^2\text{DC}^M$ ones.

As it will be shown in the following, the sufficiently high accuracy of ${}^2\text{DC}^M$ clearly makes it an asset for applications, due to its reduced computational cost in the index transformation step. However, our results indicate that two-electron picture-change errors are significant enough for deeper cores to require further attention for elements in the sixth row and beyond.

While we do not have ${}^4\text{DCG}$ reference values to which we compare the approximate Hamiltonians that include the Gaunt interaction such as ${}^2\text{DCG}^M$, we expect two-electron picture-change errors to generally follow the same trends discussed above. We, nevertheless, compare ${}^2\text{DCG}^M$ to ${}^4\text{DC}$ in Figure 4 and Table 1, as a way to underscore the importance of the Gaunt interaction for the inner edges. We clearly see that already starting with chloride, we have a non-negligible effect arising from the Gaunt interaction in ${}^2\text{DCG}^M$ that lowers the core binding energy, and can amount to nearly 2 eV for the K edge with respect to the ${}^2\text{DC}^M$ or ${}^4\text{DC}$.

We can also compare ${}^2\text{DCG}^M$ to the approaches in which the transformation to the two-component picture is done before the SCF step (X2C-AMFI and X2C-G-AMFI), as opposed to using the full atomic or molecular potential obtained from the SCF

step as in the ${}^2\text{DCG}^M$ approach. We see that the latter qualitatively follow the changes in binding energy seen for ${}^2\text{DCG}^M$ as we move across the rows, though with significant numerical differences for the K edge (with differences of over 40 eV for iodine and xenon, and over 200 eV for Astatine) as well as for the L and M edges.

Finally, we assess the effect of including the (SSISS) integrals in ${}^2\text{DC}^M$ and ${}^2\text{DCG}^M$ calculations. Since these contributions are generally very small for elements before the fourth row, and they significantly increase the cost of the SCF step, we have only investigated the fourth- and fifth-row atoms (I, Xe, and At) (Table 2).

Table 2. Contributions (in eV) to the Core Binding Energies of I^- , Xe, and At^- from (a) the Gaunt Interaction and (b) the (SSISS) Integral^a

edge	I^-		Xe		At^-	
	$\Delta E(\text{a})$	$\Delta E(\text{b})$	$\Delta E(\text{a})$	$\Delta E(\text{b})$	$\Delta E(\text{a})$	$\Delta E(\text{b})$
K	-81.57	-2.58	-86.75	-2.87	-396.04	-40.50
L ₁	-7.59	-0.39	-8.15	-0.44	-46.24	-7.28
L ₂	-13.64	-0.58	-14.62	-0.65	-78.99	-10.46
L ₃	-9.70	-0.21	-10.39	-0.23	-51.77	-3.56
M ₁	-1.05	-0.08	-1.14	-0.09	-8.72	-1.79
M ₂	-2.08	-0.21	-2.25	-0.12	-15.19	-2.35
M ₃	-1.34	-0.08	-1.46	-0.04	-9.51	-0.89
M ₄	-0.59	-0.06	-0.66	-0.03	-6.62	-0.82
M ₅	-0.26	-0.04	-0.31	0.01	-4.52	-0.09
N ₁					-1.81	-0.44
N ₂					-3.29	-0.55
N ₃					-1.81	-0.19
N ₄					-0.91	-0.10
N ₅					-0.42	0.02
N ₆					6.74	0.06
N ₇					0.48	0.12

^aThese contributions are calculated as the energy difference between CVS-EOM-CCSD calculations employing (a) the ${}^2\text{DCG}^M$ and ${}^2\text{DC}^M$ Hamiltonians and (b) the ${}^2\text{DCG}^M$ without and with the inclusion of the (SSISS) integrals at the SCF step.

For I^- and Xe, these contributions are relatively modest, with reductions of around 2.6 eV and 3 eV to their K edge binding energies, with other significant reductions for the L (0.4–0.7 eV) and M_{1,2} edges (0.1 eV). We note that there is little variation from these atomic values for the molecular systems (see dataset¹²²), in line with the much more localized nature of the small-component density compared to the large component one. Furthermore, we notice a very subtle difference (0.02 eV) between the ${}^2\text{DC}^M$ and ${}^2\text{DCG}^M$ Hamiltonians with (SSISS) integrals.

For At^- , the reductions in binding energies are much more significant: around 40 eV for the K edge, 10 eV for the L₂ edge, and well above 1 eV for the L₃ to M₂ edges. Interestingly, these contributions remain around 0.5 eV for the N edges relating to ionizations from s and p spinors, which are comparable to the energies of the M₁ and M₂ edges for I^- and Xe, which fall between 0.7 and 1 keV.

These results underline the need for explicitly accounting for these integrals (or correcting the energies for their contribution), as soon as we are interested in edges arising from ionizations of s and p spinors for fifth-row elements and beyond.

4.4. Binding Energies: Comparison to Experiment and Prior Theoretical Works. We now compare our results for the

${}^2\text{DC}^{\text{M}}$ and ${}^2\text{DCG}^{\text{M}}$ Hamiltonians, including the (SS)SS integrals, to experiment and other theoretical results for the K and L edges of Xe, XeF₂, and CH₃I, for which recent high energy, gas-phase XPS experiments have been performed (for Xe and XeF₂, results are also available for less energetic edges). Our results are found in Table 3.

Table 3. Comparison between Calculated and Experimental Gas-Phase Binding Energies (in eV) for Xe, XeF₂, and CH₃I^a

edge	model	E(a)	E(b)	E(c)	E(d)
Xe					
K	D/ ⁴ DC	34 755.91	34 690.82	34 567.44	34 563.64
	D/ ² DC ^M	34 755.91	34 690.60	34 567.22	34 563.42
	D/ ² DCCG ^M	34 669.08	34 603.85	34 567.22	34 563.42
	A/ ⁴ DC	34 755.89	34 690.63	34 567.48	34 563.68
	A/ ² DC ^M	34 755.89	34 690.63	34 567.24	34 563.44
	A/ref 127 exp. ¹²⁷	34 752.00	34 690.90	34 567.51	34 563.71
L ₁	D/ ² DC ^M	5509.35	5473.66	5460.49	
	D/ ² DCCG ^M exp. ¹²⁸	5501.17	5465.51	5460.49	5452.7
L ₂	D/ ² DC ^M	5161.45	5122.38	5109.30	
	D/ ² DCCG ^M exp. ¹²⁸	5146.77	5107.77	5109.30	5106.7
L ₃	D/ ² DC ^M	4835.59	4796.85	4787.51	
	D/ ² DCCG ^M exp. ¹²⁸	4825.15	4786.46	4784.51	4786.7
M ₄	D/ ² DC ^M	708.13	689.63	689.01	688.31
	D/ ² DCCG ^M	707.46	688.98	689.01	688.31
	A/ref 127 exp. ¹²⁷	707.5	689.6	688.98	688.28
M ₅	D/ ² DC ^M	694.90	676.71	676.32	675.62
	D/ ² DCCG ^M	694.58	676.30	676.32	675.62
	A/ref 127 exp. ¹²⁷	694.6	676.7	676.08	675.38
XeF ₂					
K	D/ ⁴ DC	34 759.79	34 694.48	34 567.08	34 566.18
	D/ ² DC ^M	34 759.79	34 693.30	34 569.90	34 566.00
	A/ ² DC ^M	34 759.76	34 694.28	34 570.89	34 566.99
	A/ref 127 exp. ¹²⁷	34 755.80	34 694.50	34 571.10	34 567.20
M ₄	D/ ² DC ^M	711.99	693.23	692.61	691.61
	A/ ² DC ^M	711.74	693.32	692.70	691.70
	A/ref 127 exp. ¹²⁹	711.9	693.9	693.28	692.28
M ₅	D/ ² DC ^M	698.49	680.22	679.93	678.93
	A/ ² DC ^M	698.46	680.30	680.01	679.01
	A/ref 127 exp. ¹²⁹	698.6	680.6	680.31	679.31
CH ₃ I					
K	D/ ² DCCG ^M exp. ¹³⁰	33 278.46	33 213.79	33 180.89	33 175.20
L ₁	D/ ² DCCG ^M exp. ¹³⁰	5244.04	5208.55	5203.95	5197.47

^aCalculations are broken down into values obtained with (a) Koopmans theorem, (b) CVS-EOM-CCSD method; (c) CVS-EOM-CCSD with atomic corrections for QED and Breit (in the case of ${}^2\text{DCCG}^{\text{M}}$ results, corrections for the gauge term) interactions; and (d) higher-order correlation corrections by Southworth and co-workers on (c). All of our results include contributions from the (SS)SS integrals. Basis sets A: ANO-RCC and D: Dyll.

As all our calculations lack contributions from QED effects, and ${}^2\text{DCG}^{\text{M}}$ does not include the gauge term that is necessary to recover the full (zero-frequency) Breit interaction, we correct our energies with the results from Koziol and Aucar,¹²⁵ who provided the values for the Breit and leading QED contributions (self-interaction and vacuum polarization) to the atomic spinors of the selected closed-shell atoms (among which is Xe). The same procedure was followed by Southworth et al.¹²⁶ to correct their calculations, which are based on a combination of one-component CVS-EOM-CCSD (and CCSDT) calculations with energy estimates for different effects (1- and 2-electron scalar relativity, spin-orbit coupling, nuclear size effects).

Taking the Xe atom, the QED effects are most important for the K and L₁ edges, lowering the associated binding energies by 42.7 and 5.5 eV, respectively. They are also non-negligible for the L₃ and M₁ edges, which are lowered, respectively, by 0.5 and 1.1 eV. The Breit interaction also lowers the binding energies and, apart from the L₁ edge (for which both are of the same magnitude), is in general much larger than the QED effects (80.7, 13.1, and 8.8 eV for the K, L₂, and L₃ edges, respectively) and still important for the M₃ edge (1.2 eV).

Since Koziol and Aucar only provide the full Breit term, we have estimated the magnitude of the gauge correction from the difference between our ${}^2\text{DCG}^{\text{M}}$ results and their Breit values. From that, we have that the gauge term increases the binding energies (6.04 eV for the K edge, 0.51 eV for the L₁ edge, and around 1.57 and 1.58 eV for the L₂ and L₃ edges, respectively).

We observe that our corrected two-component results differ from the experiment by around 2.1 eV, and if we employ the estimate for higher-order correlation contributions to EOM-CCSD from Southworth et al.,¹²⁶ which decreases the binding energies by 3.8 eV, the difference to experiment is now of -1.71 eV.

We have also investigated the use of uncontracted ANO-RCC basis, used by Southworth et al.,¹²⁷ which is slightly smaller for s and p primitive sets than the Dyll sets. We obtain differences with respect to the experiment of -1.69 eV, which is consistent with our results using the Dyll basis sets. We attribute most of the small differences between ours and prior results to the shortcomings in the treatment of the two-electron interactions in ${}^2\text{DC}^{\text{M}}$ or ${}^2\text{DCG}^{\text{M}}$ discussed previously, as calculations with ⁴DC yield results in very good agreement with those of Southworth et al.

Beyond the K edge, our corrected calculations compare rather well to the recent experimental results of Oura et al.¹²⁸ (5452.7, 5106.7, and 4786.7 eV; measurements carried out at BL29XU of SPring-8 in May 2016) with the exception of the L₁ edge, for which a larger discrepancy to experiment (7.79 eV) is observed. For the M edges, Southworth et al.¹²⁷ present experimental results for the M_{4,5} edges, and our corrected calculations differ from the experiment by -0.94 and -0.85 eV respectively. For the M₄ edge, these are quite comparable to the theoretical calculations of Southworth and co-workers,¹²⁷ now also corrected for QED effects (-0.92 eV to experiment), while for the M₅ edge both theoretical results differ by around 0.24 eV.

Given the values for the higher-order correlation effects for the K and M_{4,5} edges, and from the breakdown of relativistic, correlation, and QED effects from atomic many-body calculations on Xe at the K, L, and M edges¹³¹—which indicate non-negligible differential correlation, relaxation, and other effects (for example Auger shifts) for the different L edges, and to a lesser extent for the M edges—we consider future attempts

to investigate higher-order electron correlation corrections for the L edges to be of significant interest.

For XeF_2 , applying the same corrections as above to our ${}^2\text{DC}^M$ and ${}^2\text{DCG}^M$ calculations in the Dyall basis sets, we arrive at K and $M_{4,5}$ edge binding energies differing by roughly -0.4 eV from the experiment. These differences are smaller but consistent with those obtained for Xe, underscoring the largely atomic nature of these deep core energies. As was the case for the atom, a comparison to ${}^4\text{DC}$ results for the K edge indicates that part of the small differences between the results of Southworth et al. and our two-component ones come from the shortcomings in the treatment of two-electron interactions.

For CH_3I , we have only performed ${}^2\text{DCG}^M$ calculations first because, as illustrated above, there are no significant differences between ${}^2\text{DCG}^M$ - and ${}^2\text{DC}^M$ -based results once we account for QED and Breit or gauge contributions. Second, in the C_s point group used, the coupled-cluster wave functions are complex-valued, making calculations computationally more expensive. We obtained 33213.79 and 5208.55 eV for the K and L_1 edge binding energies, values that overestimate the experimental values by 38.59 and 11.08 eV. For the K edge, this difference is rather close to the one found for the Xe and XeF_2 species.

As QED and Breit corrections are not provided by Koziol and Aucar¹²⁵ for iodine, we have instead used those of Boudjemia et al.,¹³⁰ which amount, for the K edge, to -77.0 eV for Breit, $+6.10$ eV for the gauge term, and -39.0 eV for QED, and to -7.20 , $+0.50$, and -5.10 eV for the L_1 edge. With these corrections, we now overestimate the K and L_1 binding energies by 5.69 and 1.95 eV. The missing effect would be that of higher-order correlation corrections. However, we cannot estimate this here, and we speculate that if it follows roughly what is found for the Xe species, it would decrease the binding energies and likely take the K edge to a few eV.

4.5. Excitation Energies: Comparisons to Experiment and Prior Theoretical Works. As anticipated, the current implementation allows calculation not only of ionization energies but also of excitation energies, which will be briefly discussed hereafter. In this case, since (SSISS) integrals do not significantly affect the energies, only results for the ${}^2\text{DCG}^M$ and ${}^2\text{DC}^M$ Hamiltonian are presented, without explicit inclusion of the aforementioned integrals.

The results are shown in Table 4, which displays selected excitation energies from different edges to the lowest unoccupied molecular orbital (LUMO). We chose these particular transitions since both experimental and computational data are available in the literature,^{126,127} which can be used as reference values. In particular, the experimental values are tabulated with the theoretical energies obtained in this study.

As already observed for the ionization processes, the inclusion of the Gaunt interaction, accounting for the magnetic interaction between the electrons, lowers the excitation energy. This was expected as, in general, the inclusion of this term shifts the orbital energies of the inner core orbitals up (or lowers them, in term of absolute energy) and, at the same time, it reduces the spin-orbit coupling,⁶² which is also reflected by the results reported herein. Indeed, the difference in the excitation energy from $3d_{3/2}$ and $3d_{5/2}$ to LUMO is slightly higher when the ${}^2\text{DC}^M$ Hamiltonian is employed.

Upon inclusion of the Gaunt term in the Hamiltonian, the excitation energies differ by at most 0.4% from the corresponding experimental value, specifically for $1s(\text{F})$. This mismatch is higher than the others, which can be due to the fact

Table 4. XeF_2 ^a

edge	transition	model	E
K	$1s(\text{Xe}) \rightarrow \text{LUMO}$	$D/{}^2\text{DC}^M$	34 686.70
	$D/{}^2\text{DCG}^M$		34 599.95
	exp. ¹²⁶		34 557.4
	$1s(\text{F}) \rightarrow \text{LUMO}$	$D/{}^2\text{DC}^M$	685.91
M_4	$D/{}^2\text{DCG}^M$		685.70
	exp. ¹²⁷		682.8 ± 0.3
	$3d_{3/2} \rightarrow \text{LUMO}$	$D/{}^2\text{DC}^M$	683.71
	$D/{}^2\text{DCG}^M$		683.06
M_5	exp. ¹²⁷		682.8 ± 0.3
	$3d_{5/2} \rightarrow \text{LUMO}$	$D/{}^2\text{DC}^M$	670.34
	$D/{}^2\text{DCG}^M$		670.04
	exp. ¹²⁷		669.9 ± 0.3

^aComparison between the experimental gas-phase excitation energies and calculations (in eV). The calculated excitation energies have been obtained at the CVS-EOM-CCSD level of theory using the ${}^2\text{DC}^M$ and the ${}^2\text{DCG}^M$ Hamiltonians without contributions from the (SSISS) integrals and the Dyall (D) basis set.

that the CVS space was reduced so that this was the highest orbital, necessary to get this excitation.

Thus, this space excludes the 3d orbitals and in particular $3d_{3/2}$. The next higher disagreement between theory and experiment is found when exciting the $1s(\text{Xe})$, although by only 0.1%. Finally, the best agreement is found for both 3d orbitals, with an error of only 0.04% ($3d_{3/2}$) and 0.02% ($3d_{5/2}$) from experiment. The computed energies are thus within the experimental error. To obtain the excitation from the ($3d_{3/2}$) orbital, the same strategy of excluding all higher orbitals from the CVS space, hence excluding ($3d_{5/2}$), might be the reason for the subtly larger disagreement.

Nonetheless, these errors are actually quite small. Therefore, we can conclude that the CVS-EOM-CC method implemented in this work gives satisfactory core-excitation energies.

5. CONCLUSIONS

We have presented an implementation of the core–valence separation for the equation-of-motion coupled-cluster method in the DIRAC program. This implementation, which is based on a flexible framework for defining projection operators, enables the calculation of ionization and excitation energies for all four-component-based Hamiltonians available in DIRAC, and consequently for nonrelativistic Hamiltonians as well. We have applied our implementation to the calculation of core electron binding energies for halogen (CH_3I , X^- , and HX , $\text{X} = \text{Cl}–\text{At}$) and Xe species (Xe , XeF_2). For the latter, we also briefly explored the calculation of core excitations. With these systems, we have investigated the performance of different approximations to the original CVS approach, the basis set effects, and the performance of different classes of two-component approximations.

For highly symmetric species such as the lighter halides, for which we are able to exactly diagonalize the original similarity-transformed Hamiltonian, we show that the CVS energies closely match those from exact diagonalization at all edges. For the overall test set, our assessment of the different approximations indicates the one which more closely matches the performance of the original CVS scheme, which employs both frozen core and the removal of doubly excited determinants containing only core occupied spinors. Taken individually, these approximations yield sizeable overestimations and underestimations to the core-ionization energies, respectively.

With respect to the Hamiltonians, we observe first that the calculations in which the transformation to two-component is performed after the SCF step (${}^2\text{DC}^M$) are nearly indistinguishable from the equivalent four-component ones (${}^4\text{DC}$), though non-negligible discrepancies appear at the K edge of fifth and sixth-row elements. We have traced these discrepancies to the use of the uncorrected 2-electron operator in the two-component molecular mean-field scheme (${}^2\text{DC}^M$ and ${}^2\text{DCG}^M$). More approximate approaches in which the transformation to two-component is carried out before the SCF step (X2C-AMFI and X2C-G-AMFI), on the other hand, offer at best qualitative accuracy. Second, our results underscore the importance of explicitly considering SSSS-type integrals, in particular from the fifth row onwards.

A comparison to experimental results for Xe, XeF₂, and CH₃I underscores the importance of the QED, Breit, and higher-order correlation effects to approach the experimental results. While our calculations, including the Gaunt interaction, recover a significant fraction of the Breit interaction, the gauge term remains quite significant for the K and L edges, and it must be accounted for.

In view of these findings, we consider the four-component-based CVS-EOM-CCSD as a reliable approach for investigating core properties throughout the periodic table. Apart from its intrinsic interest, it may serve as a basis for further investigations of the reliability of more approximate schemes for atoms beyond the fifth row, as well as to verify whether such approximations result in significant changes for properties such as transition moments, that require the determination of the excited-state wave functions.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.0c01203>.

Figures of the effect of the CVS variants on the different binding energies of Xe, the effect of the REW variants on the different binding energies of Cl⁻, Br⁻, I⁻, HCl, HBr, and HI; table illustrating the operation count corresponding to the construction of the σ vector of the EOM-EE and EOM-IP equations using the dyall.acv3z basis set for Cl⁻ and At⁻; and working equations for EOM-EE and EOM-IP σ vectors (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Loïc Halbert – CNRS, UMR 8523—PhLAM—Physique des Lasers, Atomes et Molécules, Université de Lille, F-59000 Lille, France; Email: loic.halbert@univ-lille.fr

Marta L. Vidal – DTU Chemistry—Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark; orcid.org/0000-0003-0653-2078; Email: malop@kemi.dtu.dk

Avijit Shee – Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States; orcid.org/0000-0001-5042-3843; Email: ashee@umich.edu

Sonia Coriani – DTU Chemistry—Department of Chemistry, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark; orcid.org/0000-0002-4487-897X; Email: soco@kemi.dtu.dk

André Severo Pereira Gomes – CNRS, UMR 8523—PhLAM—Physique des Lasers, Atomes et Molécules,

Université de Lille, F-59000 Lille, France; orcid.org/0000-0002-5437-2251; Email: andre.gomes@univ-lille.fr

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.0c01203>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A.S.P.G. and L.H. acknowledge support from the PIA ANR project CaPPA (ANR-11-LABX-0005-01), the Franco-German project ComprIXS (Agence Nationale de la recherche ANR-19-CE29-0019, Deutsche Forschungsgemeinschaft JA 2329/6-1), I-SITE ULNE projects OVERSEE and MESONM International Associated Laboratory (LAI) (ANR-16-IDEX-0004), the French Ministry of Higher Education and Research, region Hauts de France council and European Regional Development Fund (ERDF) project CPER CLIMBIO, and the French national supercomputing facilities (grants DARI A0050801859, A0070801859), as well as discussions with Trond Saue at the 2020 DIRAC Developers' meeting and following the appearance of the manuscript as a preprint. MLV and SC acknowledge support from DTU Chemistry (start-up Ph.D. grant), the Independent Research Fund Denmark—Natural Sciences, Research Project 2, grant no. 7014-00258B, and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie European Training Network Computational Spectroscopy In Natural sciences and Engineering (COSINE), grant agreement no. 765739. Support from the COST action CM1405 MOLIM: Molecules in Motion (short-term mobility grant to M.L.V. to visit A.S.P.G.) is also acknowledged.

■ REFERENCES

- (1) Stöhr, J. *NEXAFS Spectroscopy*; Springer: Berlin, 1992.
- (2) Norman, P.; Dreuw, A. Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.
- (3) *X-Ray Absorption and X-ray Emission Spectroscopy; Theory and Applications*; van Bokhoven, J.; Lamberti, C., Eds.; Wiley & Sons, 2016.
- (4) *Synchrotron Radiation: Basics, Methods and Applications*; Mobilio, S.; Boscherini, F.; Meneghini, C., Eds.; Springer, 2014.
- (5) *X-Ray Free Electron Lasers: Applications in Materials, Chemistry and Biology*; Bergmann, U.; Yachandra, V.; Yano, J., Eds.; Energy and Environment Series 18; Royal Society of Chemistry, 2017.
- (6) Nisoli, M.; Decleva, P.; Calegari, F.; Palacios, A.; Martín, F. Attosecond Electron Dynamics in Molecules. *Chem. Rev.* **2017**, *117*, 10760–10825.
- (7) Milne, C. J.; Penfold, T. J.; Chergui, M. Recent experimental and theoretical developments in time-resolved X-ray spectroscopies. *Coord. Chem. Rev.* **2014**, *277–278*, 44–68.
- (8) Bagus, P. S. Self-Consistent-Field Wave Functions for Hole States of Some Ne-Like and Ar-Like Ions. *Phys. Rev.* **1965**, *139*, A619–A634.
- (9) Bagus, P. S.; Schaefer, H. F. Direct Near-Hartree-Fock Calculations on the 1s Hole States of NO⁺. *J. Chem. Phys.* **1971**, *55*, 1474–1475.
- (10) Besley, N. A. Modeling of the spectroscopy of core electrons with density functional theory. *WIREs Comput. Mol. Sci.* **2021**, No. e1527.
- (11) South, C.; Shee, A.; Mukherjee, D.; Wilson, A. K.; Saue, T. 4-Component relativistic calculations of L₂ ionization and excitations for the isoelectronic species UO₂²⁺, OUN⁺ and UN₂. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21010–21023.
- (12) Zheng, X.; Cheng, L. Performance of Delta-Coupled-Cluster Methods for Calculations of Core-Ionization Energies of First-Row Elements. *J. Chem. Theory Comput.* **2019**, *15*, 4945–4955.

- (13) Lee, J.; Small, D. W.; Head-Gordon, M. Excited states via coupled cluster theory without equation-of-motion methods: Seeking higher roots with application to doubly excited states and double core hole states. *J. Chem. Phys.* **2019**, *151*, No. 214103.
- (14) Besley, N. A.; Gilbert, A. T. B.; Gill, P. M. W. Self-consistent-field calculations of core excited states. *J. Chem. Phys.* **2009**, *130*, No. 124308.
- (15) Liang, Y.; Vinson, J.; Pemmaraju, S.; Drisdell, W. S.; Shirley, E. L.; Prendergast, D. Accurate X-Ray Spectral Predictions: An Advanced Self-Consistent-Field Approach Inspired by Many-Body Perturbation Theory. *Phys. Rev. Lett.* **2017**, *118*, No. 096402.
- (16) Hait, D.; Head-Gordon, M. Highly Accurate Prediction of Core Spectra of Molecules at Density Functional Theory Cost: Attaining Sub-electronvolt Error from a Restricted Open-Shell Kohn–Sham Approach. *J. Phys. Chem. Lett.* **2020**, *11*, 775–786.
- (17) Ehler, C.; Klamroth, T. PSIXAS: A Psi4 plugin for efficient simulations of X-ray absorption spectra based on the transition-potential and Δ -Kohn-Sham method. *J. Comput. Chem.* **2020**, *41*, 1781–1789.
- (18) Derricotte, W. D.; Evangelista, F. A. Simulation of X-ray absorption spectra with orthogonality constrained density functional theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.
- (19) Verma, P.; Derricotte, W. D.; Evangelista, F. A. Predicting Near Edge X-ray Absorption Spectra with the Spin-Free Exact-Two-Component Hamiltonian and Orthogonality Constrained Density Functional Theory. *J. Chem. Theory Comput.* **2016**, *12*, 144–156. PMID: 26584082.
- (20) Hunt, S. J.; Goddard, W. A. Excited States of H₂O using improved virtual orbitals. *Chem. Phys. Lett.* **1969**, *3*, 414–418.
- (21) Ågren, H.; Carravetta, V.; Vahtras, O.; Pettersson, L. G. M. Direct, atomic orbital, static exchange calculations of photoabsorption spectra of large molecules and clusters. *Chem. Phys. Lett.* **1994**, *222*, 75–81.
- (22) Ferré, N.; Assfeld, X. Application of the local self-consistent-field method to core-ionized and core-excited molecules, polymers, and proteins: True orthogonality between ground and excited states. *J. Chem. Phys.* **2002**, *117*, 4119–4125.
- (23) Lundberg, M.; Delcey, M. G. Multiconfigurational Approach to X-ray Spectroscopy of Transition Metal Complexes. In *Transition Metals in Coordination Environments: Computational Chemistry and Catalysis Viewpoints*; Broclawik, E.; Borowski, T.; Radoń, M., Eds.; Springer International Publishing: Cham, 2019; pp 185–217.
- (24) Bokarev, S. I.; Kühn, O. Theoretical X-ray spectroscopy of transition metal compounds. *WIREs Comput. Mol. Sci.* **2020**, *10*, No. e1433.
- (25) Klooster, R.; Broer, R.; Filatov, M. Calculation of X-ray photoelectron spectra with the use of the normalized elimination of the small component method. *Chem. Phys.* **2012**, *395*, 122–127.
- (26) Maganas, D.; Kowalska, J. K.; Nooijen, M.; DeBeer, S.; Neese, F. Comparison of multireference ab initio wavefunction methodologies for X-ray absorption edges: A case study on [Fe(II/III)Cl₄]^{2−/1−} molecules. *J. Chem. Phys.* **2019**, *150*, No. 104106.
- (27) Coriani, S.; Christiansen, O.; Fransson, T.; Norman, P. Coupled-Cluster Response Theory for Near-Edge X-Ray-Absorption Fine Structure of Atoms and Molecules. *Phys. Rev. A* **2012**, *85*, No. 022507.
- (28) Coriani, S.; Koch, H. Communication: X-ray Absorption Spectra and Core-Ionization Potentials within a Core-Valence Separated Coupled Cluster Framework. *J. Chem. Phys.* **2015**, *143*, No. 181103.
- (29) Coriani, S.; Koch, H. Erratum: “Communication: X-ray absorption spectra and core-ionization potentials within a core-valence separated coupled cluster framework” [*J. Chem. Phys.* **143**, 181103 (2015)]. *J. Chem. Phys.* **2016**, *145*, No. 149901.
- (30) Sadybekov, A.; Krylov, A. I. Coupled-cluster based approach for core-ionized and core-excited states in condensed phase: Theory and application to different protonated forms of aqueous glycine. *J. Chem. Phys.* **2017**, *147*, No. 014107.
- (31) Vidal, M. L.; Feng, X.; Epifanovsky, E.; Krylov, A. I.; Coriani, S. New and Efficient Equation-of-Motion Coupled-Cluster Framework for Core-Excited and Core-Ionized States. *J. Chem. Theory Comput.* **2019**, *15*, 3117–3133.
- (32) Peng, B.; Lestrange, P. J.; Goings, J. J.; Caricato, M.; Li, X. Energy-Specific Equation-of-Motion Coupled-Cluster Methods for High-Energy Excited States: Application to K-Edge X-Ray Absorption Spectroscopy. *J. Chem. Theory Comput.* **2015**, *11*, 4146.
- (33) Paul, A. C.; Myhre, R. H.; Koch, H. New and Efficient Implementation of CC3. *J. Chem. Theory Comput.* **2021**, *17*, 117–126. PMID: 33263255.
- (34) Park, Y. C.; Perera, A.; Bartlett, R. J. Equation of motion coupled-cluster for core excitation spectra: Two complementary approaches. *J. Chem. Phys.* **2019**, *151*, No. 164117.
- (35) Matthews, D. A. EOM-CC methods with approximate triple excitations applied to core excitation and ionisation energies. *Mol. Phys.* **2020**, *118*, No. e1771448.
- (36) Brabec, J.; Bhaskaran-Nair, K.; Govind, N.; Pittner, J.; Kowalski, K. Communication: Application of state-specific multireference coupled cluster to core-level excitations. *J. Chem. Phys.* **2012**, *137*, No. 171101.
- (37) Helmich-Paris, B. Simulating X-ray absorption spectra with complete active space self-consistent field linear response methods. *Int. J. Quantum Chem.* **2021**, *121*, No. e26559.
- (38) Schirmer, J.; Thiel, A. An intermediate state representation approach to K-shell ionization in molecules. I. Theory. *J. Chem. Phys.* **2001**, *115*, No. 10621.
- (39) Wenzel, J.; Wormit, M.; Dreuw, A. Calculating Core-Level Excitations and X-Ray Absorption Spectra of Medium-Sized Closed-Shell Molecules with the Algebraic-Diagrammatic Construction Scheme for the Polarization Propagator. *J. Comput. Chem.* **2014**, *35*, 1900.
- (40) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *WIREs Comput. Mol. Sci.* **2015**, *5*, 82–95.
- (41) Neville, S. P.; Averbukh, V.; Patchkovskii, S.; Ruberti, M.; Yun, R.; Chergui, M.; Stolow, A.; Schuurman, M. S. Beyond structure: ultrafast X-ray absorption spectroscopy as a probe of non-adiabatic wavepacket dynamics. *Faraday Discuss.* **2016**, *194*, 117–145.
- (42) Ekström, U.; Norman, P.; Carravetta, V.; Ågren, H. Polarization Propagator for X-Ray Spectra. *Phys. Rev. Lett.* **2006**, *97*, No. 143001.
- (43) Besley, N. A. Density Functional Theory Based Methods for the Calculation of X-ray Spectroscopy. *Acc. Chem. Res.* **2020**, *53*, 1306–1315.
- (44) Stanton, J.; Bartlett, R. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.
- (45) Sneskov, K.; Christiansen, O. Excited state coupled cluster methods. *WIREs Comput. Mol. Sci.* **2012**, *2*, 566–584.
- (46) Bartlett, R. J. Coupled-cluster theory and its equation-of-motion extensions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 126–138.
- (47) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker’s Guide to Fock Space. *Ann. Rev. Phys. Chem.* **2008**, *59*, 433–462.
- (48) Christiansen, O.; Jørgensen, P.; Hättig, C. Response Functions from Fourier Component Variational Perturbation Theory Applied to a Time-Averaged Quasienergy. *Int. J. Quantum Chem.* **1998**, *98*, 1.
- (49) Koch, H.; Jørgensen, P. Coupled Cluster Response Functions. *J. Chem. Phys.* **1990**, *93*, 3333–3344.
- (50) Coriani, S.; Pawłowski, F.; Olsen, J.; Jørgensen, P. Molecular response properties in equation of motion coupled cluster theory: A time-dependent perspective. *J. Chem. Phys.* **2016**, *144*, No. 024102.
- (51) Zuev, D.; Vecharynski, E.; Yang, C.; Orms, N.; Krylov, A. I. New algorithms for iterative matrix-free eigensolvers in quantum chemistry. *J. Comput. Chem.* **2015**, *36*, 273–284.
- (52) Huang, C.; Liu, W.; Xiao, Y.; Hoffmann, M. R. iVI: An iterative vector interaction method for large eigenvalue problems. *J. Comput. Chem.* **2017**, *38*, 2481–2499.
- (53) Huang, C.; Liu, W. iVI-TD-DFT: An iterative vector interaction method for exterior/interior roots of TD-DFT. *J. Comput. Chem.* **2019**, *40*, 1023–1037.

- (54) Stener, M.; Fronzoni, G.; de Simone, M. Time dependent density functional theory of core electrons excitations. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- (55) Tenorio, B. N. C.; Moitra, T.; Nascimento, M. A. C.; Rocha, A. B.; Coriani, S. Molecular inner-shell photoabsorption/photoionization cross sections at core-valence-separated coupled cluster level: Theory and examples. *J. Chem. Phys.* **2019**, *150*, No. 224104.
- (56) Seidu, I.; Neville, S. P.; Kleinschmidt, M.; Heil, A.; Marian, C. M.; Schuurman, M. S. The simulation of X-ray absorption spectra from ground and excited electronic states using core-valence separated DFT/MRCL. *J. Chem. Phys.* **2019**, *151*, No. 144104.
- (57) Delcey, M. G.; Sørensen, L. K.; Vacher, M.; Couto, R. C.; Lundberg, M. Efficient calculations of a large number of highly excited states for multiconfigurational wavefunctions. *J. Comput. Chem.* **2019**, *40*, 1789–1799.
- (58) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-body theory of core holes. *Phys. Rev. A: At., Mol., Opt. Phys.* **1980**, *22*, 206–222.
- (59) Faber, R.; Coriani, S. Core-valence-separated coupled-cluster-singles-and-doubles complex-polarization-propagator approach to X-ray spectroscopies. *Phys. Chem. Chem. Phys.* **2020**, *22*, 2642.
- (60) Nanda, K. D.; Vidal, M. L.; Faber, R.; Coriani, S.; Krylov, A. I. How to stay out of trouble in RIXS calculations within the equation-of-motion coupled-cluster damped response theory framework? Safe hitchhiking in the excitation manifold by means of core-valence separation. *Phys. Chem. Chem. Phys.* **2020**, *22*, 2629.
- (61) Dyall, K. G.; Faegri, K., Jr *Introduction to Relativistic Quantum Chemistry*; Oxford University Press, 2007.
- (62) Saue, T. Relativistic Hamiltonians for Chemistry: A Primer. *ChemPhysChem* **2011**, *12*, 3077.
- (63) Liu, W. Essentials of relativistic quantum chemistry. *J. Chem. Phys.* **2020**, *152*, No. 180901.
- (64) Carbone, J. P.; Cheng, L.; Myhre, R. H.; Matthews, D.; Koch, H.; Coriani, S. Chapter Eleven - An Analysis of the Performance of Coupled Cluster Methods for K-edge Core Excitations and Ionizations using Standard Basis Sets. In *State of The Art of Molecular Electronic Structure Computations: Correlation Methods, Basis Sets and More*; Ancarani, L. U.; Hoggan, P. E., Eds.; Advances in Quantum Chemistry; Academic Press, 2019; Vol. 79, pp 241–261.
- (65) Liu, J.; Matthews, D.; Coriani, S.; Cheng, L. Benchmark Calculations of K-Edge Ionization Energies for First-Row Elements Using Scalar-Relativistic Core-Valence-Separated Equation-of-Motion Coupled-Cluster Methods. *J. Chem. Theory Comput.* **2019**, *15*, 1642–1651.
- (66) Vidal, M. L.; Pokhilko, P.; Krylov, A. I.; Coriani, S. Equation-of-Motion Coupled-Cluster Theory to Model L-Edge X-ray Absorption and Photoelectron Spectra. *J. Phys. Chem. Lett.* **2020**, *11*, 8314–8321.
- (67) Keller, L.; Blum, V.; Rinke, P.; Golze, D. Relativistic correction scheme for core-level binding energies from GW. *J. Chem. Phys.* **2020**, *153*, No. 114110.
- (68) Fillaux, C.; Den Auwer, C.; Guillaumont, D.; Shuh, D. K.; Tyliczszak, T. Investigation of actinide compounds by coupling X-ray absorption spectroscopy and quantum chemistry. *J. Alloys Compd.* **2007**, *444–445*, 443–446.
- (69) Bagus, P. S.; Ilton, E. S.; Nelin, C. J. The interpretation of XPS spectra: Insights into materials properties. *Surf. Sci. Rep.* **2013**, *68*, 273–304.
- (70) Repisky, M.; Komarovskiy, S.; Kadek, M.; Konecny, L.; Ekström, U.; Malkin, E.; Kaupp, M.; Ruud, K.; Malkina, O. L.; Malkin, V. G. ReSpect: Relativistic spectroscopy DFT program package. *J. Chem. Phys.* **2020**, *152*, No. 184101.
- (71) Belpassi, L.; De Santis, M.; Quiney, H. M.; Tarantelli, F.; Storch, L. BERTHA: Implementation of a four-component Dirac-Kohn-Sham relativistic framework. *J. Chem. Phys.* **2020**, *152*, No. 164118.
- (72) Zhang, Y.; Suo, B.; Wang, Z.; Zhang, N.; Li, Z.; Lei, Y.; Zou, W.; Gao, J.; Peng, D.; Pu, Z.; Xiao, Y.; Sun, Q.; Wang, F.; Ma, Y.; Wang, X.; Guo, Y.; Liu, W. BDF: A relativistic electronic structure program package. *J. Chem. Phys.* **2020**, *152*, No. 064113.
- (73) Kutzelnigg, W.; Liu, W. Quasirelativistic theory equivalent to fully relativistic theory. *J. Chem. Phys.* **2005**, *123*, No. 241102.
- (74) Jensen, H. J. A. Douglas–Krollthe Easy Way. Talk at Conference on Relativistic Effects in Heavy Elements—REHE, Mülheim, Germany, April, 2005. Available at <https://doi.org/10.6084/m9.figshare.12046158>.
- (75) Liu, W.; Peng, D. Infinite-order quasirelativistic density functional method based on the exact matrix quasirelativistic theory. *J. Chem. Phys.* **2006**, *125*, No. 044102.
- (76) Iliáš, M.; Saue, T. An infinite-order two-component relativistic Hamiltonian by a simple one-step transformation. *J. Chem. Phys.* **2007**, *126*, No. 064102.
- (77) Peng, D.; Liu, W.; Xiao, Y.; Cheng, L. Making four- and two-component relativistic density functional methods fully equivalent based on the idea of “from atoms to molecule”. *J. Chem. Phys.* **2007**, *127*, No. 104106.
- (78) Liu, W.; Peng, D. Exact two-component Hamiltonians revisited. *J. Chem. Phys.* **2009**, *131*, No. 031104.
- (79) Sikkema, J.; Visscher, L.; Saue, T.; Iliáš, M. The molecular mean-field approach for correlated relativistic calculations. *J. Chem. Phys.* **2009**, *131*, No. 124116.
- (80) Liu, J.; Shen, Y.; Asthana, A.; Cheng, L. Two-component relativistic coupled-cluster methods using mean-field spin-orbit integrals. *J. Chem. Phys.* **2018**, *148*, No. 034106.
- (81) Liu, J.; Cheng, L. An atomic mean-field spin-orbit approach within exact two-component theory for a non-perturbative treatment of spin-orbit coupling. *J. Chem. Phys.* **2018**, *148*, No. 144108.
- (82) Cheng, L. A study of non-iterative triples contributions in relativistic equation-of-motion coupled-cluster calculations using an exact two-component Hamiltonian with atomic mean-field spin-orbit integrals: Application to uranyl and other heavy-element compounds. *J. Chem. Phys.* **2019**, *151*, No. 104103.
- (83) Asthana, A.; Liu, J.; Cheng, L. Exact two-component equation-of-motion coupled-cluster singles and doubles method using atomic mean-field spin-orbit integrals. *J. Chem. Phys.* **2019**, *150*, No. 074102.
- (84) Shee, A.; Saue, T.; Visscher, L.; S.P.Gomes, A. Equation-of-motion coupled-cluster theory based on the 4-component Dirac-Coulomb (-Gaunt) Hamiltonian. Energies for single electron detachment, attachment, and electronically excited states. *J. Chem. Phys.* **2018**, *149*, No. 174113.
- (85) Saue, T.; Bast, R.; Gomes, A. S. P.; Jensen, H. J. A.; Visscher, L.; Aucar, I. A.; Di Remigio, R.; Dyall, K. G.; Eliav, E.; Fasshauer, E.; Fleig, T.; Halbert, L.; Hedegård, E. D.; Helmich-Paris, B.; Iliáš, M.; Jacob, C. R.; Knecht, S.; Laerdahl, J. K.; Vidal, M. L.; Nayak, M. K.; Olejniczak, M.; Olsen, J. M. H.; Pernpointner, M.; Senjean, B.; Shee, A.; Sunaga, A.; van Stralen, J. N. P. The DIRAC code for relativistic molecular calculations. *J. Chem. Phys.* **2020**, *152*, No. 204104.
- (86) Pathak, H.; Sasmal, S.; Nayak, M. K.; Vaval, N.; Pal, S. Relativistic equation-of-motion coupled-cluster method for the electron attachment problem. *Comput. Theor. Chem.* **2016**, *1076*, 94–100.
- (87) Pathak, H.; Sahoo, B. K.; Sengupta, T.; Das, B. P.; Vaval, N.; Pal, S. A relativistic equation-of-motion coupled-cluster investigation of the trends of single and double ionization potentials in the He and Be isoelectronic systems. *J. Phys. B: At., Mol. Opt. Phys.* **2015**, *48*, 115009–115010.
- (88) Pathak, H.; Sasmal, S.; Nayak, M. K.; Vaval, N.; Pal, S. Relativistic equation-of-motion coupled-cluster method for the ionization problem: Application to molecules. *Phys. Rev. A* **2014**, *90*, 062501–062507.
- (89) Pathak, H.; Sahoo, B. K.; Das, B. P.; Vaval, N.; Pal, S. Relativistic equation-of-motion coupled-cluster method: Application to closed-shell atomic systems. *Phys. Rev. A* **2014**, *89*, 042510–042517.
- (90) Pathak, H.; Sasmal, S.; Nayak, M. K.; Vaval, N.; Pal, S. Relativistic equation-of-motion coupled-cluster method using open-shell reference wavefunction: Application to ionization potential. *J. Chem. Phys.* **2016**, *145*, 074110–074118.
- (91) Klein, K.; Gauss, J. Perturbative calculation of spin-orbit splittings using the equation-of-motion ionization-potential coupled-cluster ansatz. *J. Chem. Phys.* **2008**, *129*, 194106–194107.

- (92) Yang, D.-D.; Wang, F.; Guo, J. Equation of motion coupled cluster method for electron attached states with spin-orbit coupling. *Chem. Phys. Lett.* **2012**, *531*, 236–241.
- (93) Wang, Z.; Hu, S.; Wang, F.; Guo, J. Equation-of-motion coupled-cluster method for doubly ionized states with spin-orbit coupling. *J. Chem. Phys.* **2015**, *142*, 144109–144110.
- (94) Epifanovsky, E.; Klein, K.; Stopkowitz, S.; Gauss, J.; Krylov, A. I. Spin-orbit couplings within the equation-of-motion coupled-cluster framework: Theory, implementation, and benchmark calculations. *J. Chem. Phys.* **2015**, *143*, 064102–064117.
- (95) Cao, Z.; Wang, F.; Yang, M. Spin-orbit coupling with approximate equation-of-motion coupled-cluster method for ionization potential and electron attachment. *J. Chem. Phys.* **2016**, *145*, 154110–154113.
- (96) Cao, Z.; Li, Z.; Wang, F.; Liu, W. Combining spin-separated exact two-component relativistic Hamiltonian with equation-of-motion coupled-cluster for spin-orbit splittings of light and heavy elements. *Phys. Chem. Chem. Phys.* **2017**, *19*, 3713–3721.
- (97) Zhang, S.; Wang, F. Excitation Energies of UO_2^+ , NUO^+ , and NUN Based on Equation-of-Motion Coupled-Cluster Theory with Spin–Orbit Coupling. *J. Phys. Chem. A* **2017**, *121*, 3966–3975.
- (98) Akinaga, Y.; Nakajima, T. Two-Component Relativistic Equation-of-Motion Coupled-Cluster Methods for Excitation Energies and Ionization Potentials of Atoms and Molecules. *J. Phys. Chem. A* **2017**, *121*, 827–835.
- (99) Wang, F. Relativistic Equation-of-Motion Coupled-Cluster Theory (EOM-CC). In *Handbook of Relativistic Quantum Chemistry*; Liu, W., Ed.; Springer: Berlin, 2016; pp 1–27.
- (100) Guo, M.; Wang, Z.; Wang, F. Equation-of-motion coupled-cluster theory for double electron attachment with spin–orbit coupling. *J. Chem. Phys.* **2020**, *153*, No. 214118.
- (101) Pathak, H.; Sasmal, S.; Talukdar, K.; Nayak, M. K.; Vaval, N.; Pal, S. Relativistic double-ionization equation-of-motion coupled-cluster method: Application to low-lying doubly ionized states. *J. Chem. Phys.* **2020**, *152*, No. 104302.
- (102) Folkestad, S. D.; Kjønstad, E. F.; Myhre, R. H.; Andersen, J. H.; Balbi, A.; Coriani, S.; Giovannini, T.; Goletto, L.; Haugland, T. S.; Hutcheson, A.; Høyvik, L.-M.; Moitra, T.; Paul, A. C.; Scavino, M.; Skeidsvoll, A. S.; Tveten, Å. H.; Koch, H. eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *J. Chem. Phys.* **2020**, *152*, No. 184103.
- (103) Sorensen, S. L.; Zheng, X.; Southworth, S. H.; Patanen, M.; Kokkonen, E.; Oostenrijk, B.; Travnikova, O.; Marchenko, T.; Simon, M.; Bostedt, C.; Doumy, G.; Cheng, L.; Young, L. From synchrotrons for XFELs: the soft x-ray near-edge spectrum of the ESCA molecule. *J. Phys. B: At., Mol. Opt. Phys.* **2020**, *53*, No. 244011.
- (104) Rankine, C. D.; Penfold, T. J. Progress in the Theory of X-ray Spectroscopy: From Quantum Chemistry to Machine Learning and Ultrafast Dynamics. *J. Phys. Chem. A* PMID: 33729774; DOI: 10.1021/acs.jpca.0c11267.
- (105) Wenzel, J.; Holzer, A.; Wormit, M.; Dreuw, A. Analysis and Comparison of CVS-ADC Approaches up to Third Order for the Calculation of Core-Excited States. *J. Chem. Phys.* **2015**, *142*, No. 214104.
- (106) Visscher, L.; Lee, T. J.; Dyal, K. G. Formulation and implementation of a relativistic unrestricted coupled-cluster method including noniterative connected triples. *J. Chem. Phys.* **1996**, *105*, 8769–8776.
- (107) Pernpointner, M.; Visscher, L. Parallelization of four-component calculations. II. Symmetry-driven parallelization of the 4-Spinor CCSD algorithm. *J. Comput. Chem.* **2003**, *24*, 754–759.
- (108) Shee, A.; Visscher, L.; Saue, T. Analytic Gradient at the 4-component Relativistic Coupled Cluster Level with Inclusion of Spin-Orbit Coupling. *J. Chem. Phys.* **2016**, *145*, No. 184107.
- (109) DIRAC, a relativistic ab initio electronic structure program, Release DIRAC19 (2009) written by Gomes, A. S. P.; Saue, T.; Visscher, L.; Aa, H. J.; Jensen, H. J. A.; Bast, R.; Aucar, I. A.; Bakken, V.; Dyal, K. G.; Dubillard, S.; Ekström, U.; Eliav, E.; Enevoldsen, T.; Faßhauer, E.; Fleig, T.; Fossgaard, O.; Halbert, L.; Hedegård, E. D.; Heimlich-Paris, B.; Helgaker, T.; Henriksson, J.; Iliáš, M.; Jacob, ChR.; Knecht, S.; Komorovský, S.; Kullie, O.; Lærdahl, J. K.; Larsen, C. V.; Lee, Y. S.; Nataraj, H. S.; Nayak, M. K.; Norman, P.; Olejniczak, G.; Olsen, J.; Olsen, J. M. H.; Park, Y. C.; Pedersen, J. K.; Pernpointner, M.; di Remigio, R.; Ruud, K.; Salek, P.; Schimmelpennig, B.; Senjean, B.; Shee, A.; Sikkema, J.; Thorvaldsen, A. J.; Thyssen, J.; van Stralen, J.; Vidal, M. L.; Villaume, S.; Visser, O.; Winther, T.; Yamamoto, S., written by (available at <http://dx.doi.org/10.5281/zenodo.3572669>, see also <http://www.diracprogram.org>).
- (110) Dyal, K. G. Relativistic and nonrelativistic finite nucleus optimized triple-zeta basis sets for the 4p, 5p and 6p elements. *Theor. Chem. Acc.* **2002**, *108*, 335–340.
- (111) Dyal, K. G. Relativistic and nonrelativistic finite nucleus optimized triple zeta basis sets for the 4p, 5p and 6p elements. *Theor. Chem. Acc.* **2003**, *109*, 284.
- (112) Dyal, K. G. Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the light elements H–Ar. *Theor. Chem. Acc.* **2016**, *135*, No. 128.
- (113) Roos, B. O.; Lindh, R.; Malmqvist, P.-A.; Veryazov, V.; Widmark, P.-O. Main Group Atoms and Dimers Studied with a New Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.
- (114) Furry, W. H. On Bound States and Scattering in Positron Theory. *Phys. Rev.* **1951**, *81*, 115–124.
- (115) Hess, B. A.; Marian, C. M.; Wahlgren, U.; Gropen, O. A mean-field spin-orbit method applicable to correlated wavefunctions. *Chem. Phys. Lett.* **1996**, *251*, 365–371.
- (116) Marian, C. M. Spin-Orbit Coupling in Molecules. *Reviews in Computational Chemistry*; Wiley-VCH, 2001; Vol. 17, p 99.
- (117) Schimmelpennig, B. *AMFI, An Atomic Mean-field Spin-Orbit Integral Program*; University of Stockholm: Stockholm, Sweden, 1999.
- (118) Visscher, L. Approximate molecular relativistic Dirac-Coulomb calculations using a simple Coulombic correction. *Theor. Chem. Acc.* **1997**, *98*, 68–70.
- (119) Cheng, L.; Gauss, J.; Stanton, J. F. Treatment of scalar-relativistic effects on nuclear magnetic shieldings using a spin-free exact-two-component approach. *J. Chem. Phys.* **2013**, *139*, No. 054105.
- (120) *Structure of Free Polyatomic Molecules*; Kuchitsu, K., Ed.; Springer: Berlin, 1998.
- (121) Visscher, L.; Styszyński, J.; Nieuwpoort, W. C. Relativistic and correlation effects on molecular properties. II. The hydrogen halides HF, HCl, HBr, HI, and HAt. *J. Chem. Phys.* **1996**, *105*, 1987–1994.
- (122) Halbert, L.; Vidal, M. L.; Shee, A.; Coriani, S.; Gomes, A. S. P. *Dataset: Relativistic EOM-CCSD for core-excited and core-ionized state energies based on the 4-component Dirac-Coulomb(-Gaunt) Hamiltonian 2020*; <https://doi.org/10.5281/zenodo.4116366>.
- (123) Sarangi, R.; Vidal, M. L.; Coriani, S.; Krylov, A. I. On the basis set selection for calculations of core-level states: different strategies to balance cost and accuracy. *Mol. Phys.* **2020**, *118*, No. e1769872.
- (124) Herbst, M. F.; Fransson, T. Quantifying the error of the core-valence separation approximation. *J. Chem. Phys.* **2020**, *153*, No. 054114.
- (125) Koziol, K.; Aucar, G. A. QED effects on individual atomic orbital energies. *J. Chem. Phys.* **2018**, *148*, No. 134101.
- (126) Southworth, S. H.; Dunford, R. W.; Ray, D.; Kanter, E. P.; Doumy, G.; March, A. M.; Ho, P. J.; Krässig, B.; Gao, Y.; Lehmann, C. S.; Picón, A.; Young, L.; Walko, D. A.; Cheng, L. Observing pre-edge K-shell resonances in Kr, Xe, and XeF_2 . *Phys. Rev. A* **2019**, *100*, No. 022507.
- (127) Southworth, S. H.; Wehlitz, R.; Picón, A.; Lehmann, C. S.; Cheng, L.; Stanton, J. F. Inner-shell photoionization and core-hole decay of Xe and XeF_2 . *J. Chem. Phys.* **2015**, *142*, No. 224302.
- (128) Oura, M.; Gejo, T.; Nagaya, K.; Kohmura, Y.; Tamasaku, K.; Journal, L.; Piancastelli, M. N.; Simon, M. Hard x-ray photoelectron spectroscopy on heavy atoms and heavy-element containing molecules using synchrotron radiation up to 35 keV at SPring-8 undulator beamlines. *New J. Phys.* **2019**, *21*, No. 043015.
- (129) Carroll, T. X.; Shaw, R. W., Jr.; Thomas, T. D.; Kindle, C.; Bartlett, N. Electron distribution in the xenon fluorides and xenon oxide

tetrafluoride by ESCA and evidence for orbital independence in the xenon-fluorine bonding. *J. Am. Chem. Soc.* **1974**, *96*, 1989–1996.

(130) Boudjemia, N.; Jänkälä, K.; Gejo, T.; Nagaya, K.; Tamasaku, K.; Huttula, M.; Piancastelli, M. N.; Simon, M.; Oura, M. Deep core photoionization of iodine in CH₃I and CF₃I molecules: how deep down does the chemical shift reach? *Phys. Chem. Chem. Phys.* **2019**, *21*, 5448–5454.

(131) Mooney, T.; Lindroth, E.; Indelicato, P.; Kessler, E. G.; Deslattes, R. D. Precision measurements of K and L transitions in xenon: Experiment and theory for the K, L, and M levels. *Phys. Rev. A* **1992**, *45*, 1531–1543.

Annexe **D**

Implementation of Relativistic
Coupled Cluster Theory for
Massively Parallel GPU-Accelerated
Computing Architectures

Implementation of Relativistic Coupled Cluster Theory for Massively Parallel GPU-Accelerated Computing Architectures

Johann V. Pototschnig,* Anastasios Papadopoulos,* Dmitry I. Lyakh, Michal Repisky, Loïc Halbert, André Severo Pereira Gomes,* Hans Jørgen Aa Jensen,* and Lucas Visscher*



Cite This: <https://doi.org/10.1021/acs.jctc.1c00260>



Read Online

ACCESS |



Metrics & More

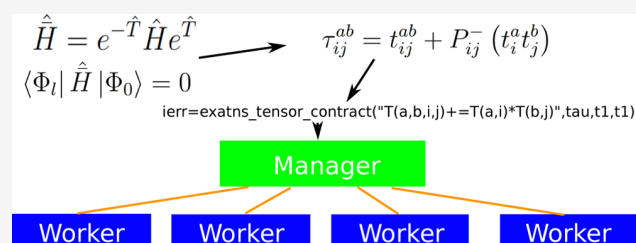


Article Recommendations



Supporting Information

ABSTRACT: In this paper, we report reimplementation of the core algorithms of relativistic coupled cluster theory aimed at modern heterogeneous high-performance computational infrastructures. The code is designed for parallel execution on many compute nodes with optional GPU coprocessing, accomplished via the new ExaTENSOR back end. The resulting ExaCorr module is primarily intended for calculations of molecules with one or more heavy elements, as relativistic effects on the electronic structure are included from the outset. In the current work, we thereby focus on exact two-component methods and demonstrate the accuracy and performance of the software. The module can be used as a stand-alone program requiring a set of molecular orbital coefficients as the starting point, but it is also interfaced to the DIRAC program that can be used to generate these. We therefore also briefly discuss an improvement of the parallel computing aspects of the relativistic self-consistent field algorithm of the DIRAC program.



1. INTRODUCTION

Computational chemistry is a standard tool in the analysis, design, and synthesis of molecular systems.¹ In particular, density functional theory (DFT) is used in a routine fashion in academic and industrial applications. While often sufficiently accurate, DFT does not allow for molecule-specific validations of the accuracy of its predictions. However, this is possible for the wave-function-based methods, such as coupled cluster (CC) theory, for which extensions of the single-particle basis combined with an increase of the excitation level in the CC ansatz lead to a systematic improvement of the accuracy. For organic molecules, CC methods can nowadays predict molecular structures to a precision that is better than 1 pm in bond lengths and better than one degree in bond angles.^{2,3} Furthermore, the efficient equation-of-motion (EOM) treatment of electronically excited states^{4,5} makes it possible to study photochemical processes and aid the interpretation of spectroscopic data. The standard approaches to compute ground-state energies, molecular properties, and electronically excited states have all been generalized to relativistic theory as well, yielding methods that can provide very high accuracy in the electronic structure part of a calculation. This is demonstrated in numerous small-molecule applications^{6–10} for which steep scaling with the system size of the coupled cluster algorithm is not an issue. This rapid increase in computational requirements does, however, in practice, prevent the application of relativistic CC with a fully spin-orbit-coupled reference wave function to systems that contain more than about ten atoms. In nonrelativistic CC computa-

tions, it is, of course, possible to go to larger system sizes, but relativistic corrections can then only be incorporated approximately (usually to some order in perturbation theory and often in combination with effective core potentials). With the current implementation, we want to enable the treatment of larger molecular systems with an all-electron correlated relativistic method that can be used to estimate the accuracy of different approximations for systems with significant correlation and relativistic effects.

Further improvements of the relativistic algorithms are well possible, however, as many reduced-scaling techniques from nonrelativistic algorithms can be taken over in a slightly modified form. One example is the use of density fitting (DF)^{11–13} or Cholesky decomposition^{14–17} to reduce the size of the two-electron integral tensors. Here, relativistic treatments require handling of the density of small components of the Dirac wave functions or, equivalently, fitting of the relativistic correction terms to a two-electron operator in the two-component formulation.¹⁸ Another example is the use of the Laplace transform in Møller–Plesset perturbation theory,¹⁷ where the effects of spin–orbit coupling are visible in the form of (quaternion) imaginary contributions to the density

Received: March 15, 2021

matrices. In both examples, one observes a steep increase in the computational cost of the algorithm but also notes that formal scaling with the system size is identical to that of the nonrelativistic algorithm. Because numerically small contributions to tensor elements can be neglected by the use of screening techniques and many additional terms are only significant in the vicinity of heavy atoms, scaling can in principle be further improved. On the other hand, one may observe that inclusion of only a single heavy atom already presents a challenge due to the number of electrons that has to be correlated in a coupled cluster treatment. This can be illustrated by comparing the CO₂ molecule to the uranyl ion, UO₂²⁺. Both are linear triatomic systems with the oxygen atoms contributing a total of 12 valence electrons. In CO₂, this yields a total of 16 valence electrons that are to be correlated, while in uranyl, one needs to correlate at least 24 electrons¹⁹ and preferably 34 electrons²⁰ due to the large uranium atom.

Both the increase of the number of electrons to be correlated and the switch from real to complex algebra make relativistic calculations rather demanding. However, they are very well suited for deployment on supercomputers because the key algorithms can be formulated as contractions of large tensors, which can be carried out with a relatively high computational efficiency. To be able to realize the full potential of both reduced-scaling techniques and parallel computing, as is nowadays common²¹ in non- or scalar-relativistic approaches, it is advantageous to first create a modern implementation of the relativistic coupled cluster algorithm.

The legacy CC code of DIRAC, RELCCSD,²² allows for parallelization²³ but does not scale well on a larger number of nodes as it was designed for clusters of the early 2000s. An advantage of this code is the use of spatial symmetry, which reduces the computational cost and is helpful in interpreting molecular spinors and electronic transitions. Both aspects are less relevant when applying the coupled cluster approach to large molecular systems that possess (almost) no symmetry. In our reimplementation, we therefore do not consider molecular symmetries but instead focus on data and compute parallelism. The ExaCorr implementation that we describe here is based on the ExaTENSOR library,²⁴ a distributed numerical tensor algebra library for GPU-accelerated HPC platforms developed at the Oak Ridge Leadership Computing Facility (OLCF).

The main body of post-Hartree–Fock quantum–chemical machinery is based on numerical tensor algebra. For the commonly used coupled cluster singles and doubles (CCSD) model, it is possible to formulate²⁵ all operations as tensor contractions of at most four-dimensional tensors. This strict adherence to formulation in terms of tensor contractions is the key to computationally efficient implementation that we present here. It should be regarded both as a platform for future developments and a tool to generate reference data to validate approximate methods in which a large number of two-electron integrals and/or excitation amplitudes are reduced by, for instance, rank reduction^{26,27} or Laplace transformation.²⁸

The paper is structured in the following way. In Section 2, we briefly summarize the coupled cluster algorithms that we consider in the current work. This is followed by Section 3 in which the implementation of the algorithms is discussed. Section 4 is devoted to the details of the computations we used to test the implementation. In Section 5, we present calculations for validation of the correctness of the results by comparing with the reference RELCCSD implementation as well as calculations aimed at showing the computational

scaling. The conclusion follows, which includes a discussion of the follow-up work.

2. THEORY

2.1. Relativistic Theory. A prerequisite for a relativistic coupled cluster calculation is a set of two- or four-component molecular spinors obtained by solving the relativistic Dirac–Hartree–Fock equation. In the four-component case, this equation reads

$$\hat{f}_D \psi = \begin{bmatrix} \hat{V}_{eN} + \hat{J} - \hat{K} & c(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}) - \hat{K} \\ c(\boldsymbol{\sigma} \cdot \hat{\mathbf{p}}) - \hat{K} & \hat{V}_{eN} + \hat{J} - \hat{K} - 2mc^2 \end{bmatrix} \begin{bmatrix} \psi^L \\ \psi^S \end{bmatrix} = E \begin{bmatrix} \psi^L \\ \psi^S \end{bmatrix} \quad (1)$$

in which $\boldsymbol{\sigma}$ is the vector of the 2×2 Pauli matrices and ψ^S and ψ^L are the small and large component parts of the full 4-spinor ψ , respectively. Each of these is itself a 2-spinor with as individual components ψ_α^X and ψ_β^X ($X = L, S$). \hat{V}_{eN} represents the nuclear–electron interaction, usually defined with a Gaussian model of the nuclear charge distribution,²⁹ and the local \hat{J} and nonlocal \hat{K} operators describe the electron–electron interaction in the mean-field approximation

$$\hat{J} = \sum_j^{N_o} \int \psi_j^\dagger(\mathbf{x}_2) \hat{g}(\mathbf{x}_1, \mathbf{x}_2) \psi_j(\mathbf{x}_2) d\mathbf{x}_2 \quad (2)$$

$$\hat{K} \psi_i^X(\mathbf{x}_1) = \sum_j^{N_o} \sum_Y^{L,S} \int \psi_j^\dagger Y(\mathbf{x}_2) \hat{g}(\mathbf{x}_1, \mathbf{x}_2) \psi_i^X(\mathbf{x}_2) d\mathbf{x}_2 \psi_j^Y(\mathbf{x}_1) \quad (3)$$

where N_o denotes the number of occupied spinors. The two-electron interaction operator $\hat{g}(1, 2)$ is the Coulomb (–Gaunt) operator

$$\hat{g}(1, 2) = \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{ij}} \left(-\frac{e^2}{4\pi\epsilon_0} \frac{\boldsymbol{\alpha}_i \cdot \boldsymbol{\alpha}_j}{r_{ij}} \right) \quad (4)$$

Before proceeding to the coupled cluster stage, all operators are transformed using the exact two-component (X2C) method that allows re-expression of the four-component spinors into a two-component picture. Three main variants of the X2C method are used in the current work. The first one, termed X2C-1e, is based on the simple X2C transformation of the one-electron Dirac Hamiltonian that is combined with the nonrelativistic Coulomb operator to describe the electron–electron interactions.^{30,31} Since X2C-1e omits all two-electron relativistic corrections and leaves the relativistic scalar and spin–orbit coupling operators associated with the nuclear potential unscreened, the second variant extends X2C-1e by an explicit addition of the atomic mean-field two-electron potential (done via the AMFI code³²). This approach is termed X2C-AMFI and is the default X2C Hamiltonian in DIRAC. In both X2C approaches, the transformation to the two-component picture is carried out before the Hartree–Fock procedure, and therefore, the two-electron molecular integrals that involve small component basis functions are never computed. In contrast, these types of integrals do enter in the third variant, named X2Cmmf,¹⁸ as the X2C transformation is carried out after solving the Hartree–Fock equations, and therefore, the full molecular potential is used

to define the X2C transformation. This makes X2Cmmf more accurate than the X2C-AMFI (or X2C-1e) approach. Moreover, the two-electron spin-orbit contributions of electrons that will not be explicitly taken into account in the correlation treatment (hereafter referred to as the core or frozen electrons) are treated exactly.

Although (obviously) more expensive due to the mean-field part of the calculations, the X2Cmmf procedure has the same favorable computational characteristics as the X2C-AMFI (or X2C-1e) procedure in the post-HF steps, with the advantage that it yields results that are very close to the full four-component treatment.^{5,18} In the current implementation, the X2Cmmf approach functions as the high-level reference method, while in the DIRAC code, the use of X2C-AMFI and a nonrelativistic treatment (to compare with other coupled cluster implementations) is also supported. Currently, the X2Cmmf approach also allows for an approximate inclusion of the Gaunt interaction,¹⁸ and an implementation of the full Dirac-Coulomb-Gaunt operator for use in very precise benchmark calculations is planned as well.

All of the aforementioned methods apply the no-virtual pair approximation such that the Hamiltonian to be used for the coupled cluster treatment is written in the second quantization as

$$\hat{H} = E^{\text{core}} + \sum_{pq}^{\text{valence}} h_{pq}^{\text{core}} a_p^\dagger a_q + \frac{1}{4} \sum_{pqrs}^{\text{valence}} V_{rs}^{pq} a_p^\dagger a_r^\dagger a_s a_q \quad (5)$$

with V_{rs}^{pq} being the antisymmetrized two-electron integrals

$$V_{rs}^{pq} = \langle pr | \hat{g} | qs \rangle - \langle pr | \hat{g} | ls \rangle \quad (6)$$

and valence in the summation indicating the spinors that are active in the coupled cluster calculation (omitting frozen occupied spinors as well as deleted virtual spinors). The E^{core} constant contains the energy of the core electrons as well as the nuclear repulsion term. The operator h^{core} describes the interaction between the frozen core electrons and the valence electrons and contains the Dirac kinetic energy and nuclear-electron interaction terms defined above as well. The main difference with the nonrelativistic treatments that use an identical expression, is the fact that the tensors h^{core} and \mathbf{V} are defined in complex algebra, whereas in the nonrelativistic treatments, it is usually possible to employ real algebra.

In nonrelativistic quantum chemistry, the Hamiltonian is spin-free, which makes it possible to separate the spatial and spin degrees of freedom and solve equations for spatial orbitals. In relativistic computations, such a separation is not possible because relativistic spinors cannot be written as a simple product of a spatial and spin function. However, in the absence of magnetic fields, one may still use time-reversal symmetry, also known as Kramers symmetry, as each spinor can be related to another with the same energy.^{33,34} Use of this symmetry implies a Kramers-restricted (KR) algorithm in which the occupation of each of the two spinors that comprise a Kramers pair is kept identical in defining the mean-field potential. In contrast, a Kramers-unrestricted (KU) algorithm treats the spinors independent of each other and allows obtaining the so-called spin polarization effect.³⁵

As the use of Kramers symmetry has little advantage in coupled cluster calculations,^{22,36} and we intend to keep the implementation modular and independent of the program used to generate the spinors, we will henceforth assume that all spinors are unrelated to each other, demanding only

orthogonality between them. The consequence of these choices is an increase of the amount of data that needs to be dealt with: in a nonrelativistic restricted CCSD algorithm, the largest tensor appearing in the amplitude equations is of size n_{vir}^4 (with n_{vir} being the number of the virtual spatial orbitals), while in the relativistic treatment, this increases to 16 or 256 n_{vir}^4 complex numbers, with a corresponding increase in memory requirements and the number of floating point operations to generate and contract this tensor.

2.2. Coupled Cluster Algorithms. The wave function in the coupled cluster method is defined as

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}} |\Phi_0\rangle \quad (7)$$

where $|\Phi_0\rangle$ is the single-determinant wave function. The cluster operator \hat{T} is most commonly restricted to the single and double hole-particle excitations

$$\hat{T} = \hat{T}_1 + \hat{T}_2; \quad \hat{T}_1 = \sum_i \sum_a t_i^a a_a^\dagger a_i; \quad \hat{T}_2 = \frac{1}{4} \sum_{ij} \sum_{ab} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i \quad (8)$$

defining the coupled cluster singles and doubles method (CCSD). The energy and cluster amplitudes are computed using equations

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = E \quad (9)$$

$$\langle \Phi_l | \hat{H} | \Phi_0 \rangle = 0; \quad |\Phi_l\rangle = \hat{\tau}_l |\Phi_0\rangle \quad (10)$$

with $\hat{\tau}_l$ denoting a generic excitation operation yielding any (singly, doubly, etc.) excited determinant $|\Phi_l\rangle$, and where the similarity-transformed Hamiltonian

$$\hat{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}} \quad (11)$$

is employed. The working equations for this formalism are well known and can for instance be found in the paper²² describing the RELCCSD program that we used as a reference implementation. In contrast to this code, for the current implementation, we assume that the working memory of the parallel computer is large enough to keep all tensors in memory. Furthermore, we formulate all operations as tensor contractions to enable efficient use of the ExaTensor library. Some intermediates were therefore also altered, resulting in the working equations listed in the [Supporting Information](#). To allow for faster calculations, the CC2 approximation is implemented according to the working equations in the [Supporting Information](#). To speed up the convergence of the CC solver, the direct inversion in the iterative subspace (DIIS) algorithm³⁷ was implemented. Triple excitations are necessary to achieve chemical accuracy, but they are computationally expensive. A widely applied compromise is thus to add them perturbatively.^{22,38,39} The relevant working equations can be found in the [Supporting Information](#).

To obtain first-order molecular electronic properties at the coupled cluster level, we use the Lagrange formalism,⁴⁰ which requires solving the equations for the Lagrange multipliers $\{\lambda\}$

$$L^{\text{CC}}(\mathbf{t}, \lambda) = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle + \sum_l \lambda_l \langle \Phi_l | \hat{H} | \Phi_0 \rangle \quad (12)$$

These are obtained from the stationary conditions

$$\frac{\partial L^{\text{CC}}}{\partial \lambda} = 0 \quad (13)$$

$$\frac{\partial L^{\text{CC}}}{\partial \mathbf{t}} = 0 \quad (14)$$

where eq 13 represents the CC equations. Note that this definition of the Lagrangian neglects orbital relaxation, which is assumed to be partly covered by the \hat{T}_1 operator.⁴¹ Equation 14 is solved to obtain the values of the Lagrange multipliers after which the expectation value of any one-body operator \hat{O} can be computed by computing the one-body density matrix γ^{41}

$$\gamma_p^q = \langle \Phi_0 | (1 + \Lambda) e^{-\hat{T}} a_q^\dagger a_p e^{\hat{T}} | \Phi_0 \rangle \quad (15)$$

$$\langle \hat{O} \rangle = \sum_{p,q} \gamma_p^q o_{pq} \quad (16)$$

The symmetrized one-body density matrix is transformed to the atomic orbital basis and then contracted with the matrix representation of the appropriate property operator \hat{O} . The resulting working equations are listed in the Supporting Information.

3. IMPLEMENTATION

In this part, the details of the implementation are presented. To run coupled cluster computations, molecular spinors for a reference state are required. Section 3.1 describes the computation of these molecular spinors as well as highlights the changes required for improving the performance for larger molecular systems that have become feasible with the new implementation. The implementation utilizes two separate libraries to perform the compute-intensive operations. The calculation of two-electron integrals in the atomic basis is performed by the efficient InteRest library (Section 3.2), while the tensor contractions that comprise the majority of the coupled cluster algorithm are performed with the ExaTENSOR library (Section 3.3). Input handling and interfacing to the self-consistent field (SCF) programs are discussed in Section 3.4, while the transformation of two-electron integrals from the atomic to the molecular basis is described in Section 3.5. Finally, details regarding the coupled cluster code are discussed in Section 3.6.

3.1. Generation of the Molecular Spinors. Molecular spinors are required for the ExaCorr coupled cluster module, and they thus need to be efficiently generated for large system sizes. Because of the fast evaluation of two-electron integrals by the InteRest module, the efficient parallel implementation of the AO-to-MO transformation, and the fast solution of the CC equations described below, for DIRAC calculations, the Fock matrix diagonalizations required in the self-consistent field (SCF) stage became a bottleneck. As this step was not parallelized, it became excruciatingly slow for large AO basis spaces.

Historically, before DIRAC, the well-known double point groups as formulated by Wigner⁴² were used in the pioneering 4c relativistic molecular codes. When the SCF optimization was implemented in DIRAC,^{34,43} we used instead a more general quaternion description, which in fact relies on the simpler (single) point group irreps for quaternion basis function components.⁴³ This implementation has, with small adjustments, been used until work on ExaCorr started, and each SCF-DIIS iteration has thus been based on a direct MPI parallel construction of Fock matrices based on the DALTON implementation,⁴⁴ followed by a sequential quaternion generalization of the Fock matrix diagonalization (see

Appendix D in ref 43). This procedure gave satisfactory scaling with the number of MPI nodes for calculations of up to approximately 2000 AOs, which were feasible with the RELCCSD module. However, the new ExaCorr CC module described in this paper allows for larger applications and significantly large AO spaces and it became paramount that one should be able to do SCF calculations with 1000–5000 AO basis functions (and more in the future) in a small fraction of the wall time needed for the AO-to-MO transformation and the CC calculations. Analyzing the SCF performance for such larger systems with large numbers of compute nodes, it turned out that the parallel Fock matrix construction is acceptably efficient, but it was no surprise that the sequential quaternion matrix diagonalization needed in the MO-based DIIS algorithm required revision.

In this subsection, we describe how this diagonalization bottleneck was removed by tuning the sequential QDIAG code and addition of OpenMP structures in QDIAG. In relativistic quantum chemistry, there are mostly two approaches used for diagonalization. On the one hand, quaternions can be used to get matrix representation in real numbers, which is used here and in a recent publication dealing with large-scale quaternion matrix diagonalization.⁴⁵ On the other hand, complex numbers and routines can be used, which are applied in ReSpect.

The implementation of the QDIAG routines in DIRAC by Saue⁴³ in 1995 was based on his clever quaternion generalization of the complex diagonalization routines in EISPACK, where the EISPACK routines are direct transcriptions of the original ALGOL versions. However, ALGOL just as C and C++ uses row-major storage of matrices, while FORTRAN uses column-major storage. Therefore, the EISPACK routines were very inefficient for larger matrices because of many cache misses caused by the large strides in memory. A necessary first step was consequently to rewrite the QDIAG routines by transposing the access to all matrices followed by improvements of the logical structure. This change by itself already caused significant improvement in the sequential performance. The resulting implementation was then suitable for OpenMP parallelization. Initial timings of a large application on the TITAN supercomputer that was performed with 800 cores indicated that OpenMP parallelization with just eight OpenMP threads was sufficient to reduce the time spent in diagonalization to less than 11 min, compared to an overall wallclock time of 66 min in one SCF iteration (outputs of this and other benchmark runs are provided in a separate repository⁴⁶). Additional timings on the SUMMIT supercomputer also showed that the wall time spent in diagonalization is much less than that needed for other steps like Fock matrix construction, and therefore, it was deemed unnecessary to also program additional GPU and/or MPI parallelization.

3.2. InteRest Integral Library. In the ExaTENSOR library (described below), it is possible to call an external library to initialize a particular tensor with the desired values. This mechanism allows for efficient parallel computation of the electron repulsion integrals (ERIs). A prerequisite is, however, that this external library is sufficiently modular, a requirement that could not be met by the legacy HERMIT integral generator used in DIRAC. We therefore interfaced the InteRest library⁴⁷ to enable parallel computation of the ERIs arising from relativistic and nonrelativistic theories.

As discussed in ref 35, all commonly applied basis types in relativistic calculations are of multicomponent spinor nature

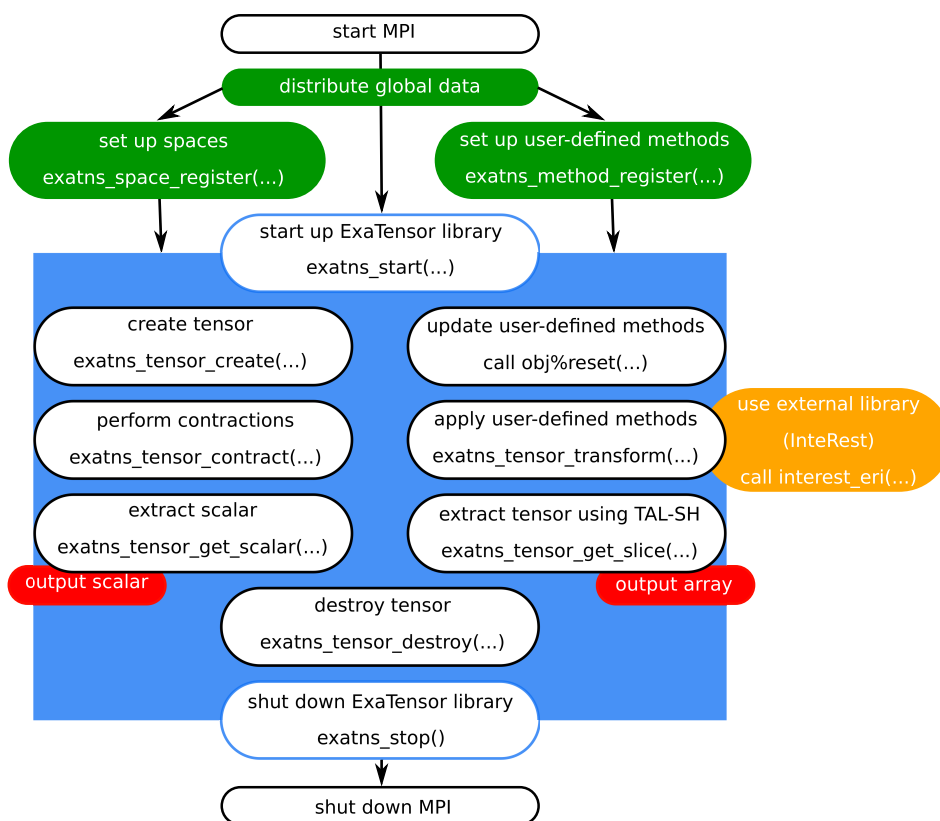


Figure 1. ExaCorr computational workflow based on the ExaTensor library.

and can uniformly be formulated in terms of real quaternion functions ($\mathbb{H}_{\mathbb{R}}(\mathbb{R}^3)$) or complex quaternion functions ($\mathbb{H}_{\mathbb{C}}(\mathbb{R}^3)$) over the field of real numbers \mathbb{R} . A product of any two quaternion basis functions $X_a(\vec{r})$ and $X_b(\vec{r})$ defines the so-called quaternion overlap distribution function $\Omega_{ab}(\vec{r}) \equiv X_a^*(\vec{r})X_b(\vec{r})$ in terms of which one can formulate and design an efficient algorithm for evaluation of nonrelativistic and relativistic ERIs.³⁵ For instance, if Ω refers to a restricted kinetically balanced (RKB) basis,⁴⁸ then $\Omega \in \mathbb{H}_{\mathbb{R}}(\mathbb{R}^3)$ comprises four real quaternion components. Then, a single quadruplet of ERIs defined similarly to the nonrelativistic case as

$$[\Omega_{ab}|\Omega_{cd}] \equiv \iint \frac{\Omega_{ab}(r_1)\Omega_{cd}(r_2)}{r_{12}} dr_1 dr_2 \quad (17)$$

requires the evaluation and processing of 25 times more real scalar integrals than in the nonrelativistic case. InteRest utilizes the Obara–Saika integration technique over Cartesian Gaussians⁴⁹ to compute all of these scalar integrals in parallel and groups them into four integral classes [LLLL], [SSLL], [LLSS], and [SSSS] according to their values, which gradually decrease in powers of c^{-2} .³⁵ At the expense of going from real to complex quaternion functions, the presented uniform formalism for relativistic ERI evaluation can also be applied in the solid-state domain.⁵⁰ Additional basis requirements needed for magnetic property calculations, such as the restricted magnetic balance⁵¹ (RMB) in combination with the gauge-including atomic orbitals (RMB-GIAO),⁵² can also be handled with the discussed integral scheme. A thorough discussion on this topic is given in ref 35.

3.3. ExaTensor and TAL-SH Backends. The ExaCorr module provides two distinct implementations of coupled cluster methods, one intended for execution on a single shared-memory node with an optional GPU acceleration and another one for execution on many such nodes (distributed parallelism), thus supporting a broad variety of computer platforms, from simple workstations to leadership HPC systems. Both implementations use the ExaTensor library²⁴ as a massively parallel GPU-accelerated processing backend for numerical tensor algebra operations, although there are some differences in the interface between the single- and multinode API. For the single-node runs (with OpenMP multithreading and/or GPU acceleration), only the single-node component of ExaTensor, the TAL-SH library,⁵³ is used. The single-node implementation is more efficient when MPI parallelization is not needed. It also serves as a validation reference for the corresponding multinode implementation. The ExaTensor library is written in a mix of Fortran-2008 and C/C++. It depends on BLAS, LAPACK, OpenMP, CUDA, and MPI (MPI is not necessary for single-node runs, while CUDA is only necessary for the GPU-enabled builds).

Figure 1 outlines the computational workflow where the ExaCorr module offloads all computationally expensive operations (primarily tensor contractions) to the ExaTensor library. Essentially, the high-level interface of the ExaTensor library allows for the creation, destruction, addition, and contraction of distributed tensors via a single API call per operation, thus making it possible to directly translate tensor equations into the library calls. Such direct translation of quantum many-body equations into a human-readable code drastically accelerates the implementation of new coupled cluster methods in the ExaCorr module. Additionally,

ExaTENSOR provides API for user-defined transformations on distributed tensors, which are often necessary in the coupled cluster algorithms. As described in Figure 1, the general computational workflow of a coupled cluster method implemented in ExaCorr starts with a replication of some global data, like molecular spinor coefficients, diagonal elements of the Fock matrix, etc., which normally do not consume much memory. Then, all vector spaces necessary for defining many-body tensors are explicitly registered, such as the space of atomic orbitals, occupied molecular spinors, virtual molecular spinors, etc. After that, all necessary ExaCorr-specific unary tensor transformations are registered as well. These extensions of the ExaTENSOR library are implemented as extensions of an abstract tensor transformation class provided by the ExaTENSOR interface. Once this is done, the ExaTENSOR parallel runtime (domain-specific virtual processor²⁴) is started within a provided MPI communicator. After initialization, ExaTENSOR will begin accepting commands to perform distributed tensor algebra operations that realize a given coupled cluster algorithm. Importantly, the ability to implement user-defined tensor transformations facilitates the use of external libraries within ExaTENSOR, for example, the InteRest library,⁴⁷ which was straightforwardly integrated with ExaTENSOR to enable parallel computation of the Coulomb integrals. Finally, once the given coupled cluster workload has been executed to completion, a local copy of the resulting scalar (e.g., energy, property) or tensor (e.g., density matrix) of interest can be retrieved. At the very end, the ExaTENSOR parallel runtime is explicitly shut down and control is handed back to the stand-alone ExaCorr or DIRAC program.

In ExaTENSOR, a tensor is formally defined as a multi-indexed vector from a linear space constructed as a direct product of basic (single-indexed) vector spaces. Such a multi-indexed vector (tensor) is represented by an array of complex numbers. The number of basic spaces in the direct product space defines the order of tensors living in that direct product space (note that in physics the tensor order is usually called the tensor rank). Each tensor dimension is thus associated with a specific basic space (or its subspace) from the defining direct product space. In practice, one must explicitly register all necessary basic vector spaces by calling the `exatns_space_register` API function provided by ExaTENSOR. To construct a basic vector space, one simply needs to provide a basis for that space or just specify its dimension. One can also construct a subspace of a registered basic vector space, thus enabling construction of tensor slices. Importantly, the definition of basic vector spaces requires their splitting into a number of subspaces chosen by the user, thus inducing the splitting of tensors into tensor slices. These slices are called elementary tensor blocks. All tensors are stored as distributed collections of such elementary tensor blocks. In the current implementation, the segment size used for splitting a basic vector space into a direct sum of its subspaces can be controlled by a keyword (see the Supporting Information).

Another prerequisite of coupled cluster algorithms is the necessity of custom tensor transformations (or initializations), like initialization of the Coulomb integral tensor, import of pre-existing many-body tensors (e.g., the Fock matrix), Jacobi preconditioning during amplitude updates, etc. Each such tensor initialization or transformation can easily be injected into ExaTENSOR by implementing user-defined tensor classes extending the abstract class `tens_method_uni_t`,

followed by their registration with `exatns_method_register`. These user-defined subclasses can further be classified as either static or dynamic. The objects of static subclasses do not change their internal state after the registration with the ExaTENSOR runtime, whereas the objects of dynamic subclasses are allowed to change their internal state after the registration, thus enabling further flexibility and dynamic behavior during the execution of a tensor algorithm.

Once all necessary basic vector spaces/subspaces and user-defined tensor methods have been pre-registered, one may proceed to the execution of actual tensor operations on distributed tensors. A tensor is created via calling `exatn_tensor_create`, where a user provides which space/subspace each tensor dimension is associated with. Inside the ExaTENSOR parallel runtime, each tensor is recursively decomposed into smaller slices, which are distributed across all nodes (the tensor decomposition is induced by the direct sum structure of the vector spaces defining tensor dimensions). A tensor can then be initialized to either a scalar value or some custom value via a user-defined initialization method (`exatns_tensor_init`). There are three main tensor operations currently provided by ExaTENSOR: user-defined unary tensor transformation (`exatns_tensor_transform`), tensor addition (`exatns_tensor_add`), and tensor contraction (`exatns_tensor_contract`). These are sufficient for implementing the majority of coupled cluster algorithms. Both tensor addition and tensor contraction API take symbolic strings specifying the addition/contraction index pattern, for example

$$S(a, b, i, j) += V(a, b, c, d) * T(c, d, i, j)$$

for a partial contraction over indices c and d , or

$$E() += V + (a, b, i, j) * T(a, b, i, j)$$

for a full contraction over all indices in which the complex conjugate values of the tensor V are used (indicated by the $+$ symbol). Note that ExaTENSOR can handle arbitrary permutations of indices in tensor contraction specifications. Explicit tensor reordering is usually not necessary but can be achieved with the following permuted tensor addition specification

$$A(a, b, c, d) += B(b, a, d, c)$$

This reordering is employed in the creation of the antisymmetrized ERIs of eq 6 after the AO-to-MO transformation is completed.

In principle, tensor operations submitted to ExaTENSOR are processed asynchronously by all available MPI processes but one can also invoke bulk synchronization by calling `exatns_sync` to ensure the completion of all outstanding computations (a global barrier). Once all necessary computations have been completed, one can retrieve a local copy of the computed scalar (e.g., energy, property) via `exatns_tensor_get_scalar`. If one needs a slice of some computed tensor instead (e.g., density matrix), `exatns_tensor_get_slice` will return a local copy of the requested tensor slice. All created tensors need to be explicitly destroyed once no longer needed via `exatns_tensor_destroy`.

For the sake of completeness, let us briefly discuss the current parallelization algorithm used by ExaTENSOR for a distributed execution of tensor contractions on many GPU-accelerated HPC nodes. We should immediately note that the

current algorithm is not communication optimal and also has other inefficiencies that we are currently addressing. In the current work, however, our goal was simply to extend the size of molecules that can be treated with relativistic coupled cluster theory via utilizing large-scale GPU-accelerated HPC platforms. The work on optimizing our parallel algorithms and their execution will follow in the near future.

The ExaTENSOR task-based parallel runtime consists of two types of MPI processes: managers and workers. Managers accept incoming tensor instructions (tensor operations, e.g., tensor contractions) and decompose them into smaller pieces (tasks), recursively. The tensor operation decomposition is induced by the decomposed structure of the participating tensors, determined during tensor creation. At the lowest level of task granularity, the generated elementary tensor operations are distributed across all workers (specialized MPI processes). The current task distribution algorithm is based on data affinity (tasks gravitate toward the workers owning the largest operand) and dynamic load balancing. The latter is replaced with static load balancing for tensor contractions with large output tensors due to the inefficiency of the MPI_Accumulate operation. Finally, the elementary tensor operations (tasks) received from managers are executed by workers, which includes allocation of memory resources, remote data prefetch or accumulation, and actual execution on either a multi-threaded CPU or one or more GPUs in a fully asynchronous manner.

All computational aspects discussed above are taken care of by the ExaTENSOR parallel runtime and cannot be changed by the user of ExaCorr. Job-specific tuning and optimization of the parallelization are, however, possible by setting environment variables and/or specific keywords in the input. This provides control over the amount of memory used on a single node, whether GPUs are to be used, how OpenMP threads are distributed and mapped to CPU cores, etc.

3.4. Molecular Spinors: Interface to DIRAC and ReSpect. The ExaCorr module was designed with modularity in mind, so it would be easy to interface with other quantum-chemical packages. For convenience, we currently use the build infrastructure of DIRAC, but the code can also be compiled and used as a stand-alone program since the minor dependencies on some specific modules of DIRAC can be easily removed.

ExaCorr requires two files with information to be present: a job input file and a file containing information about the molecular spinors. A complete diagram of the interface is depicted in Figure 2.

The input file (exacc.inp) contains the options controlling the coupled cluster computations and should at least include the definition of the active occupied and active virtual spinor spaces. Spinors outside this active space are considered as belonging to the frozen core (for the occupied spinors) or as deleted (for the virtual spinors). In the following, we will consider the occupied and virtual spaces as pertaining to these (potentially reduced) subspaces of the full spinor spaces defined in the molecular spinor file. Examples for additional options are convergence thresholds, choice of coupled cluster wave functions (CCD/CCSD/CC2/CCSD(T)), a switch to enable the computation of the density matrix, and several more technical keywords. A complete list can be found in the Supporting Information. These options can also be set in the DIRAC input (dirac.inp) if the ExaCorr module is called directly from DIRAC.

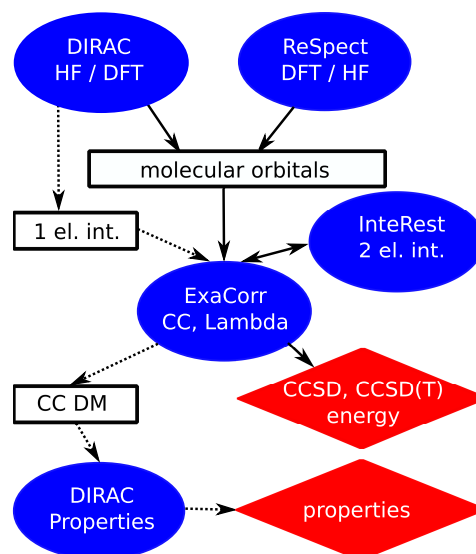


Figure 2. Workflow of ExaCorr computations; details can be found in the text.

The second file can either be DIRAC's molecular spinor file, DFCEOF, or the RSD_MOS file from ReSpect.³⁵ These interface files contain three different sets of data defining the canonical molecular spinors: (i) information about the basis set, (ii) the coefficients of the molecular spinors, and (iii) the spinor energies thereof for the Fock matrix expression used in the generating SCF procedure. An optional input file (MRCONEE) containing one-electron integrals can be generated by the MOLTRA module in DIRAC. This additional data can be used to recompute the Fock operator for open-shell cases, for which the DIRAC definition,⁵⁴ used to define the spinor energies, differs from the simple KU formalism assumed in ExaCorr. Results of the CC calculations are provided in the form of a text output file and an effective density matrix, in case the lambda equations are solved as well. This density matrix can be used by DIRAC to compute a wide range of molecular properties. As DIRAC assumes a KR formalism, the latter type of calculation is currently limited to Kramers symmetric (closed-shell) systems.

3.5. Index Transformation Algorithms. For relativistic calculations in which the size of the AO space is usually an order of magnitude larger than the MO -space, the transformation of the two-electron integrals from the atomic to the molecular basis can amount to a significant fraction of the overall computational expense. There are different approaches to implement these transformations differing in memory requirements and operation count. In ExaCorr, the current default is the standard Yoshimine⁵⁵ scheme with n^5 scaling, which reads for the Coulomb interaction in the X2C models as

$$(p\lambda_{\sigma_1}|\mu\nu) = \sum_{\kappa}^{n_{\text{AO}}} C_{\kappa\sigma_1 p}^* (\kappa_{\sigma_1} \lambda_{\sigma_1} |\mu\nu) \quad (18)$$

$$(pq|\mu\nu) = \sum_{\lambda}^{n_{\text{AO}}} \sum_{\sigma_1} C_{\lambda\sigma_1 q} (p\lambda_{\sigma_1} |\mu\nu) \quad (19)$$

$$(pq|rv_{\sigma_2}) = \sum_{\mu}^{n_{\text{AO}}} C_{\mu\sigma_2 r}^* (pq|\mu_{\sigma_2} \nu_{\sigma_2}) \quad (20)$$

$$\langle pq|rs \rangle = \sum_{\nu} \sum_{\sigma_2}^{n_{\text{AO}}} C_{\nu\sigma_2}(pq|r\nu_{\sigma_2}) \quad (21)$$

where $p, q, r,$ and s are the molecular spinors; $\kappa, \lambda, \mu,$ and ν are the spatial atomic orbitals; and σ_1 and σ_2 denote the spin for electrons 1 and 2, respectively. In this procedure, we make use of the fact that the AO spinors are defined as simple products of spatial and spin functions so that the spin integration reduces to additional summations in the second and fourth steps of transformation. The antisymmetrization and reordering, $\langle pr||qs \rangle = \langle pq|rs \rangle - \langle ps|rq \rangle$, is done after the index transformation is completed. By making use of permutational symmetry, only six unique classes of molecular integrals are used, which can be generated on the basis of three classes of half-transformed integrals (vv, vo, and oo, where v and o stand for virtual and occupied spaces, respectively). The number of atomic orbitals (n_{AO}) can become quite large, which makes it impractical to work with the complete atomic two-electron integral tensor ($\kappa\lambda|\mu\nu$). The default in the current code is to slice the last index and work with a subspace thereof. This does not increase the operation count of the algorithm and is in practice sufficient to reduce the memory footprint due to handling of the AO integral tensor. This choice has the benefit of keeping large spaces for the other indices, making the tensor contractions optimally efficient.

3.6. Coupled Cluster Implementation. The general approach for applying the ExaTensor library is outlined in Figure 1. After reading and processing the input data, basis set information, spinor energies, MO coefficients and, optionally, the one-electron integrals are stored as global variables and broadcasted to all nodes using MPI. Three different spaces are defined, for the atomic orbitals, the occupied spinors, and the virtual spinors. In addition, the ExaCorr-specific methods are registered in the ExaTENSOR library interface. Apart from the already-mentioned ERI generation by InteRest, this, e.g., comprises methods to initialize a tensor with MO coefficients, initialize a tensor with one-electron integrals, scale a tensor with denominators, or project on a subspace. After these preliminary steps, the ExaTENSOR library is started and MO integrals are computed by transforming the ERIs to the MO basis. The MP2 amplitudes are subsequently computed to obtain an initial guess for the CC amplitudes. These CC amplitudes are then refined in an iterative procedure, the working equations for which can be found in the Supporting Information or ref 25. As the convergence of the nonlinear coupled cluster iterations can be slow, we have implemented the DIIS scheme³⁷ and are also assessing the less memory-demanding CROP algorithm.⁵⁶ In the current implementation, all necessary tensors are created before the iterative procedure starts, allowing for a priori assessment of the maximum memory footprint of the run.

Triple excitations require tensors of the size $n_{\text{occ}}^3 n_{\text{vir}}^3$. For the full triples, these amplitudes need to be determined iteratively, which requires a significant amount of memory and number of operations. For the perturbative treatment of the triple excitations considered here, memory requirements can be reduced by splitting the occupied space and using three nested loops over these subspaces (of dimension n_{red}) to evaluate all contributions. This results in a memory requirement of $n_{\text{red}}^3 n_{\text{vir}}^3$ in addition to the memory required for the coupled cluster amplitudes, the two-electron integrals, and the Fock matrix. Permutational symmetry is used to speed up the computation

by only computing unique blocks. The equations for the triples corrections are listed in the Supporting Information.

For the calculation of molecular properties, the equations for the Lagrange multipliers λ have to be solved, which can be done in much the same way as described above for the CC equations, including the use of DIIS to reduce the number of iterations. The λ and amplitude tensors are then combined according to the equations in the Supporting Information (see also ref 41) to obtain the one-particle density matrix. In the case of the TAL-SH implementation, the tensor elements can be accessed directly and written to file. For ExaTENSOR, a local copy is first created in the form of a TAL-SH tensor, which is then written. The properties module in DIRAC can read these data and compute the properties.

4. COMPUTATIONAL DETAILS

All calculations were performed with development versions of the DIRAC⁵⁷ and ExaCorr software package; details on the particular revisions used in the calculations described below are apparent in the respective output files that are provided within a separate repository.⁴⁶ The geometries of the systems are also included in this collection.

The reference orbitals and single-determinant reference wave functions have mostly been obtained by the SCF implementation in DIRAC, which is a Kramers-restricted implementation. To enforce Kramers symmetry for systems that have an odd number of electrons or have near-degeneracies at the Fermi level, an average-of-configuration (AOC) approach is used in DIRAC.^{54,57,58} In contrast, the ReSpect code performs Kramers-unrestricted (KU) calculations, in which the Kramers symmetry is not imposed.³⁵ For the utilization of spinors generated by the AOC procedure in DIRAC, some additional features are needed for the interface, as the definition of the Fock matrix in DIRAC differs from the KU Fock matrix, the definition assumed in ReSpect and ExaCorr. For closed-shell molecules, the difference between the AOC and KU Fock matrix expressions disappears and spinor energies can be read in from the DIRAC program and are sufficient to define the reference Hamiltonian. For open-shell molecules, one may either employ the ReSpect code or another code that has a compatible KU Fock matrix definition or recomputes the Fock matrix during the CC stage of the calculation. Both cases result in the use of a KU Hartree–Fock expression for a given reference determinant that is chosen by the user of the program. This is important for perturbation treatments, because the diagonal of the Fock matrix is then used to define the zeroth-order Hamiltonian.

Unless otherwise noted, we employed uncontracted Dyall basis sets of double (dyall.v2z), triple (dyall.v3z), or quadruple (dyall.v4z) zeta quality.^{59–61} The set of spinors included in the correlated calculations generally consists of a subset of the total set of spinors. By default, these are selected by energy thresholds corresponding to relatively high-lying occupied and low-lying virtuals, with energies between -10 and 20 hartree.

In the case of lanthanide monofluorides (LnFs) and the uranium hexafluorides (UF₆) dimer, geometries were optimized at DFT level with the ADF⁶² code using the scalar-relativistic ZORA Hamiltonian,⁶³ the Perdew–Burke–Ernzerhof (PBE) functional,⁶⁴ and triple zeta basis sets with one polarization function (TZP).

Further, molecule-specific, computational details are listed below.

4.1. Lanthanide Monofluorides. For LaF and YbF, the AOC-SCF approach was applied using DIRAC either employing X2C-AMFI^{30,32} or the one-component nonrelativistic (NR) Hamiltonian. In the case of EuF, KU calculations were performed with ReSpect³⁵ using the one-component nonrelativistic or X2C-1e Hamiltonian.³¹ Several thresholds for the occupied and virtual spinors were considered for the double zeta basis set, and the same values were employed for larger basis sets.

4.2. Argon Binding to Gold. For argon atoms bound to gold clusters, we employed the X2Cmmf Hamiltonian,¹⁸ with the structures being taken from ref 65. The default energy thresholds in the coupled cluster step have been employed. The numbers of correlated electrons for the systems considered are 46 (AuAr⁺), 60 (AuAr₂⁺), 74 (AuAr₃⁺), 88 (AuAr₄⁺), and 102 (AuAr₅⁺).

4.3. Uranium Hexafluoride Dimer. In the case of UF₆ and (UF₆)₂ calculations, the X2Cmmf Hamiltonian^{30,31} was applied, except for some smaller scaling investigations for which we used X2C-AMFI.^{30,32} In these smaller computations, the cc-pVDZ and dyall.v2z basis sets were selected for F and U, respectively. The larger, more accurate computations employed the corresponding triple zeta basis sets. The energy thresholds for the included spinors were −35 and 80 hartree for smaller computations and −10 and 8 hartree for the triple zeta ones.

In the computations, the distance between uranium and fluorine in a monomer was fixed to an experimental value of 1.996 Å.⁶⁶ A restricted optimization using these monomers was performed for (UF₆)₂, and the structures were applied in the coupled cluster computations. In this case, we added the dispersion correction by Grimme to the PBE functional. Additionally, an optimization without restrictions was performed for the UF₆ dimer at the DFT level to estimate the U–F bond distances for this level of theory.

4.4. Uranyl Tris-nitrate Complex. Calculations of the *zz* component of electric field gradient (EFG) at the U nucleus (*q_{zz}*) for the uranyl tris-nitrate ([UO₂(NO₃)₃][−]) complex have been performed at the X-ray structure for the RbUO₂(NO₃)₃ crystal,⁶⁷ employing the X2C-AMFI^{30,32} and taking into account the picture change of the EFG operator. In addition to CC, we have performed DFT calculations with the B3LYP, PBE0, and CAMB3LYP density functionals.

For the property CC calculations, we considered occupied spinors with energies higher than or equal to (a) −6 hartree (106 electrons, in which the U 5d is correlated as done for other uranyl complexes⁶⁸), (b) −22 hartree (156 electrons, in which the U 4f and all electrons for the light atoms are correlated), and (c) −4500 hartree, which amounted to correlating all 202 electrons. These three occupied spaces are combined with virtual spinors with energies up to and including (a) 5 hartree for both double and triple bases (543 and 649 virtuals, respectively); for only double zeta bases, (b) 20 hartree (680 virtuals), (c) 50 hartree (818 virtuals), (d) 150 hartree (896 virtuals); and for triple zeta bases, (e) 7 hartree (944 virtuals). The total numbers of virtual spinors are 1286 and 2076 for double and triple zeta bases, respectively. We have not performed CCSD calculations with quadruple zeta bases.

5. RESULTS AND DISCUSSION

Our first goal was to verify the correctness of the new implementation. To do so, we compared the results of the new

implementation using TAL-SH or ExaTENSOR to the results obtained by the RELCCSD implementation in DIRAC.^{22,57,58} Comparisons of the energies for H₂O, LiO, and CuAr_{*n*}⁺ are in the Supporting Information. To check the property implementation, we compared the dipole moment, EFG, and the nuclear quadrupole coupling constant (NQCC) of CHFCIBr and UF₆ for different implementations, which can also be found in the Supporting Information; the output files are provided in a separate repository.⁴⁶

A few systems were selected to show the capabilities of the new implementation in the investigation of heavy-element systems. First, we consider the ionization energies of three lanthanide monofluorides (LnF) species, LaF, YbF, and EuF since, from a methodological perspective, these calculations allow us to demonstrate the usage of our implementation for both closed- and open-shell configurations. Second, the binding of argon atoms to gold cations has been studied including triples corrections, which are necessary to achieve chemical accuracy. In the subsequent section, results for uranium hexafluoride and its dimer are presented as well as some information about the scaling of the new code. Finally, the electric field gradient of the uranyl tris-nitrate complex was computed as an example for evaluation of electronic properties in a larger molecule.

5.1. Ionization Energies of Lanthanide Monofluorides. Lanthanides are often treated using density functional theory, but results are shown to have a strong dependence on the exchange–correlation functional that is selected.⁶⁹ Coupled cluster theory can provide more accurate and precise results and has been applied in conjunction with more approximate methods to account for relativistic effects, like the one-component Douglas–Kroll Hamiltonian⁷⁰ and effective core potentials.⁷¹ The current implementation, and in particular the interface for the ReSpect code, provides a way to investigate the (generally open-shell) ground states for such systems with full inclusion of relativistic and core correlation effects in coupled cluster theory.

Before proceeding to the discussion of our results for the ionization energies themselves, we shall discuss the requirements, in terms of the number of occupied and virtual spinors necessary for obtaining reliable results. For this, we have decided to consider two sets of equilibrium structures, one for the neutral and the other for the ionized species. Our structures, obtained at the DFT level using the PBE functional, are shown in Table 1, together with experimental values and

Table 1. Experimental Reference Values and Structures Used in the Computations for Lanthanide Monofluorides and DFT Bond Distances Applied in the Computations

	<i>r_e</i> (exp.)	<i>r_e</i> (PBE)	<i>r_e</i> (cation, PBE)	<i>r_e</i> (ECP, CCSD(T)) ⁷¹
LaF	2.0234 ⁷²	2.0293	2.0150	2.0215
EuF	2.083 ⁷³	2.0676	1.9992	2.0750
YbF	2.016516 ⁷⁴	1.9868	1.9345	2.0204

prior theoretical values. For these systems, DFT produces the experimentally observed trend, with EuF having the largest bond distance and YbF the smallest. As expected, there are some deviations as well, with the DFT bond distances being smaller than the experimental values for EuF and YbF but slightly larger for LaF.

Now, concerning the coupled cluster calculations themselves, we first investigated the convergence of the energies

with the number of active occupied and virtual spinors. The reason for such an investigation is that employing the complete set of virtual spinors is typically not needed in relativistic calculations of heavy elements. This is due to the use of uncontracted basis sets, which leads to a significant number of virtual spinors being mostly located in the chemically inactive core region. These types of spinors can be deleted without affecting the results much. In the current work, we identify such spinors by a simple energy criterium, relying on the observation that the large kinetic energy of these solutions puts them in the upper range of energies obtained by Fock matrix diagonalization. More advanced schemes, such as use of approximate natural orbitals are also possible and under development. Regarding the choice of occupied spinors to be included, one needs to take into account that for lanthanides, the closeness (in radial extent) of the open-shell 4f and other electrons that would otherwise be considered as core (4s–4d) may require that they are correlated alongside the (5s, 5p) valence.

We present in Table 2 the results of such an investigation for the YbF, which had previously been investigated by some of

Table 2. Ionization Potential in eV of YbF for Different Numbers of Correlated Spinors Employing the dyall.v2z Basis Set^a

threshold _{low}	threshold _{high}	nocc	nvir	% occ	% vir	CCSD
–20	2.3	49	89	63	21	4.49
–20	6	49	137	63	32	5.89
–20	150	49	267	63	63	5.89
–20	10 000	49	367	63	86	5.89
–60	10	61	155	78	36	5.90
–60	20	61	195	78	50	5.90
–60	150	61	267	78	63	5.90
–3	40	31	213	40	50	5.95
–20	40	49	213	63	50	5.90
–40	40	51	213	65	50	5.90
–60	40	61	213	78	50	5.90
–400	40	77	213	99	50	5.90
exp						5.91 ± 0.05 ⁷⁵

^aThe number of occupied and virtual spinors refers to the neutral molecule; for the cation, one of these occupied spinors becomes a virtual spinor. The Δ SCF ionization potential computed using the reference determinant wave functions was 5.48 eV. The spinor thresholds are listed in atomic units.

us⁶¹ and which was found to be particularly sensitive to electron correlation treatment. We provide equivalent tables for LaF and EuF as the Supporting Information due to the fact that these exhibit the same trends as discussed below.

From Table 2, we can identify two main trends: (a) Employing a too small virtual space (comprising around 21% of the total number of virtuals), even with a fairly large number of occupied, yields a (strong) underestimation of the ionization energy at the CCSD level (–1.41 eV). A modest increase in the number of virtuals (including around 30% of the virtuals) greatly reduces this underestimation and brings values closer to the experimental value. Further increases in the number of virtuals past 60% yield no significant difference in the CCSD ionization energies. (b) Employing a converged virtual space (>30%) but not enough occupied spinors overestimates the ionization energies, though not by much (around +0.05 eV). Possible choices for the occupied space are to correlate only

F(2s²2p⁶) and Yb(5s²5p⁶4f¹⁴) (comprising 40% of the occupied space) or to include the Yb(4s²4p⁶4d¹⁰) shells as well (63% of the occupied space).

For reliable results, we find that the following electrons need to be correlated: La(4s²4p⁶5s²4d¹⁰5p⁶), Eu-(3s²3p⁶4s²3d¹⁰4p⁶4d¹⁰5s²5p⁶6s¹4f⁷), and Yb-(4s²3d¹⁰4p⁶5s²4d¹⁰5p⁶6s¹4f¹⁴), which can be achieved by employing energy thresholds of –20, –200, and –60 hartree for LaF, EuF, and YbF, respectively. The 2s² and 2p⁶ of F are always included, and the 1s² is omitted for LaF. In the case of Yb, the neutral molecule is an open-shell system, while the cation only has closed shells. The opposite is true for La. EuF was considered as an example with a high-spin state; the neutral molecule as well as the cation has several open shells.

Following our analysis of what are the minimum requirements in terms of occupied and virtual spinors for obtaining converged ionization energies, we investigate the adiabatic and vertical ionization potentials for the three molecules, computed for the basis set of increased quality, for two classes of Hamiltonians (nonrelativistic and X2C). The values are listed in Table 3.

The largest change of the ionization potential is due to the inclusion of relativistic contributions. Regarding LaF, the ionization potential is larger by about 1.3 eV (CCSD) or 1.4 eV (SCF) for the X2C-1e Hamiltonian than that for the nonrelativistic one. For EuF/YbF, the changes are somewhat smaller; increases of about 0.4/0.5 and 0.3/0.4 eV were obtained for the coupled cluster and Δ SCF, respectively.

The inclusion of electron correlation by CCSD results in an increase of the ionization potential by about 1, 0.4, and 0.5 eV for LaF, EuF, and YbF, respectively. These are the values for the X2C-1e Hamiltonian; the changes in the nonrelativistic case are similar. The perturbative triples corrections are the smallest for EuF, probably due to the relatively simple high-spin ground states of both the neutral and the cation, which can be well described by the Kramers-unrestricted reference wave functions obtained with ReSpect. They increase the ionization potential by less than 0.06 eV. The triples add about 0.1 eV in the case of LaF for X2C-1e and slightly more for the nonrelativistic Hamiltonian. While the triples in the nonrelativistic case are similar (between 0.1 and 0.2 eV) for YbF, the values for X2C are much larger. The fourth-order correction increases the IP by about 1.7 eV, the fifth-order correction results in values about 0.7 eV below the CCSD ones. A similar observation has been reported in ref 61. These large values are an indication that a perturbative inclusion of triples is insufficient, in agreement with the large T_1 values, see Table 3. This is probably caused by a mixing of excited states with closed and open f-shells, which was observed to cause a large change in ground-state polarizability of the Yb atom⁷⁸ and a change in the nuclear quadrupole coupling constant in ref 79.

For an increase of the basis set size, the ionization potential of the reference determinant becomes smaller except for the vertical transition of LaF using the nonrelativistic Hamiltonian. Going from double to triple zeta basis sets, the coupled cluster ionization potential increases for YbF and EuF, although the changes are below 0.03 eV in the latter case. Regarding LaF, a decrease of the ionization potential is observed at the coupled cluster level.

The adiabatic IP should be smaller than the vertical one if the equilibrium distances are correct. For the HF reference, this is never the case as the electron correlation is missing in

Table 3. Vertical (a) and Adiabatic (b) Ionization Energies (in eV) for the Lanthanide Monofluorides^a

		basis	ΔSCF		CCSD		T1	CCSD+T		CCSD(T)		CCSD-T		
			(a)	(b)	(a)	(b)		(a)	(b)	(a)	(b)	(a)	(b)	
LaF	NR	2z	3.41	3.44	5.03	5.06	0.01	5.15	5.18	5.11	5.14	5.10	5.13	
		3z	3.43	3.44	4.71	4.71	0.02	4.90	4.91	4.84	4.84	4.82	4.74	
		4z	3.43	3.43	4.61	4.60	0.02	4.81	4.81	4.74	4.74	4.72	4.72	
		∞z	3.42	3.43	4.53	4.53		4.74	4.74	4.66	4.66	4.65	4.64	
	X2C	2z	4.93	4.96	5.91	5.93	0.01	5.96	5.99	5.96	5.98	5.96	5.98	
		3z	4.87	4.87	5.87	5.87	0.01	5.98	5.98	5.96	5.96	5.96	5.95	
		4z	4.86	4.86	5.87	5.86	0.01	5.99	5.99	5.97	5.97	5.97	5.96	
	exp.						6.00	5.99	5.98	5.97	5.97	5.96		
EuF	NR	2z	4.73	4.75	5.08	5.12	0.01	5.10	5.15	5.10	5.15	5.11	5.15	
		3z	4.72	4.74	5.10	5.13	0.01	5.13	5.17	5.13	5.17	5.13	5.17	
	X2C	2z	5.04	5.07	5.46	5.51	0.01	5.51	5.57	5.51	5.56	5.51	5.56	
		3z	5.02	5.05	5.48	5.52	0.01	5.54	5.58	5.53	5.57	5.53	5.57	
		exp.						5.9 ± 0.3 ⁷⁷						
	YbF	NR	2z	5.04	5.09	5.39	5.39	0.01	5.43	5.43	5.42	5.43	5.42	5.43
			3z	4.93	4.98	5.40	5.43	0.01	5.47	5.49	5.45	5.47	5.46	5.48
4z			4.93	4.98	5.42	5.44	0.01	5.59	5.60	5.56	5.57	5.57	5.58	
∞z			4.93	4.98	5.44	5.45		5.67	5.68	5.64	5.65	5.65	5.66	
X2C		2z	5.48	5.49	5.90	5.87	0.05	6.51	6.44	5.57	5.56	5.67	5.65	
		3z	5.44	5.46	6.00	5.98	0.09	7.76	7.73	5.30	5.29	5.76	5.75	
		4z	5.44	5.46	6.00	5.97	0.05	6.82	6.78	5.57	5.56	5.75	5.74	
	exp.						6.13	6.08	5.77	5.75	5.75	5.73		
							5.91 ± 0.05 ⁷⁵							

^aThe active ranges for LaF, EuF, and YbF are -20 to 40 , -200 to 200 , and -60 to 40 hartree, respectively. The X2C Hamiltonians used were X2C-AMFI for LaF and YbF (spinors obtained with the DIRAC code) and X2C-1e for EuF (spinors obtained with the ReSpect code). The complete basis set limit values (∞z) have been obtained with a two-point extrapolation formula based on the 3z and 4z values.

contrast to the DFT that was used to determine the bond distances. In the case of the coupled cluster, the correct order of the vertical and adiabatic IP is obtained for the X2C-1e Hamiltonian and triple or quadrupole zeta basis set, indicating that the DFT bond distances are close. EuF always shows the wrong order, probably because the bond distances are not accurate enough, which is probably also reflected in the larger differences between the theoretical and experimental values in Table 1.

The best estimates from Table 3 are 5.97, 5.57, and 5.97 eV for LaF, EuF, and YbF, respectively, while values of 6.3, 5.9, and 5.91 eV were determined in experiments (Table 1). One of the reasons for these discrepancies is the large experimental uncertainties (especially for LaF and EuF), while also the neglect of zero point energies in our values will play a role. Considering these sources of errors, the energies show an acceptable agreement.

Systems with open shells, like treated above, can be difficult to describe using coupled cluster since CC is based on a single reference determinant. The t1 amplitudes recover a portion of the static correlation, which makes treatment possible when there is one dominant determinant, but in cases with several important configurations, multireference methods are necessary.⁸⁰ A related difficulty appearing in a two-component treatment is that the spinors are no longer eigenfunctions of \hat{S}_Z . At the SCF level, this can be handled by an average-of-configuration approach,⁵⁴ which occupies all relevant configurations, resulting in spinors with varying spin-up and spin-down contributions. Making a proper selection of such spinors to form a single-determinant reference wave is, however, difficult in the general case. The exception is cases with only a single unpaired electron in which either spinor of the singly

occupied Kramers pair can be taken to construct the reference determinant. In the current work, molecules were selected that are still rather easy to treat. YbF and LaF have only one unpaired electron and thus belong to this important special case of simple open-shell molecules. For neutral and positively charged EuF, one determinant can qualitatively correctly describe the ground states if we allow for a KU spinor optimization that is able to converge to a “high-spin” state. This is possible with ReSpect.

While the use of Kramers restriction in an averaged SCF is feasible for simple open-shell systems, it does lead to an inconsistency in the definition of the reference Fock operator and orbital energies between the KR SCF program and the KU CC implementation. This is formally not a problem as our working equations do not require the use of a diagonal Fock operator but make working with denominators consisting of spinor energy differences between occupied and virtual spinors more complicated. With unmodified orbital energies, the energy difference between the highest occupied spinor (one of the two open-shell spinors) and the lowest unoccupied spinor (its Kramers partner) would be zero. There are several ways to deal with this complication. One is to recompute the spinor energies according to the KU Fock matrix expression. This will induce an energy gap and make it possible to apply denominators. This simple approach was applied to the results for LaF and YbF presented in Table 3.

5.2. Binding Energies of Argon Atoms to Gold Atoms.

Gold is one of the most nonreactive metals in the periodic table, and noble gases are also exceptionally inert. Nevertheless, the AuNe⁺ dimer was reported in 1977⁸¹ and early computations suggested a covalent bond between gold cations and noble gas atoms,⁸² which is supported by recent

Table 4. Total Binding Energy (ΔE , in eV) of AuAr_n Systems with dyall Basis Sets of Different Cardinal Numbers^a

system	basis	V	ΔE					
			HF	MP2	CCSD	CCSD + T	CCSD(T)	CCSD - T
AuAr^+	2z	136	-0.1341	-0.5260	-0.4620	-0.5267	-0.5165	-0.5156
	3z	230	-0.1401	-0.5993	-0.4699	-0.5519	-0.5408	-0.5400
	4z	400	-0.1456	-0.6408	-0.4817	-0.5689	-0.5586	-0.5581
	∞z		-0.1496	-0.6711	-0.4904	-0.5814	-0.5716	-0.5713
AuAr_2^+	2z	168	0.0845	-1.2292	-1.0356	-1.1981	-1.1723	-1.1706
	3z	288	-0.0219	-1.4124	-1.0890	-1.2813	-1.2563	-1.2550
	4z	508	-0.0399	-1.4938	-1.1161			
	∞z		-0.0531	-1.5532	-1.1360			
AuAr_3^+	2z	200	-0.0140	-1.3416	-1.1491	-1.3100	-1.2875	-1.2860
	3z	346	-0.0816	-1.5202	-1.1899	-1.3902	-1.3669	-1.3652
	4z	616	-0.1019	-1.6124	-1.2268			
	∞z		-0.1167	-1.6797	-1.2538			
$\text{AuAr}_4^+(3\text{D})$	2z	232	-0.1061	-1.4673	-1.2681	-1.4267	-1.4090	-1.4078
	3z	404	-0.1367	-1.6474	-1.3009	-1.5092	-1.4890	-1.4872
	4z	724	-0.1590	-1.7527	-1.3491			
	∞z		-0.1753	-1.8295	-1.3843			
$\text{AuAr}_4^+(2\text{D})$	2z	232	-0.1605	-1.4326	-1.2530	-1.4024	-1.3863	-1.3852
	3z	404	-0.1783	-1.5967	-1.2769	-1.4750	-1.4563	-1.4545
	4z	724	-0.2001	-1.6978	-1.3243			
	∞z		-0.2160	-1.7715	-1.3588			
AuAr_5^+	2z	264	-0.2217	-1.5945	-1.3898	-1.5398	-1.5301	-1.5750
	3z	462	-0.2059	-1.7730	-1.4126	-1.6238	-1.6100	-1.6081

^aIn all cases, spinors with energies between -10 and 20 hartree have been included in the correlation treatment. The number of virtual spinors in each case (V) is shown; see Section 4 for the number of correlated electrons for each species. The complete basis set limit values (∞z) have been obtained with a two-point extrapolation formula based on the 3z and 4z values.

experimental results.⁸³ A theoretical study observed strengthening of the binding in small gold clusters if noble gas atoms are attached.⁸⁴ This covalency is in part attributed to the relativistic nature of the heavy Au; this makes it necessary to include these contributions in theoretical studies. Recently, the significant influence of argon atoms on the IR spectra and bonding of small gold clusters was observed.⁸⁵ Here, we want to compute the interaction of a single gold atom with argon using the reliable coupled cluster method in combination with the X2Cmmf Hamiltonian. A summary of the current state of research on noble gas-noble metal compounds can be found in a recent review.⁸⁶

First, the energy of the AuAr dimer was computed for different internuclear distances. The equilibrium bond distances for the AuAr dimer were determined by fitting a Morse potential to about 5 points. For the dyall.v4z basis set, equilibrium bond distances of 2.50, and 2.47 Å were obtained by CCSD and CCSD(T), respectively. The MP2 value is about 0.1 Å smaller than the CCSD one, and the HF value is about 0.3 Å larger. The triples correction reduces the equilibrium distance by about 0.03 Å, a detailed table is in the Supporting Information. This general trend is observed for all basis sets, while the bond distance is about 0.05 smaller for 3z than for 2z. The structures of the larger systems were obtained by density functional theory.⁶⁵ Coupled cluster binding energies are listed in Table 4. There is a strong dependence on the method; Hartree–Fock underestimates the CCSD binding energies by 0.4–1.5 eV, MP2 overestimates them by 0.04–0.4 eV. The triples correction are also significant; they increase the binding energy by 0.16/0.20 eV for the fourth-order +T and 0.14/0.18 for the fifth-order (T)/-T considering the dyall.v2z/dyall.v3z basis set, excluding the AuAr dimer with smaller triples contributions. The dimer constitutes a special case as the

structures were optimized at the coupled cluster level. For this reason, significant HF binding energies were obtained as they are computed for longer bond distances as the CC ones. If the basis set is increased or extrapolated, the CCSD energy increases by about 0.04 eV, except for the dimer with smaller changes. The growth of the CCSD+T/(T)/-T energy is about 0.08 eV in going from the double to triple zeta basis set, excluding the AuAr dimer. The energy per argon atom reaches its maximum for AuAr_2^+ with about 0.78 eV at the CBS CCSD level of theory. For AuAr_4^+ , two structures have been computed to assess the relative stability of a planar and a 3D arrangement. Independent of the basis set and method, the three-dimensional structures are found to be lower in energy.

These preliminary findings will be incorporated in a larger investigation of Ar bound to gold clusters in conjunction with infrared multiphoton dissociation experiments.^{87,88} For such investigations, it is essential to be able to have reliable benchmarks of DFT calculations, which will become possible with this new implementation,

For these systems, it is possible to compare the performance of the new implementation to the RELCCSD code in DIRAC.^{22,57,58} In Table 5, the timings for the three parts of the coupled cluster computation are listed, the integral transformation, the CCSD iterations, and the computation of the perturbative triples. The new ExaCorr implementation is always faster than the RELCCSD implementation without symmetry. Including the symmetry significantly reduces the computational time, especially for systems with linear symmetry (AuAr^+ , AuAr_2^+). In these systems, the integral transformation of RELCCSD with symmetry takes about the same time as in ExaCorr, while the CCSD iterations take less time. However, once the system has less symmetry (AuAr_3^+ , Cs), ExaCorr is significantly faster than RELCCSD for which

Table 5. Wall Time in Seconds for the New Implementation and the RELCCSD Release 2019 Reference Implementation, with Symmetry (R19s) and without Using Symmetry (R19)^a

step	system	nodes	R19	R19s	ExaCorr
integrals	AuAr ⁺	4	371	71	106
	AuAr ⁺	8	308	70	58
	AuAr ₂ ⁺	8	715	87	77
	AuAr ₃ ⁺	8	1602	810	133
	AuAr ₃ ⁺	20	1450	703	87
CCSD	AuAr ⁺	4	327	7	84
	AuAr ⁺	8	232	6	71
	AuAr ₂ ⁺	8	1129	7	73
	AuAr ₃ ⁺	8	2512	849	183
	AuAr ₃ ⁺	20	1585	586	143
triples	AuAr ⁺	4	2055	42	1555
	AuAr ⁺	8	3130	46	1294
	AuAr ₂ ⁺	8	>5200	90	2666
	AuAr ₃ ⁺	20	>5200	>5200	5821

^aAll computations were performed using the dyall.v2z basis set. There are 243/42 atomic orbitals, 32/14 occupied and 104/32 virtual spinors per gold/argon atom.

calculation of the triples correction turned out to be infeasible for this system.

5.3. Binding Energy of Uranium Hexafluoride Dimers.

Uranium hexafluoride is used in the gaseous form in enrichment methods for nuclear fuels. To simulate the behavior of this gas under different conditions, an accurate description of the intermolecular interaction potential is important. Early attempts to describe the interaction of molecules were based on potentials derived from thermophysical data and spectroscopy.^{89,90} In quantum chemistry, the properties^{91–94} and reaction pathways⁹⁵ of the monomer have been mainly studied employing relativistic DFT. To describe the interaction of two such units, it is important to account for relativistic⁹¹ as well as dispersion effects accurately. As the electronic structure of the dimer is not problematic and well described by a single reference determinant, coupled cluster theory can be used to provide accurate reference data. Since the computations are rather expensive, due to the number of electrons that needs to be correlated, this particular system is well suited for testing our implementation.

First, we performed computations for the UF₆ monomer on different numbers of nodes. Tables 6 and 7 display the

Table 6. Time in Seconds for Integral Transformation (t_I) and Solving the Coupled Cluster (t_{CCD} , t_{CCSD}) and Λ Equations (t_Λ) for UF₆ Using the dyall.v2z Basis Set^a

n	t_I (75)	t_I (50)	t_{CCD} (75)	t_{CCD} (50)	t_{CCSD} (75)	t_Λ (75)
16	841	1179	843	1314	2046	
24	654	902	711	1045	1869	1865
32	505	764	617	925	1645	1649
48	413	686	594	861	1512	1493
64	375	634	600	810	1545	1412

^aThe CCD, CCSD, and Λ equations, took resp. 10, 20, and 21 iterations to solve. For the selected thresholds of -35 to 80 hartree, 110 occupied and 474 virtual spinors are included. n is the number of Summit nodes. The number in parentheses in the header is the segment size used by ExaTENSOR for chunking the occupied and virtual vector spaces.

Table 7. Time in Seconds for Integral Transformation (t_I) and Coupled Cluster (t_{CCSD}) for UF₆ Applying the dyall.v3z Basis Set^a

n	t_I (75)	t_{CCSD} (75)
32	1685	1279
64	1191	1205
96	817	1081
128	687	973

^aThe CCSD equations took 21 iterations to solve. For thresholds of -10 to 8 hartree, 70 occupied and 554 virtual spinors are active. n is the number of Summit nodes. The number in parentheses in the header is the segment size used by ExaTENSOR for chunking the occupied and virtual vector spaces.

obtained timings for the double and triple zeta basis set, respectively. As evident from Table 6, our code scales up to 48 nodes for such a small system before the ExaTENSOR worker processes begin to starve and/or have load balancing issues. Naively, the total number of tasks as well as load balancing could in principle be improved with finer task granularity, which can be achieved by decreasing the dimension segment size from 75 to 50, but the increased inter- and intranode communication overheads then in fact lead to an overall increase of the computational time, as can be seen in Table 6. The better performance and scaling of large tensor contractions enhance the difference between the CCD and CCSD formalisms. While the additional tensor contractions related to the inclusion of single excitations are at most of order n^5 , CCSD iterations took noticeably more time than the CCD ones as these additional contractions were computationally less efficient at the time of experimentation due to poor load balancing. Since then, the efficiency of the singles tensor contractions has been improved via a more even work distribution. The lambda equation iterations are slightly faster than the CCSD ones but otherwise behave similarly in terms of scaling. As the size of the AO basis is much larger than that of the MO basis, in particular, the first stages of the integral transformation can make up a large portion of the computational time. This is more important for larger AO sets. In Table 7, one may notice that for the smallest node count, this step even dominates the calculation. Therefore, index transformations require special attention and will be the first target for improvements using techniques like Cholesky decomposition that allow for reduction of the operation count without impacting the accuracy.

Currently, the scalability and efficiency of our GPU-accelerated implementation is hindered by a number of factors. First, in order to amortize the cost of the Host-to-Device memory transfers, we have to maintain a large granularity of the stored tensor blocks. Although it is beneficial for absolute efficiency within a node, it can also cause work starvation at larger node counts when the system is relatively small. Second, for tensor contractions with a large output tensor (e.g., any four-dimensional output tensor), the static load balancing mechanism is activated due to the inefficiency of MPI_Accumulate operation required in the dynamic load balancing mechanism (the latter is now used for tensor contractions with a small output tensor). The static load balancing mechanism is based on the task-binding affinity dictated by the output tensor, which can lead to a noticeable load imbalance when the total number of the output tensor blocks is not divided evenly by the number of the worker

processes. Finally, the third and the most important issue we observe is the serialization overhead caused by the generation of the global task list done by the manager processes. In particular, all runs reported in this work utilized only a single manager process. For the molecular systems reported here, the generation of the global task list could take a noticeable time as compared to the actual task execution time by all worker processes. This explains the low scalability slope observed in Tables 6 and 7, resulting in low parallel efficiency. We are currently addressing this issue by switching to multimanager configurations. Finally, we also observed occasional synchronization overheads for MPI-3 one-sided communications.

To further assess performance on larger node counts, a larger system with a higher operation count is necessary. We therefore also investigated the $(\text{UF}_6)_2$ dimer to have a case with approximately 64 times more floating point operations to process. This system has not yet been treated at the coupled cluster level of theory, but a dimer interaction potential was computed with DFT,⁹⁶ including relativistic effects via a relativistic effective core potential. Concerning the relative position and alignment of the two UF_6 monomers, there are three minima that are depicted in Figure 3. They are

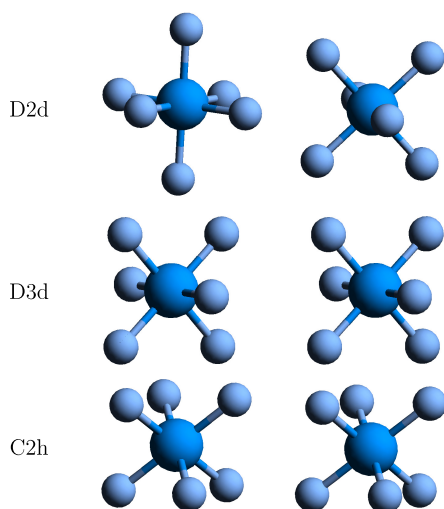


Figure 3. Orientation of the UF_6 dimers.

designated by the symmetry of the complex. The energies and U–U bond distances are listed in Table 8. For the D2d complex, the smallest U–U distance and the highest binding energy were obtained. The C2h complex had the largest separation of the uranium atoms in the equilibrium, and for D3d, the smallest binding energies were obtained. The trend of the energies is the same for the different levels of theory, but the absolute values vary strongly. In the case of Hartree–Fock,

the complexes are barely bound, while MP2 overestimates the binding energy and the CCSD and DFT values are rather close ranging from 0.13 to 0.19 eV.

In the v_3z computation of the dimer, 140 active occupied and 1108 active virtual spinors are taken into account in the coupled cluster computation. With a chosen segment size of 70 for both spaces, the t_2 tensor consists of 1024 blocks. This means that we reach the principal scaling limit at 1024 MPI processes or, equivalently, 512 Summit nodes. If one compares the timings for runs with 385 and 513 nodes in Table 8, one can see a speedup for the integral transformation, but not for the coupled cluster iterations. The time dominating tensor contractions in the integral transformation are those in which indices are transformed from the atomic orbital basis to the virtual space, and these contain a relatively large number of tensor blocks to parallelize over, whereas in the coupled cluster iterations, most tensors are smaller than t_2 . Due to the necessity of avoiding remote accumulates and maintaining large granularity of tensor blocks for GPU processing, the coupled cluster workload simply does not have enough work items to efficiently parallelize over more than 385 nodes (for this specific molecular system). On the other hand, due to high memory demands, we could not use fewer nodes for this particular calculation. Such a situation will be characteristic for molecules with a large virtual-to-occupied spinor ratio, like our current UF_6 dimer system with a significantly reduced occupied space (letting the occupied space include all electrons would restore the scaling to higher node counts). Currently, we are working on improving our original algorithm to address this issue.

Finally, to conclude our performance analysis, Tables 9 and 10 show the scalability of a few representative individual tensor

Table 9. Strong Scaling Benchmark: Times in Seconds of Three Representative Tensor Contractions with 560 AO, 560 Virtual, and 280 Occupied Spinors^a

no. of nodes	contraction 1	contraction 2	contraction 3
32	35.70	11.82	96.68
64	23.45	8.25	56.99
128	22.75	4.32	36.08

^aExecution configuration: 2 MPI processes per node and 3 GPU per MPI process. The number of nodes refers only to Summit nodes running worker processes.

contractions with an increased number of occupied spinors (280) that were performed after some very recent improvements in the ExaTENSOR library. In contrast to previous calculations, these runs were performed with multiple ExaTENSOR manager processes instead of a single manager process, and they dynamically switched between static and

Table 8. UF_6 Dimer Bonding Energies (dE, eV) for a System with 140 Occupied and 1108 Virtual Spinors for Different Computational Methods^a

sym	U–U	dE				<i>n</i>	<i>t</i> ₁	<i>t</i> _{CC/it.}
		DFT	HF	MP2	CCSD			
D3d	5.144	−0.136	0.037	−0.154	−0.131	385	9244	687
D2d	5.139	−0.160	−0.003	−0.189	−0.178	513	8641	758
C2h	5.290	−0.150	−0.001	−0.163	−0.151	1025	8356	665

^aThe U–F distances have been fixed at 1.996 Å;⁶⁶ the U–U distances are optimized with DFT and are listed in the first column in Å. *n* is the number of Summit nodes.

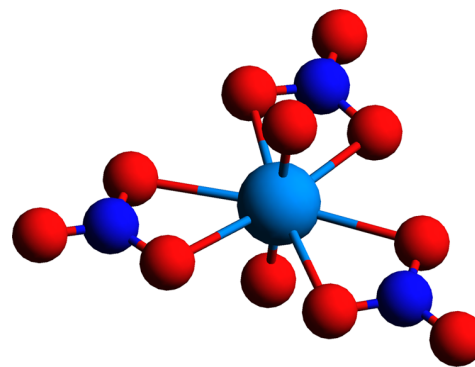
Table 10. Strong Scaling Benchmark: Times in Seconds of Three Representative Tensor Contractions with 840 AO, 840 Virtual, and 280 Occupied Spinors^a

no. of nodes	contraction 1	contraction 2	contraction 3
128	87.66	18.11	177.13
256	71.88	10.32	125.52
512	80.64	7.70	89.11

^aExecution configuration: 2 MPI processes per node and 3 GPU per MPI process. The number of nodes refers only to Summit nodes running worker processes.

dynamic load balancing, as described in Section 3.3. Contraction 1 is a representative contraction of the last stage of the integral transformation: $V(v, v, v, v) + = C(v, a) * H(v, v, v, a)$, where v designates a virtual index placeholder and a designates an atomic index placeholder. Contraction 2 is a representative of a singles projection contraction: $Z(v, o) + = V(v, v, o, v) * T(v, v, o, o)$, where o designates an occupied index placeholder. Contraction 3 is the cost-dominating term in the doubles projection: $Z(v, v, o, o) + = V(v, v, v, v) * T(v, v, o, o)$. As one can see from Table 9, the parallel efficiency of Contraction 1 is not improved as compared to the overall parallel efficiency of the integral transformation reported in Table 7. However, the parallel efficiency of Contraction 2 and Contraction 3, which are two representative contractions from the coupled cluster singles and doubles equations, is much better than that of the original full CCSD run reported in Table 7. Indeed, for Contraction 2, we observe around 68% parallel efficiency from 32 to 128 nodes. For Contraction 3, we similarly observe around 67% parallel efficiency from 32 to 128 nodes, which is encouraging. Unfortunately, due to a shortage in computational time and some technical issues, we have not yet had a chance to rerun the full CCSD iterations with an updated algorithm, which will be done in the near future. For the sake of completeness, we have also run Contraction 3 with an increased number of virtual spinors (1120) on 512 and 1024 worker nodes, observing the execution times of 504.96 and 365.08 s, respectively. Although this confirms the scalability of larger tensor contractions to larger node counts, the corresponding parallel efficiency in the weak scaling regime is only 39%, which indicates communication overhead among other inefficiencies described above. At last, let us report the absolute efficiency of the cost-dominating Contraction 3, which is 44 and 30% on 32 and 128 nodes in Table 9, respectively, and 15% on 512 nodes in Table 10, which includes all reported inefficiencies.

5.4. EFG of Uranium in the Uranyl Tris-nitrate Complex. As a final example of possible applications, we now turn our attention to the calculation of the q_{zz} component of the EFG tensor on the uranium atom for the $[\text{UO}_2(\text{NO}_3)_3]^-$ complex, for which there are experimental values⁹⁷ in the solid state. Initial theoretical investigations of EFGs for actinyl species focused on the bare uranyl ion (UO_2^{2+}),^{19,20} where it was found that a qualitative agreement with the experiment was only achieved if the effect of the equatorial ligands was taken into account (even if through point-charge embedding¹⁹). These studies nevertheless revealed that the $U q_{zz}$ value had a dominant contribution from the so-called $U(6p)$ core-hole, arising from the depletion of charge because of the overlap between the $O(2p)$ and the high-lying antibonding $U(6p_\sigma) + O(2s)$ spinors (Figure 4).

**Figure 4.** Structure of the uranyl complex derived from the X-ray structure for the $\text{RbUO}_2(\text{NO}_3)_3$ crystal.⁶⁷

Explicit inclusion of the contributions from the equatorial ligands to the $U q_{zz}$ value and the associated analysis of orbital contributions to the EFG was, to the best of our knowledge, first performed by Belanzoni and co-workers,⁹⁸ employing the BP86 generalized gradient approximation (GGA) functional, the ZORA-4 Hamiltonian, and QZ4P bases. They have, first, identified that the $U(6p)$ core-hole yielded a positive contribution to the EFG, though in bare uranyl these were offset by negative contributions due to the nonspherical electron distribution in the valence $5f$ shell caused by the $U-O$ bonding. Moreover, positive contributions due to the ligands arise from the tails of the $U(6p)$ spinor, which extends significantly to the region of the nitrate ligands, as well as from electron donation by the nitrate groups into $U(5f_{\text{orb}}, 6d_s)$, which have lobes in the equatorial plane. These calculations were found to underestimate the $U q_{zz}$ experimental value by around 4 au.

More recently, Autschbach and co-workers⁹⁹ have employed the X2C-1e Hamiltonian, triple zeta quality basis sets (U , ANO(-h); light atoms, TZVPP), and density functional approximations (DFAs) including Hartree-Fock exchange, such as B3LYP and CAMB3LYP, to revisit the $U q_{zz}$ on the $[\text{UO}_2(\text{HCO}_3)_3]^-$ complex that resembles fairly well the structural motif in $[\text{UO}_2(\text{NO}_3)_3]^-$, though for which, unfortunately, we are not aware of any experimental values. These results have demonstrated the importance of accounting for picture change effects in the representation of the EFG operator (increases in $U q_{zz}$ of around 8 au, fairly consistent among the different methods), as well as the importance of including Hartree-Fock exchange in the DFAs from going from GGAs⁹⁸ to hybrids⁹⁹ for obtaining larger values for $U q_{zz}$. They have also confirmed the large effect of the ligands on $U q_{zz}$ found by Belanzoni and co-workers, as $U q_{zz}$ goes from a negative value (between -8 and -7 au depending on the DFA) in bare uranyl to a positive one (see Table 12).

Though the results by Autschbach and co-workers suggest that DFAs would work rather well in this case, it is well known in the literature¹⁰⁰ that it can be difficult for these to correctly represent EFGs for transition metals^{101,102} or lanthanides,⁷⁹ and as such it would be highly desirable to perform EFG calculations at the coupled cluster level more routinely, if not to provide benchmark values for systems larger than diatomics or triatomics. In this respect, our calculations are, to the best of our knowledge, the first effort to obtain EFGs at the CCSD level for uranyl species while explicitly including the equatorial ligands—in the pioneering calculation by de Jong and co-workers,¹⁹ the structure of uranyl was investigated with CC,

Table 11. Comparison of Times to Solution (TTS, in Hours, Equal to the Total Wall Time for Each Calculation and the Part Spent in the T and Λ Equations), T_1 Diagnostic, and q_{zz} Component of the EFG Tensor (in Atomic Units) for CCSD Expectation Value Calculations for the Uranyl Tris-nitrate Complex ($[\text{UO}_2(\text{NO}_3)_3]^-$) with the X2C-AMFI Hamiltonian^a

calculation	O	x_o	V	x_v	M	C	NormC	TTS				
								total	T + Λ	T_1	q_{zz}	
						dyall.v2z						
1	106	0.53	534	0.42	160	9	1.00	3h34	2h55	0.0126	10.02	
2	156	0.77	534	0.42	160	20	2.17	5h33	5h06	0.0103	9.70	
3	156	0.77	818	0.64	360	109	5.30	11h21	10h43	0.0102	8.48	
4	202	1.00	534	0.42	400	33	1.45	8h02	7h29	0.0091	9.73	
5	202	1.00	680	0.53	2050	87	0.75	12h28	11h58	0.0090	8.89	
6	202	1.00	818	0.64	2050	183	1.56	18h23	17h49	0.0089	8.59	
7	202	1.00	896	0.68	2050	263	2.25	23h04	22h21	0.0089	8.54	
						dyall.v3z						
8	106	0.53	694	0.33	480	26	0.95	6h	4h08	0.0139	10.29	
9	156	0.72	694	0.33	480	57	2.06	10h24	8h02	0.0114	10.04	
10	156	0.72	944	0.46	2050	193	1.65	17h48	15h44	0.0104	9.99	

^aO, number of occupied spinors correlated; V, number of virtual spinors correlated; M, number of ranks in the parallel calculation; C, cost estimate for the calculation (scaled by 1.0×10^{14}); NormC, cost divided by the number of ranks and normalized with respect to the value for the smallest calculation. See the text for details.

but the EFG was only calculated at the Hartree–Fock level. As such, our calculations are also the first to investigate the effects of increasing the number of correlated occupied and virtual spinors on the EFG values of uranyl complexes.

Our results are shown in Tables 11 and 12. We note that we have restricted ourselves to the X2C-AMFI Hamiltonian to remain close to the setup of prior calculations.⁹⁹ Furthermore, due to constraints on the available memory and other resources, we were only able to carry out calculations with significantly extended virtual spaces for double zeta calculations. From Table 11, the first interesting result is the behavior of the T_1 diagnostic. For the smallest calculation (1), we have a value slightly greater than 0.012, which is higher than what is usually observed for molecules containing light elements but in line with our experience^{7,61,68,103} of CCSD calculations on heavy elements in which we did not correlate electrons below the U 5d shell. When we increase the number of correlated electrons, the T_1 value progressively decreases and reaches around 0.009 for the calculations in which all electrons are correlated (4–7). This would suggest, therefore, that in prior calculations the relatively high T_1 values (with respect to the prescriptions originally put forth for light-element molecules) arise mostly due to an incomplete occupied correlating space, instead of being the sign of a growing multireference character for the heavy elements.

Concerning the value of q_{zz} , we have fairly large variations going from the calculations in which the correlation space is rather small (1) to the largest calculation performed (7). Comparing calculations in which the number of occupied spinors is increased but the number of virtual spinors is kept constant (1 and 2 and 1 and 4) to those in which the reverse is true (2 and 3), we have that the size of the virtual space is the most significant variable to control. From our calculations, we observe that it is only after including roughly 60% of the virtual space in the correlating space (more than 680 virtual spinors for double zeta bases) that the q_{zz} values start to converge to a value around 8.5 au, and though we have a much more limited set of data for triple zeta bases, we seem to observe a similar pattern, with rather large variations in the q_{zz} values with a change in the number of correlated virtual spinors. Unfortunately, with the largest calculations we carried out,

we are not able to include more than roughly 45% of the virtual space in the correlating space, which appears not to be sufficient for obtaining a converged q_{zz} value.

Strictly speaking, the calculations presented in Table 11 cannot be used to characterize the strong scaling of the code. This is because the memory required to store the two-electron integral tensor grows sharply as we increase the number of correlated occupied and virtuals, requiring that the number of Summit nodes used was changed for each run to fit this tensor in. It is nevertheless possible to extract some information on the code's weak scaling. To do so, we define the following two metrics: (a) the cost of calculation n , taken to be proportional to the number of operations of the costlier contractions for both amplitude and Λ equations ($C[n] = O^2V^4[n]$, with O being the number of correlated occupied spinors and V being the number of correlated virtual spinors) and (b) cost per rank M employed in the parallel calculation, normalized by the cost of calculation 1, the smallest considered here ($\text{NormC}[n] = C[n]/(M^*C[1])$). We see that the significant cost increase resulting from the augmentation of the occupied and virtual spaces (a 28-fold increase from calculation 1 to calculation 7) can be offset by an increase in the number of ranks in the parallel calculation (a 13-fold increase comparing the same calculations) such that the time to the solution remains within reasonable bounds (less than 24 h). At the same time, NormC for the different calculations remains, with a few exceptions, between 1 and 2, and that including both the double and triple zeta calculations. This indicates that the code can handle high workloads in roughly the same manner as it does for much smaller workloads.

Finally, in Table 12, we compare our best CCSD results to uncorrelated (Hartree–Fock), DFT calculations performed with DIRAC for hybrid DFAs, to the results from Belanzoni and co-workers⁹⁸ and Autschbach and co-workers,⁹⁹ and to the experiment.⁹⁷ We observe that the Hartree–Fock calculations provide largely overestimated values with respect to the experiment, with differences of above 6 au for the double, triple, and quadruple zeta bases, whereas all DFT calculations provide somewhat underestimated values—among the functionals compared, B3LYP fares the worst (differences from -1.74 to -1.49 au), while PBE0 (differences from -0.80 and

Table 12. Comparison between Electronic Structure Methods and Experimental Results for the q_{zz} Component of the EFG Tensor for the Uranyl Tris-nitrate Complex ($[\text{UO}_2(\text{NO}_3)_3]^-$) for Basis Sets of Different Cardinal Numbers^b

method	2z	3z	4z
HF	14.67	14.85	14.97
B3LYP	6.64	6.79	6.89
CAMB3LYP	7.24	7.40	7.50
PBE0	7.58	7.75	7.85
CCSD	8.54		
ref 98, BP86, ZORA-4			4.11
ref 99 ^a , HF, X2C-1e		15.17	
ref 99 ^a , B3LYP, X2C-1e		6.81	
ref 99 ^a , CAMB3LYP, X2C-1e		8.33	
exp ⁹⁷		8.38 ± 0.13	

^aThe coupled cluster calculation employs 202 occupied and 896 virtual spinors. All values in are atomic units. The X2C-AMFI Hamiltonian is employed in all calculations. ^bCalculations on $[\text{UO}_2(\text{HCO}_3)_3]^-$.

−0.53 au) fares slightly better than CAMB3LYP (differences from −1.14 and −0.88 au). Taken together, these results indicate that for the mean-field approaches, an increase in basis set quality translates into an increase in q_{zz} of slightly under 0.2 au going from double to triple zeta, with an additional increase of 0.1 au when going from triple to quadruple zeta and that all DFT results appear to move in the direction of the experiment. We observe that our Hartree–Fock and DFT values are slightly smaller than those obtained by Autschbach and co-workers,⁹⁹ though a somewhat larger discrepancy is seen for CAMB3LYP than that for B3LYP or Hartree–Fock, which suggests something other than purely the difference in the molecular system could be at play here and that the very good agreement to experiment with CAMB3LYP in the literature seems fortuitous.

Our best double zeta CCSD result (calculation 7) shows only a slight overestimation (0.16 au) with respect to the experiment, which is a significantly better result than any of the mean-field ones. That said, since we were not able to perform accurate calculations with more flexible basis sets and, from the DFT trends, it would not be surprising that these results, compared to future calculations, would be found to underestimate the q_{zz} value at least a few tenths of an atomic unit, pointing to an overall overestimation of the experimental results at the CCSD level. Once we implement approaches that allow us to efficiently truncate our (artificially, due to the use of uncontracted basis sets) large virtual spaces without the loss of accuracy, we intend to revisit this question.

6. CONCLUSIONS

A reimplement of the Kramers-unrestricted coupled cluster method was presented and shown to be able to exploit hundreds of GPU-accelerated nodes while correctly reproducing the results of the prior implementation for a range of test cases. With this implementation, which currently does not exploit any rank reduction techniques or other approximations, we were already able to investigate systems for which about 200 electrons and around 1000 virtual molecular spinors were taken into account in the correlated calculations.

Current functionalities include CC2, CCD, and CCSD wave functions and perturbative triples corrections to the ground-

state. Ground-state expectation values are also available for CCD and CCSD through the computation of one-body density matrices. The code is now interfaced to the DIRAC and ReSpect packages, but interfaces to other software packages employing Gaussian-type atomic orbitals can be implemented in a straightforward manner as they only require functionality to read in the molecular spinors.

Unlike the original implementation, this reimplement does not exploit double point group symmetry. For the aforementioned functionality, this is a disadvantage for small, highly symmetric systems (i.e., around 10 atoms or 100 electrons), though we consider that in practice this shortcoming is offset by large-scale parallelization that allows us to treat such systems upon distortions (bent/twisted configurations) with triple or quadruple zeta basis sets without hitting the wall of transitioning from real to complex algebra (as is the case in the original code when symmetry is reduced), inefficient communication over hundreds of nodes, and the extensive use of disk storage of intermediate quantities. For larger systems, which often have little to no symmetry, the current implementation makes correlated calculations feasible.

We have employed the code to investigate the properties of different heavy-element complexes: ionization energies for lanthanide monofluorides (LaF, EuF, and YbF), energies of formation of gold–argon clusters of different sizes, and the electric field gradient (q_{zz}^U) at the uranium nucleus in a uranyl tris-nitrate complex.

For the lanthanide monofluorides, we have shown that we need to correlate at least the Ln (4s4p4d4f) electron shell to obtain converged ionization energies. We observed that for LaF and EuF, the perturbative triples-corrected ionization energies for the triple zeta quality basis sets are already within the experimental error bounds, with adiabatic values showing a slightly better agreement to the experiment than vertical ones. For YbF, we have encountered the same issue previously described in the literature, in that spin–orbit-coupled calculations show a surprisingly large T_1 diagnostic value around the equilibrium distance, which makes perturbative triples results unreliable, but we see a smooth convergence of the CCSD values toward the experimental values.

For the gold–argon clusters, substantial binding energy is obtained at the coupled cluster level for closed-shell noble metal Au^+ and noble gas Ar, with negligible values at the Hartree–Fock level (below 0.6 eV) and overestimation of the binding energy up to 0.5 eV for MP2. Therefore, a selection of a reliable method is important. For AuAr_4^+ , two different structures that are close in energy have been computed and the 3D structure is about 0.2/0.3 eV lower than the 2D structure at the CCSD/CCSD(T) level.

For the uranyl tris-nitrate complex, we have found that CCSD wave functions are capable of providing q_{zz}^U values in very good agreement to the experiment, with non-negligible improvements over DFT calculations. This improvement, however, appears to come at the cost of including high-lying virtual spinors in the CCSD calculation. We have been limited to performing CCSD calculations with double zeta basis due to the number of virtual spinors that appears to be required to converge the q_{zz}^U value. While the DFT results suggest non-negligible basis set effects due to the use of uncontracted basis sets, for reasons of computational cost, it was not possible to include enough high-lying virtual spinors in triple zeta calculations and therefore the calculations we were able to perform are still far from converged. We are currently working

on implementing approaches to compress the virtual space while retaining precision in the correlated treatment to enable efficient calculations with larger basis sets. We expect that these will allow us not only to revisit this system in the near future but also enable calculations on significantly larger systems.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.1c00260>.

Relevant equations for the implementation, more detailed comparison between the current implementation and RELCCSD results, tables determining the active space of LaF and EuF, CCSD bond distances for AuAr⁺, DFT results for (UF₆)₂, and details on setting up the input and code examples for the usage of TAL-SH and ExaTENSOR (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Johann V. Pototschnig – Department of Chemistry and Pharmaceutical Sciences, Faculty of Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands;

orcid.org/0000-0002-9982-0556;

Email: j.v.pototschnig@vu.nl

Anastasios Papadopoulos – Department of Chemistry and Pharmaceutical Sciences, Faculty of Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands;

Email: papadopoulos@kofo.mpg.de

André Severo Pereira Gomes – Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers, Atomes et Molécules, F-59000 Lille, France; orcid.org/0000-0002-5437-2251; Email: andre.gomes@univ-lille.fr

Hans Jørgen Aa Jensen – Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense M, Denmark; Email: hjj@sdu.dk

Lucas Visscher – Department of Chemistry and Pharmaceutical Sciences, Faculty of Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands;

orcid.org/0000-0002-7748-6243; Email: l.visscher@vu.nl

Authors

Dmitry I. Lyakh – National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States

Michal Repisky – Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, UiT The Arctic University of Norway, N-9037 Tromsø, Norway;

orcid.org/0000-0003-0776-4137

Loïc Halbert – Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers, Atomes et Molécules, F-59000 Lille, France

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jctc.1c00260>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User

Facility supported under Contract DE-AC05-00OR22725. Some computer codes used in this research (ExaTENSOR, TAL-SH, partly ExaCorr) were developed during the OLCF-4 Center for Accelerated Application Readiness (CAAR) program funded by the US Department of Energy at the Oak Ridge National Laboratory. ASPG and LH acknowledge support from PIA ANR project CaPPA (ANR-11-LABX-0005-01), the Franco-German project ComprIXS (Agence nationale de la recherche ANR-19-CE29-0019, Deutsche Forschungsgemeinschaft JA 2329/6-1), the I-SITE ULNE project OVERSEE, the French Ministry of Higher Education and Research, region Hauts de France council, the European Regional Development Fund (ERDF) project CPER CLIMBIO, and the French national supercomputing facilities (grants DARI A0070801859 and Joliot Curie grands challenges 2019 gch0417). ASPG, LH, JVP, and LV acknowledge support from MESONM International Associated Laboratory (LAI) (ANR-16-IDEX-0004). J.V.P. acknowledges funding from the Austrian Science Fund (FWF, J 4177-N36). M.R. acknowledges the funding support from the Research Council of Norway through a Center of Excellence Grant (Grant no. 262695).

■ REFERENCES

- (1) Thiel, W.; Hummer, G. Methods for computational chemistry. *Nature* **2013**, *504*, 96–97.
- (2) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. The accurate determination of molecular equilibrium structures. *J. Chem. Phys.* **2001**, *114*, 6548–6556.
- (3) Coriani, S.; Marchesan, D.; Gauss, J.; Hättig, C.; Helgaker, T.; Jørgensen, P. The accuracy of ab initio molecular geometries for systems containing second-row atoms. *J. Chem. Phys.* **2005**, *123*, No. 184107.
- (4) Bartlett, R. J. Coupled-cluster theory and its equation-of-motion extensions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 126–138.
- (5) Shee, A.; Saue, T.; Visscher, L.; Severo Pereira Gomes, A. Equation-of-motion coupled-cluster theory based on the 4-component Dirac-Coulomb (-Gaunt) Hamiltonian. Energies for single electron detachment, attachment, and electronically excited states. *J. Chem. Phys.* **2018**, *149*, No. 174113.
- (6) Haase, P. A. B.; Eliav, E.; Iliáš, M.; Borschevsky, A. Hyperfine Structure Constants on the Relativistic Coupled Cluster Level with Associated Uncertainties. *J. Phys. Chem. A* **2020**, *124*, 3157–3169.
- (7) Kervazo, S.; Réal, F.; Virost, F.; Severo Pereira Gomes, A.; Vallet, V. Accurate Predictions of Volatile Plutonium Thermodynamic Properties. *Inorg. Chem.* **2019**, *58*, 14507–14521.
- (8) Denis, M.; Haase, P. A. B.; Timmermans, R. G. E.; Eliav, E.; Hutzler, N. R.; Borschevsky, A. Enhancement factor for the electric dipole moment of the electron in the BaOH and YbOH molecules. *Phys. Rev. A* **2019**, *99*, No. 042512.
- (9) Zelovich, T.; Borschevsky, A.; Eliav, E.; Kaldor, U. Relativistic coupled cluster calculation of Mössbauer isomer shifts of iodine compounds. *Mol. Phys.* **2017**, *115*, 138–143.
- (10) van Stralen, J. N. P.; Visscher, L. The nuclear quadrupole moment of ¹¹⁵In from molecular data. *J. Chem. Phys.* **2002**, *117*, 3103–3108.
- (11) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496–4501.
- (12) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
- (13) Förster, A.; Visscher, L. Low-Order Scaling G0W0 by Pair Atomic Density Fitting. *J. Chem. Theory Comput.* **2020**, *16*, 7381–7399.
- (14) Beebe, N. H. F.; Linderberg, J. Simplifications in the generation and transformation of two-electron integrals in molecular calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683–705.

- (15) Pedersen, T. B.; Aquilante, F.; Lindh, R. Density fitting with auxiliary basis sets from Cholesky decompositions. *Theor. Chem. Acc.* **2009**, *124*, 1–10.
- (16) Peng, B.; Kowalski, K. Highly Efficient and Scalable Compound Decomposition of Two-Electron Integral Tensor and Its Application in Coupled Cluster Calculations. *J. Chem. Theory Comput.* **2017**, *13*, 4179–4192.
- (17) Helmich-Paris, B.; Repisky, M.; Visscher, L. Relativistic Cholesky-decomposed density matrix MP2. *Chem. Phys.* **2019**, *518*, 38–46.
- (18) Sikkema, J.; Visscher, L.; Saue, T.; Iliáš, M. The molecular mean-field approach for correlated relativistic calculations. *J. Chem. Phys.* **2009**, *131*, No. 124116.
- (19) de Jong, W. A.; Visscher, L.; Nieuwpoort, W. C. On the bonding and the electric field gradient of the uranyl ion. *J. Mol. Struct.: THEOCHEM* **1998**, *458*, 41–52.
- (20) Infante, I.; Visscher, L. QM/MM study of aqueous solvation of the uranyl fluoride [UO₂F] complex. *J. Comput. Chem.* **2004**, *25*, 386–392.
- (21) Calvin, J. A.; Peng, C.; Rishi, V.; Kumar, A.; Valeev, E. F. Many-Body Quantum Chemistry on Massively Parallel Computers. *Chem. Rev.* **2021**, *121*, 1203–1231.
- (22) Visscher, L.; Lee, T. J.; Dyall, K. G. Formulation and implementation of a relativistic unrestricted coupled-cluster method including noniterative connected triples. *J. Chem. Phys.* **1996**, *105*, 8769–8776.
- (23) Pernpointner, M.; Visscher, L. Parallelization of four-component calculations. II. Symmetry-driven parallelization of the 4-Spinor CCSD algorithm. *J. Comput. Chem.* **2003**, *24*, 754–759.
- (24) Lyakh, D. I. Domain-specific virtual processors as a portable programming and execution model for parallel computational workloads on modern heterogeneous high-performance computing architectures. *Int. J. Quantum Chem.* **2019**, *119*, No. e25926.
- (25) Papadopoulos, S. Working towards a Massively Parallel Equation-of-Motion Coupled-Cluster Implementation in DIRAC. M.Sc. Thesis, Vrije Universiteit: Amsterdam, 2019.
- (26) Parrish, R. M.; Hohenstein, E. G.; Schunck, N. F.; Sherrill, C. D.; Martinez, T. J. Exact Tensor Hypercontraction: A Universal Technique for the Resolution of Matrix Elements of Local Finite-Range N-Body Potentials in Many-Body Quantum Problems. *Phys. Rev. Lett.* **2013**, *111*, No. 132505.
- (27) Peng, B.; Kowalski, K. Low-rank factorization of electron integral tensors and its application in electronic structure theory. *Chem. Phys. Lett.* **2017**, *672*, 47–53.
- (28) Almlöf, J. Elimination of energy denominators in Møller–Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319–320.
- (29) Visscher, L.; Dyall, K. G. Dirac-Fock atomic electronic structure calculations using different nuclear charge distributions. *At. Data Nucl. Data Tables* **1997**, *67*, 207–224.
- (30) Iliáš, M.; Saue, T. An infinite-order two-component relativistic Hamiltonian by a simple one-step transformation. *J. Chem. Phys.* **2007**, *126*, No. 064102.
- (31) Konecny, L.; Kadek, M.; Komorovsky, S.; Malkina, O. L.; Ruud, K.; Repisky, M. Acceleration of Relativistic Electron Dynamics by Means of X2C Transformation: Application to the Calculation of Nonlinear Optical Properties. *J. Chem. Theory Comput.* **2016**, *12*, 5823–5833.
- (32) Schimmelpfennig, B. *AMFI, An Atomic Mean-field Spin-Orbit Integral Program*; University of Stockholm: Stockholm, Sweden, 1999.
- (33) Kramers, H. A. Théorie générale de la rotation paramagnétique dans les cristaux. *Proc. R. Neth. Acad. Arts Sci.* **1930**, *33*, 959–972.
- (34) Saue, T.; A Jensen, H. J. Quaternion symmetry in relativistic molecular calculations: The Dirac-Hartree-Fock method. *J. Chem. Phys.* **1999**, *111*, 6211–6222.
- (35) Repisky, M.; Komorovsky, S.; Kadek, M.; Konecny, L.; Ekström, U.; Malkin, E.; Kaupp, M.; Ruud, K.; Malkina, O. L.; Malkin, V. G. ReSpecT: Relativistic spectroscopy DFT program package. *J. Chem. Phys.* **2020**, *152*, No. 184101.
- (36) Visscher, L.; Dyall, K. G.; Lee, T. J. Kramers-restricted closed-shell CCSD theory. *Int. J. Quantum Chem.* **1995**, *56*, 411–419.
- (37) Pulay, P. Convergence acceleration of iterative sequences. The case of scf iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398.
- (38) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (39) Deegan, M. J. O.; Knowles, P. J. Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories. *Chem. Phys. Lett.* **1994**, *227*, 321–326.
- (40) Arponen, J. Variational principles and linked-cluster exp S expansions for static and dynamic many-body problems. *Ann. Phys.* **1983**, *151*, 311–382.
- (41) Shee, A.; Visscher, L.; Saue, T. Analytic one-electron properties at the 4-component relativistic coupled cluster level with inclusion of spin-orbit coupling. *J. Chem. Phys.* **2016**, *145*, No. 184107.
- (42) Wigner, E. P. *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*; Academic Press, 1959.
- (43) Saue, T. Principles and Applications of Relativistic Molecular Calculations. Ph.D. Thesis, University of Oslo, 1996. May be retrieved from <http://diracprogram.org>.
- (44) DALTON, A Molecular Electronic Structure Program, see <http://daltonprogram.org>.
- (45) Shiozaki, T. An efficient solver for large structured eigenvalue problems in relativistic quantum chemistry. *Mol. Phys.* **2017**, *115*, 5–12.
- (46) Pototschnig, J. V.; Papadopoulos, A.; Lyakh, D. I.; Repisky, M.; Halbert, L.; Gomes, A. S. P.; Jensen, H. J. A.; Visscher, L. Dataset: Implementation of Relativistic Coupled Cluster Theory for Massively Parallel GPU-Accelerated Computing Architectures, 2021, <https://doi.org/10.5281/zenodo.4589358>.
- (47) Repisky, M. *InteRest, An Integral Library for Relativistic Quantum Chemistry*; Wiley Online Library, 2018.
- (48) Stanton, R. E.; Havriliak, S. Kinetic balance: A partial solution to the problem of variational safety in Dirac calculations. *J. Chem. Phys.* **1984**, *81*, 1910–1918.
- (49) Obara, S.; Saika, A. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.* **1986**, *84*, 3963–3974.
- (50) Kadek, M.; Repisky, M.; Ruud, K. All-electron fully relativistic Kohn-Sham theory for solids based on the Dirac-Coulomb Hamiltonian and Gaussian-type functions. *Phys. Rev. B* **2019**, *99*, No. 205103.
- (51) Komorovský, S.; Repiský, M.; Malkina, O. L.; Malkin, V. G.; Malkin Ondřík, I.; Kaupp, M. A fully relativistic method for calculation of nuclear magnetic shielding tensors with a restricted magnetically balanced basis in the framework of the matrix Dirac-Kohn-Sham equation. *J. Chem. Phys.* **2008**, *128*, No. 104101.
- (52) Komorovský, S.; Repiský, M.; Malkina, O. L.; Malkin, V. G. Fully relativistic calculations of NMR shielding tensors using restricted magnetically balanced basis and gauge including atomic orbitals. *J. Chem. Phys.* **2010**, *132*, No. 154101.
- (53) Lyakh, D. I. TAL-SH: Numerical Tensor Algebra Library for Shared-Memory Heterogeneous Nodes Equipped with Multicore CPU and NVIDIA GPU. https://github.com/DmitryLyakh/TAL_SH.
- (54) Thyssen, J. Development and Applications of Methods for Correlated Relativistic Calculations of Molecular Properties. Ph.D. Thesis, University of Southern Denmark, 2001.
- (55) Yoshimine, M. Construction of the hamiltonian matrix in large configuration interaction calculations. *J. Comput. Phys.* **1973**, *11*, 449–454.
- (56) Ziółowski, M.; Weijo, V.; Jørgensen, P.; Olsen, J. An efficient algorithm for solving nonlinear equations with a minimal number of trial vectors: Applications to atomic-orbital based coupled-cluster theory. *J. Chem. Phys.* **2008**, *128*, No. 204105.
- (57) Gomes, A. S. P.; Saue, T.; Visscher, L.; Jensen, H. J. A.; Bast, R. DIRAC, A Relativistic Ab Initio Electronic Structure Program, Release DIRAC19,2019, written by, with contributions from I. A.

- Aucar, V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, E. Faßhauer, T. Fleig, O. Fossgaard, L. Halbert, E. D. Hedegård, B. Heimlich-Paris, T. Helgaker, J. Henriksson, M. Iliáš, Ch. R. Jacob, S. Knecht, S. Komorovský, O. Kullie, J. K. Lærdahl, C. V. Larsen, Y. S. Lee, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, J. M. H. Olsen, Y. C. Park, J. K. Pedersen, M. Pernpointner, R. di Remigio, K. Ruud, P. Salek, B. Schimmelpfennig, B. Senjean, A. Shee, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, M. L. Vidal, S. Villaume, O. Visser, T. Winther, and S. Yamamoto. Available at: <http://dx.doi.org/10.5281/zenodo.3572669>.
- (58) Saue, T.; Bast, R.; Gomes, A. S. P.; Jensen, H. J. A.; Visscher, L.; Aucar, I. A.; Di Remigio, R.; Dyall, K. G.; Eliav, E.; Fasshauer, E.; Fleig, T.; Halbert, L.; Hedegård, E. D.; Helmich-Paris, B.; Iliáš, M.; Jacob, C. R.; Knecht, S.; Laerdahl, J. K.; Vidal, M. L.; Nayak, M. K.; Olejniczak, M.; Olsen, J. M. H.; Pernpointner, M.; Senjean, B.; Shee, A.; Sunaga, A.; van Stralen, J. N. P. The DIRAC code for relativistic molecular calculations. *J. Chem. Phys.* **2020**, *152*, No. 204104.
- (59) Dyall, K. G. Relativistic Quadruple-Zeta and Revised Triple-Zeta and Double-Zeta Basis Sets for the 4p, 5p, and 6p Elements. *Theor. Chem. Acc.* **2006**, *115*, 441–447.
- (60) Dyall, K. G. Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the actinides Ac-Lr. *Theor. Chem. Acc.* **2007**, *117*, 491–500.
- (61) Gomes, A. S. P.; Dyall, K. G.; Visscher, L. Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the lanthanides La-Lu. *Theor. Chem. Acc.* **2010**, *127*, 369–381.
- (62) ADF2019, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, <https://www.scm.com>.
- (63) van Lenthe, E.; Ehlers, A.; Baerends, E.-J. Geometry optimizations in the zero order regular approximation for relativistic effects. *J. Chem. Phys.* **1999**, *110*, 8943–8953.
- (64) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (65) Pototschnig, J. V.; Visscher, L.; Jamshidi, Z. Bonding of argon to gold cluster, 2021, in preparation.
- (66) Kimura, M.; Schomaker, V.; Smith, D. W.; Weinstock, B. Electron-Diffraction Investigation of the Hexafluorides of Tungsten, Osmium, Iridium, Uranium, Neptunium, and Plutonium. *J. Chem. Phys.* **1968**, *48*, 4001–4012.
- (67) Barclay, G. A.; Sabine, T. M.; Taylor, J. C. The crystal structure of rubidium uranyl nitrate: A neutron-diffraction study. *Acta Crystallogr.* **1965**, *19*, 205–209.
- (68) Gomes, A. S. P.; Jacob, C. R.; Réal, F.; Visscher, L.; Vallet, V. Towards systematically improvable models for actinides in condensed phase: the electronic spectrum of uranyl in Cs₂UO₂Cl₄ as a test case. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15153–15162.
- (69) Aebersold, L. E.; Yuwono, S. H.; Schoendorff, G.; Wilson, A. K. Efficacy of Density Functionals and Relativistic Effective Core Potentials for Lanthanide-Containing Species: The Ln54 Molecule Set. *J. Chem. Theory Comput.* **2017**, *13*, 2831–2839.
- (70) Sekiya, M.; Noro, T.; Koga, T.; Shimazaki, T. Relativistic segmented contraction basis sets with core-valence correlation effects for atoms La-57 through Lu-71: Sapporo-DK-nZP sets (n = D, T, Q). *Theor. Chem. Acc.* **2012**, *131*, No. 1247.
- (71) Solomonik, V. G.; Smirnov, A. N. Toward Chemical Accuracy in ab Initio Thermochemistry and Spectroscopy of Lanthanide Compounds: Assessing Core-Valence Correlation, Second-Order Spin-Orbit Coupling, and Higher Order Effects in Lanthanide Diatomics. *J. Chem. Theory Comput.* **2017**, *13*, 5240–5254.
- (72) Bernard, A.; Effantin, C.; d'Incan, J.; Verges, J. The 1¹I₁, 2¹Σ⁺ – > X¹Σ⁺ transitions of LaF. *J. Mol. Spectrosc.* **2000**, *202*, 163–165.
- (73) Dmitriev, Y. N.; Kaledin, L. A.; Kobylansky, A. I.; Kulikov, A. N.; Shenyavskaya, E. A.; Gurvich, L. V. Electronic spectra of diatomic molecules containing f-elements: GdO, EuF and UO. *Acta Phys. Hung.* **1987**, *61*, 51–54.
- (74) Dickinson, C. S.; Coxon, J. A.; Walker, N. R.; Gerry, M. C. L. Fourier transform microwave spectroscopy of the 2²Σ⁺ ground states of YbX (X=F, Cl, Br): Characterization of hyperfine effects and determination of the molecular geometries. *J. Chem. Phys.* **2001**, *115*, 6979–6989.
- (75) Kaledin, L. A.; Heaven, M. C.; Field, R. W. Thermochemical properties (D₀⁰ and IP) of the lanthanide monohalides. *J. Mol. Spectrosc.* **1999**, *193*, 285–292.
- (76) Zmbov, K. F.; Margrave, J. L. Mass Spectrometric Studies of Scandium Yttrium Lanthanum and Rare-earth Fluorides. *Adv. Chem. Ser.* **1968**, *72*, 267.
- (77) Zmbov, K. F.; Margrave, J. L. Mass Spectrometric Studies at High Temperatures. 13. Stabilities of Samarium Europium and Gadolinium Mono- and Difluorides. *J. Inorg. Nucl. Chem.* **1967**, *29*, No. 59.
- (78) Dzuba, V. A.; Derevianko, A. Dynamic polarizabilities and related properties of clock states of the ytterbium atom. *J. Phys. B: At, Mol. Opt. Phys.* **2010**, *43*, No. 074011.
- (79) Pašteka, L. F.; Mawhorter, R. J.; Schwerdtfeger, P. Relativistic coupled-cluster calculations of the ¹⁷³Yb nuclear quadrupole coupling constant for the YbF molecule. *Mol. Phys.* **2016**, *114*, 1110–1117.
- (80) Lyakh, D. I.; Musial, M.; Lotrich, V. F.; Bartlett, R. J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.* **2012**, *112*, 182–243.
- (81) Kapur, S.; Müller, E. W. Metal-neon compound ions in low field evaporation. *Surf. Sci.* **1977**, *62*, 610–620.
- (82) Pyykkö, P. Predicted Chemical Bonds between Rare Gases and Au+. *J. Am. Chem. Soc.* **1995**, *117*, 2067–2070.
- (83) Shayeghi, A.; Johnston, R. L.; Rayner, D. M.; Schäfer, R.; Fielicke, A. The Nature of Bonding between Argon and Mixed Gold-Silver Trimers. *Angew. Chem., Int. Ed.* **2015**, *54*, 10675–10680.
- (84) Ghiringhelli, L. M.; Levchenko, S. V. Strengthening gold-gold bonds by complexing gold clusters with noble gases. *Inorg. Chem. Commun.* **2015**, *55*, 153–156.
- (85) Ferrari, P.; Hou, G.-L.; Lushchikova, O. V.; Calvo, F.; Bakker, J. M.; Janssens, E. The structures of cationic gold clusters probed by far-infrared spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11572–11577.
- (86) Pan, S.; Jana, G.; Merino, G.; Chattaraj, P. K. Noble-Noble Strong Union: Gold at Its Best to Make a Bond with a Noble Gas Atom. *ChemistryOpen* **2019**, *8*, 173–187.
- (87) Jamshidi, Z.; Far, M. F.; Maghari, A. Binding of Noble Metal Clusters with Rare Gas Atoms: Theoretical Investigation. *J. Phys. Chem. A* **2012**, *116*, 12510–12517.
- (88) Jamshidi, Z.; Lushchikova, O. V.; Bakker, J. M.; Visscher, L. Not Completely Innocent: How Argon Binding Perturbs Cationic Copper Clusters. *J. Phys. Chem. A* **2020**, *124*, 9004–9010.
- (89) Beu, T. A.; Onoe, J.; Takeuchi, K. Calculations of structure and IR-spectrum for small UF₆ clusters. *J. Chem. Phys.* **1997**, *106*, 5910–5919.
- (90) Beu, T. A.; Onoe, J.; Takeuchi, K. Structure and frequency shift calculations for small UF₆ clusters. *J. Mol. Struct.* **1997**, *410–411*, 295–298.
- (91) Malli, G. L.; Styszynski, J. Ab initio all-electron Dirac-Fock-Breit calculations for UF₆. *J. Chem. Phys.* **1996**, *104*, 1012–1017.
- (92) Kovács, A.; Konings, R. J. M. Theoretical study of UX₆ and UO₂X₂ (X=F, Cl, Br, I). *J. Mol. Struct.: THEOCHEM* **2004**, *684*, 35–42.
- (93) Batista, E. R.; Martin, R. L.; Hay, P. J.; Peralta, J. E.; Scuseria, G. E. Density functional investigations of the properties and thermochemistry of UF₆ and UF₅ using valence-electron and all-electron approaches. *J. Chem. Phys.* **2004**, *121*, 2144–2150.
- (94) Manzhos, S.; Carrington, T.; Laverdure, L.; Mosey, N. Computing the Anharmonic Vibrational Spectrum of UF₆ in 15 Dimensions with an Optimized Basis Set and Rectangular Collocation. *J. Phys. Chem. A* **2015**, *119*, 9557–9567.
- (95) Peluzo, B. M. T. C.; Galvão, B. R. L. Theoretical study on the structure and reactions of uranium fluorides. *J. Mol. Model.* **2018**, *24*, No. 197.
- (96) Gagliardi, L.; Willetts, A.; Skylaris, C.-K.; Handy, N. C.; Spencer, S.; Ioannou, A. G.; Simper, A. M. A Relativistic Density Functional Study on the Uranium Hexafluoride and Plutonium

Hexafluoride Monomer and Dimer Species. *J. Am. Chem. Soc.* **1998**, *120*, 11727–11731.

(97) Monard, J. A.; Huray, P. G.; Thomson, J. O. Mössbauer studies of electric hyperfine interactions in ^{234}U , ^{236}U , ^{238}U . *Phys. Rev. B* **1974**, *9*, 2838–2845.

(98) Belanzoni, P.; Baerends, E. J.; Van Lenthe, E. The uranyl ion revisited: the electric field gradient at U as a probe of environmental effects. *Mol. Phys.* **2005**, *103*, 775–787.

(99) Autschbach, J.; Peng, D.; Reiher, M. Two-Component Relativistic Calculations of Electric-Field Gradients Using Exact Decoupling Methods: Spin-orbit and Picture-Change Effects. *J. Chem. Theory Comput.* **2012**, *8*, 4239–4248.

(100) Autschbach, J. Relativistic calculations of magnetic resonance parameters: background and some recent developments. *Philos. Trans. R. Soc., A* **2014**, *372*, No. 20120489.

(101) Autschbach, J.; Zheng, S.; Schurko, R. W. Analysis of electric field gradient tensors at quadrupolar nuclei in common structural motifs. *Concepts Magn. Reson., Part A* **2010**, *36A*, 84–126.

(102) Srebro, M.; Autschbach, J. Does a Molecule-Specific Density Functional Give an Accurate Electron Density? The Challenging Case of the CuCl Electric Field Gradient. *J. Phys. Chem. Lett.* **2012**, *3*, 576–581. PMID: 26286152.

(103) Tecmer, P.; Gomes, A. S. P.; Ekström, U.; Visscher, L. Electronic spectroscopy of UO_2^{2+} , NUO^+ and NUN : an evaluation of time-dependent density functional theory for actinides. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6249–6259.

Annexe **E**

Diagrammes : Linear Response pour
la polarisabilité (α) et Matrice densité
Coupled-Cluster (Double)

Diagrammes

$P_-(ab)$, opérateur de permutation $P_-(ab)f(\dots, a, b, \dots) = f(\dots, a, b, \dots) - f(\dots, b, a, \dots)$

h : nombre de *hole line* ; l : nombre de *loop*.

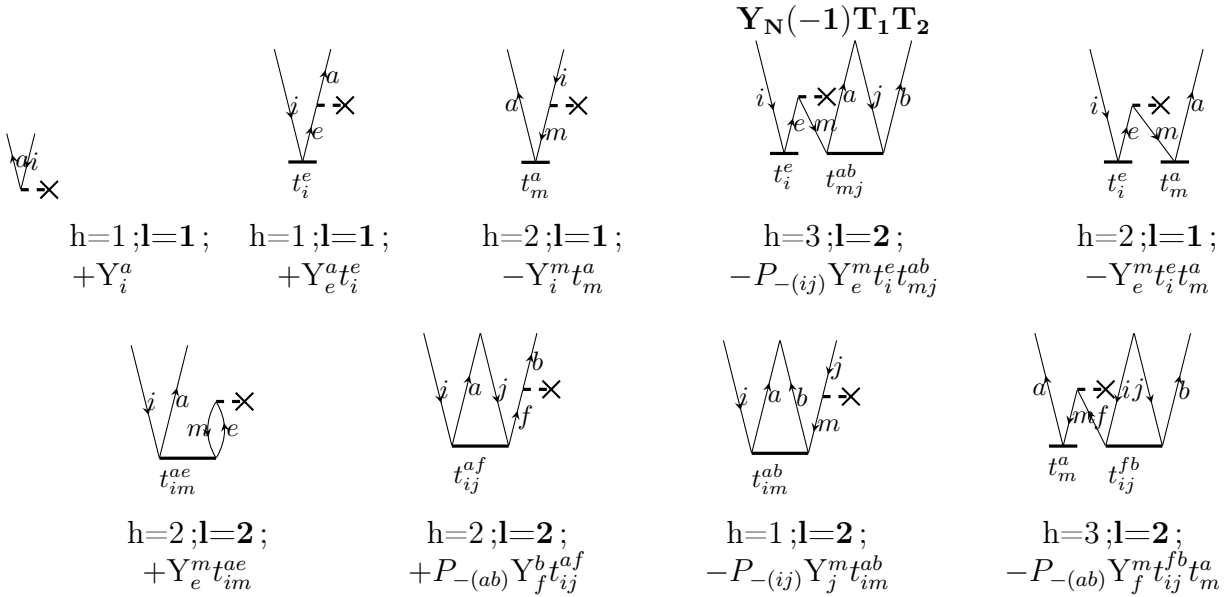
$\{a \equiv b\}$: équivalence des lignes a et b .

$\{ab \equiv cd\}$: équivalence de couple de ligne ab et cd .

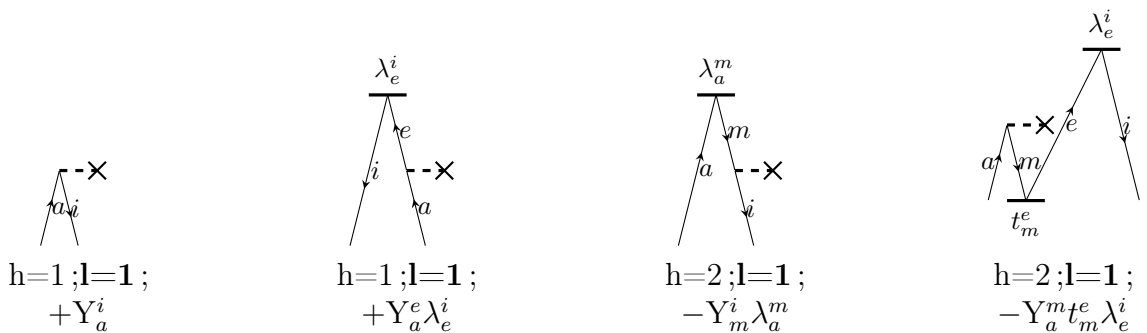
L'intermédiaire \mathcal{W}_{vv}^{oo} est V_{vv}^{oo} .

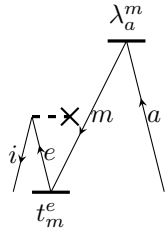
Diagrammes pour LR-CC

ξ_μ^Y

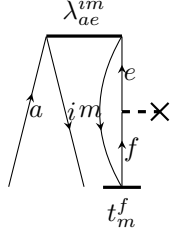


η_μ^Y

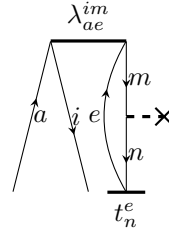




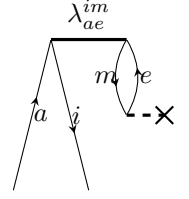
$$h=2; l=1; \\ -Y_e^{ie} t_m^e \lambda_a^m$$



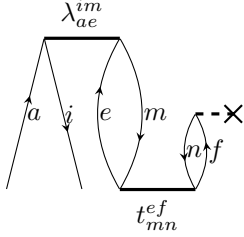
$$h=2; l=2; \\ +t_m^f \lambda_{ae}^{im} Y_f^e$$



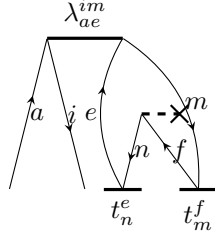
$$h=3; l=2; \\ -t_n^e \lambda_{ae}^{im} Y_m^n$$



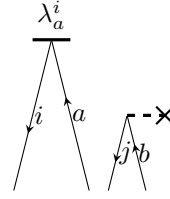
$$h=2; l=2; \\ +\lambda_{ae}^{im} Y_m^e$$



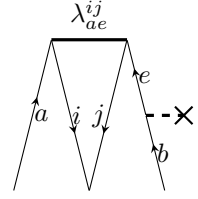
$$h=3; l=3; \\ +t_{mn}^{ef} \lambda_{ae}^{im} Y_f^n$$



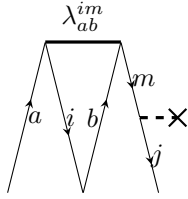
$$h=3; l=2; \\ -t_n^e t_m^f \lambda_{ae}^{im} Y_f^n$$



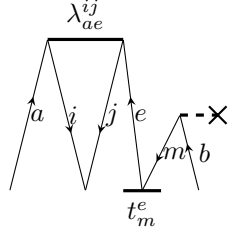
$$h=2; l=2; \\ +P_{-(ij)} P_{-(ab)} \lambda_a^i Y_b^j$$



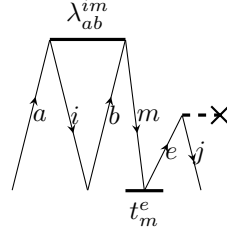
$$h=2; l=2; \\ +P_{-(ab)} \lambda_{ae}^{ij} Y_b^e$$



$$h=3; l=2; \\ -P_{-(ij)} \lambda_{ab}^{im} Y_m^j$$

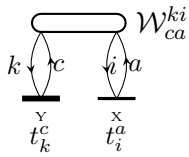


$$h=3; l=2; \\ -P_{-(ab)} t_m^e \lambda_{ae}^{ij} Y_b^m$$

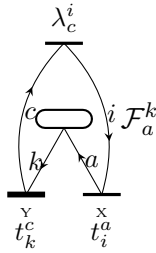


$$h=3; l=2; \\ -P_{-(ij)} t_m^e \lambda_{ab}^{im} Y_e^j$$

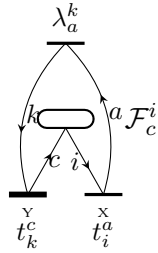
$t^X F$



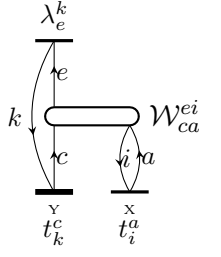
$$h=2; l=2; \\ +t_k^y t_i^x \mathcal{W}_{ca}^{ki} \\ B_c^k = +t_i^x \mathcal{W}_{ca}^{ki}$$



$$h=2; l=1; \\ -\lambda_c^i t_k^y t_i^x \mathcal{F}_a^k \\ B_c^k = -\lambda_c^i t_i^x \mathcal{F}_a^k$$



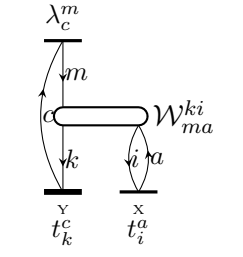
$$h=2; l=1; \\ -\lambda_a^k t_k^y t_i^x \mathcal{F}_c^i \\ B_c^k = -\lambda_a^k t_i^x \mathcal{F}_c^i$$



$$h=2; l=2;$$

$$+\lambda_e^k t_k^c t_i^a W_{ca}^{ei}$$

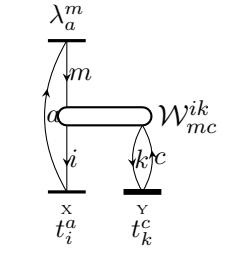
$$B_c^k = +\lambda_e^k t_i^a W_{ca}^{ei}$$



$$h=3; l=2;$$

$$-\lambda_c^m t_k^c t_i^a W_{ma}^{ki}$$

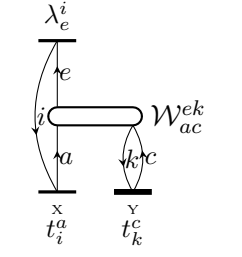
$$B_c^k = -\lambda_c^m t_i^a W_{ma}^{ki}$$



$$h=3; l=2;$$

$$-\lambda_a^m t_k^c t_i^a W_{mc}^{ik}$$

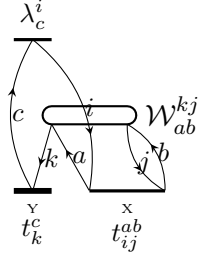
$$B_c^k = -\lambda_a^m t_i^a W_{mc}^{ik}$$



$$h=2; l=2;$$

$$+\lambda_e^i t_k^c t_i^a W_{ac}^{ek}$$

$$B_c^k = +\lambda_e^i t_i^a W_{ac}^{ek}$$

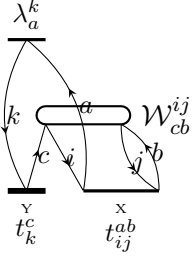


$$h=3; l=2;$$

$$\{a \equiv b\};$$

$$-\frac{1}{2} t_k^c \lambda_c^i t_{ij}^{ab} W_{ab}^{kj}$$

$$B_c^k = -\frac{1}{2} \lambda_c^i t_{ij}^{ab} W_{ab}^{kj}$$

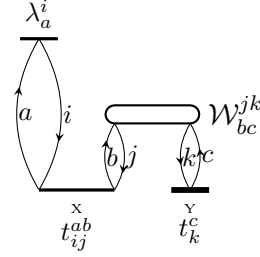


$$h=3; l=2;$$

$$\{i \equiv j\};$$

$$-\frac{1}{2} \lambda_a^k t_k^c t_{ij}^{ab} W_{cb}^{ij}$$

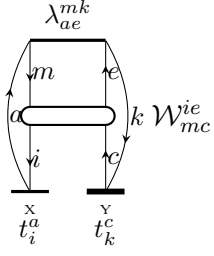
$$B_c^k = -\frac{1}{2} \lambda_a^k t_{ij}^{ab} W_{cb}^{ij}$$



$$h=3; l=3;$$

$$+\lambda_a^i t_k^c t_{ij}^{ab} W_{bc}^{jk}$$

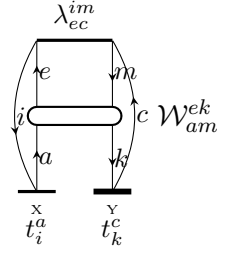
$$B_c^k = +\lambda_a^i t_{ij}^{ab} W_{bc}^{jk}$$



$$h=3; l=2;$$

$$-\lambda_{ae}^{mk} t_i^a t_k^c W_{mc}^{ie}$$

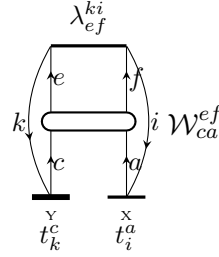
$$B_c^k = -\lambda_{ae}^{mk} t_i^a W_{mc}^{ie}$$



$$h=3; l=2$$

$$-\lambda_{ec}^{im} t_k^c t_i^a W_{am}^{ek}$$

$$B_c^k = -\lambda_{ec}^{im} t_i^a W_{am}^{ek}$$

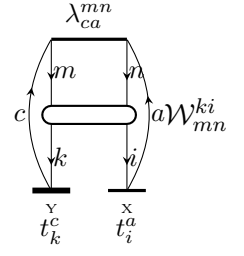


$$h=2; l=2;$$

$$\{e \equiv f\};$$

$$+\frac{1}{2} \lambda_{ef}^{ki} t_k^c t_i^a W_{ca}^{ef}$$

$$B_c^k = +\frac{1}{2} \lambda_{ef}^{ki} t_i^a W_{ca}^{ef}$$

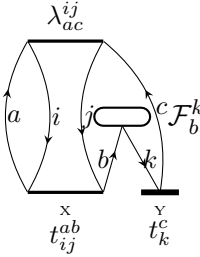


$$h=2; l=2;$$

$$\{m \equiv n\};$$

$$+\frac{1}{2} \lambda_{ca}^{mn} t_k^c t_i^a W_{mn}^{ki}$$

$$B_c^k = +\frac{1}{2} \lambda_{ca}^{mn} t_i^a W_{mn}^{ki}$$

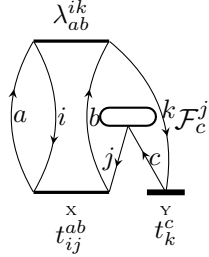


$$h=3; l=2;$$

$$\{i \equiv j\};$$

$$-\frac{1}{2} \lambda_{ac}^{ij} t_k^c t_{ij}^{ab} F_b^k$$

$$B_c^k = -\frac{1}{2} \lambda_{ac}^{ij} t_{ij}^{ab} F_b^k$$

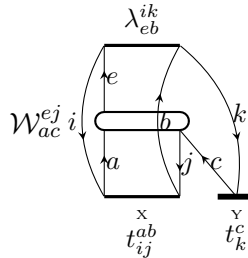


$$h=3; l=2;$$

$$\{a \equiv b\};$$

$$-\frac{1}{2} \lambda_{ab}^{ik} t_k^c t_{ij}^{ab} F_c^j$$

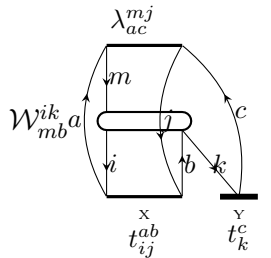
$$B_c^k = -\frac{1}{2} \lambda_{ab}^{ik} t_{ij}^{ab} F_c^j$$



$$h=3; l=2;$$

$$-\lambda_{eb}^{ik} t_{ij}^{ab} t_k^c W_{ac}^{ej}$$

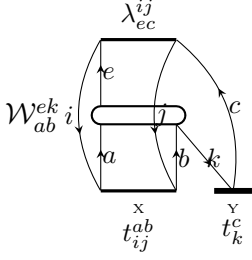
$$B_c^k = -\lambda_{eb}^{ik} t_{ij}^{ab} W_{ac}^{ej}$$



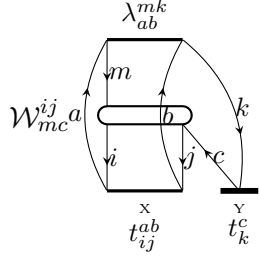
$$h=4; l=2;$$

$$+\lambda_{ac}^{mj} t_{ij}^{ab} t_k^c W_{mb}^{ik}$$

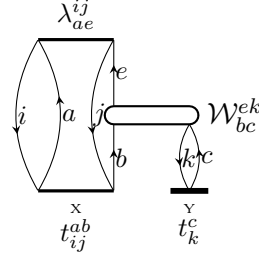
$$B_c^k = +\lambda_{ac}^{mj} t_{ij}^{ab} W_{mb}^{ik}$$



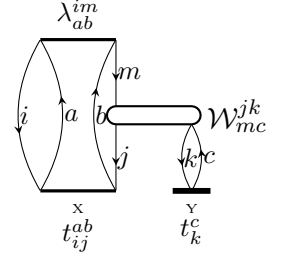
$$\begin{aligned}
 & h=3; l=2; \\
 & \{a \equiv b\} \{i \equiv j\}; \\
 & -\frac{1}{4} \lambda_{ec}^{ij} t_{ij}^a t_{ij}^b t_k^c \mathcal{W}_{ab}^{ek} \\
 & B_c^k = -\frac{1}{4} \lambda_{ec}^{ij} t_{ij}^a t_{ij}^b \mathcal{W}_{ab}^{ek}
 \end{aligned}$$



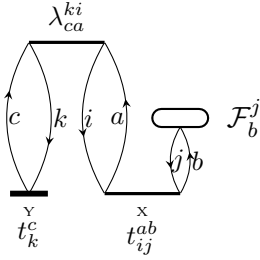
$$\begin{aligned}
 & h=4; l=2; \\
 & \{i \equiv j\} \{a \equiv b\} \\
 & +\frac{1}{4} \lambda_{ab}^{mk} t_{ij}^a t_{ij}^b t_k^c \mathcal{W}_{mc}^{ij} \\
 & B_c^k = +\frac{1}{4} \lambda_{ab}^{mk} t_{ij}^a t_{ij}^b \mathcal{W}_{mc}^{ij}
 \end{aligned}$$



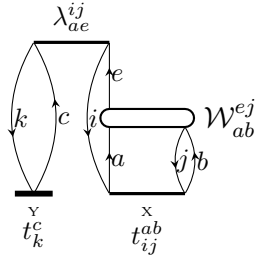
$$\begin{aligned}
 & h=3; l=3; \\
 & \{i \equiv j\} \\
 & +\frac{1}{2} \lambda_{ae}^{ij} t_{ij}^a t_{ij}^b t_k^c \mathcal{W}_{bc}^{ek}; \\
 & B_c^k = +\frac{1}{2} \lambda_{ae}^{ij} t_{ij}^a t_{ij}^b \mathcal{W}_{bc}^{ek}
 \end{aligned}$$



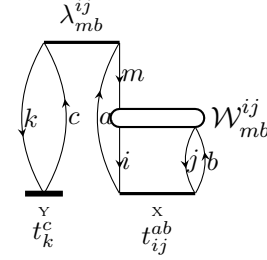
$$\begin{aligned}
 & h=4; l=3; \\
 & \{a \equiv b\}; \\
 & -\frac{1}{2} \lambda_{ab}^{im} t_{ij}^a t_{ij}^b t_k^c \mathcal{W}_{mc}^{jk}; \\
 & B_c^k = -\frac{1}{2} \lambda_{ab}^{im} t_{ij}^a t_{ij}^b \mathcal{W}_{mc}^{jk}
 \end{aligned}$$



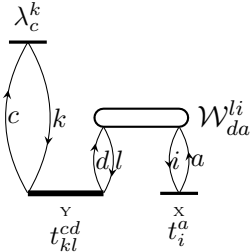
$$\begin{aligned}
 & h=3; l=3; \\
 & +\lambda_{ca}^{ki} t_{ij}^a t_{ij}^b \mathcal{F}_b^j \text{NC}
 \end{aligned}$$



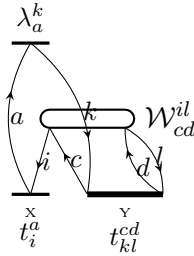
$$\begin{aligned}
 & h=3; l=3; \\
 & +\lambda_{ae}^{ij} t_{ij}^a t_{ij}^b \mathcal{W}_{ab}^{ej} \text{NC}
 \end{aligned}$$



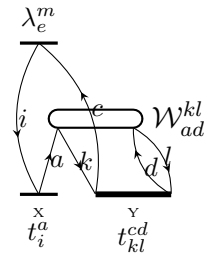
$$\begin{aligned}
 & h=4; l=3; \\
 & -\lambda_{mb}^{ij} t_{ij}^a t_{ij}^b \mathcal{W}_{mb}^{ij} \text{NC}
 \end{aligned}$$



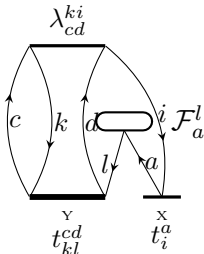
$$\begin{aligned}
 & h=3; l=3; \\
 & +\lambda_{c}^{k} t_{kl}^c t_i^a \mathcal{W}_{da}^{li} \\
 & B_{cd}^{kl} = +P_{-(cd)} P_{-(kl)} \lambda_{c}^{k} t_i^a \mathcal{W}_{da}^{li}
 \end{aligned}$$



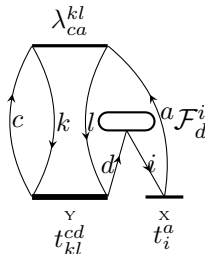
$$\begin{aligned}
 & h=3; l=2; \\
 & \{c \equiv d\} \\
 & -\frac{1}{2} \lambda_a^k t_{kl}^c t_i^a \mathcal{W}_{cd}^{il} \\
 & B_{cd}^{kl} = -P_{-(kl)} \lambda_a^k t_i^a \mathcal{W}_{cd}^{il}
 \end{aligned}$$



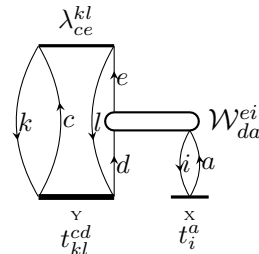
$$\begin{aligned}
 & h=3; l=2; \\
 & \{k \equiv l\} \\
 & -\frac{1}{2} \lambda_c^m t_{kl}^c t_i^a \mathcal{W}_{ad}^{kl} \\
 & B_{cd}^{kl} = -P_{-(cd)} \lambda_c^m t_i^a \mathcal{W}_{ad}^{kl}
 \end{aligned}$$



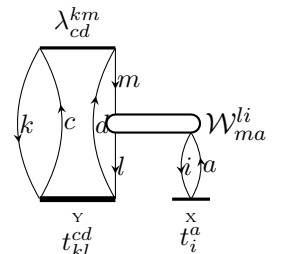
$$\begin{aligned}
 & h=3; l=2; \\
 & \{c \equiv d\}; \\
 & -\frac{1}{2} \lambda_{cd}^{ki} t_{kl}^c t_i^a \mathcal{F}_a^l \\
 & B_{cd}^{kl} = -P_{-(kl)} \lambda_{cd}^{ki} t_i^a \mathcal{F}_a^l
 \end{aligned}$$



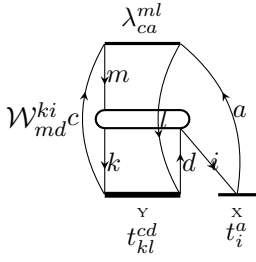
$$\begin{aligned}
 & h=3; l=2; \\
 & \{k \equiv l\}; \\
 & -\frac{1}{2} \lambda_{ca}^{kl} t_{kl}^c t_i^a \mathcal{F}_d^i \\
 & B_{cd}^{kl} = -P_{-(cd)} \lambda_{ca}^{kl} t_i^a \mathcal{F}_d^i
 \end{aligned}$$



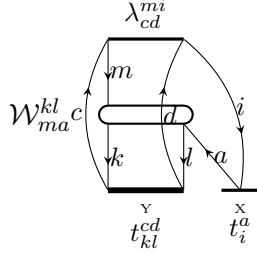
$$\begin{aligned}
 & h=3; l=3; \\
 & \{k \equiv l\} \\
 & +\frac{1}{2} \lambda_{ce}^{kl} t_{kl}^c t_i^a \mathcal{W}_{da}^{ei}; \\
 & B_{cd}^{kl} = +P_{-(kd)} \lambda_{ce}^{kl} t_i^a \mathcal{W}_{da}^{ei}
 \end{aligned}$$



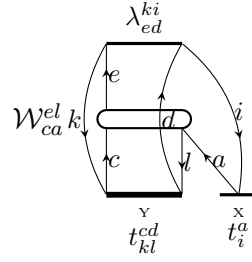
$$\begin{aligned}
 & h=4; l=3; \\
 & \{c \equiv d\} \\
 & -\frac{1}{2} \lambda_{cd}^{km} t_{kl}^c t_i^a \mathcal{W}_{ma}^{li}; \\
 & B_{cd}^{kl} = \\
 & -P_{-(kl)} \lambda_{cd}^{km} t_i^a \mathcal{W}_{ma}^{li}
 \end{aligned}$$



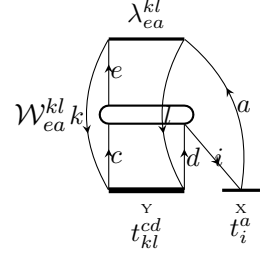
$$\begin{aligned}
 & h=4; l=2; \\
 & +\lambda_{ca}^{ml} t_{kl}^{cd} t_i^a \mathcal{W}_{md}^{ki} \\
 & B_{cd}^{kl} = \\
 & +P_{-(cd)} P_{-(kl)} \lambda_{ca}^{ml} t_i^a \mathcal{W}_{md}^{ki}
 \end{aligned}$$



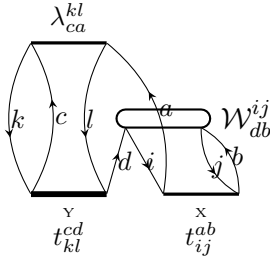
$$\begin{aligned}
 & h=4; l=2; \\
 & \{c \equiv d\}; \{k \equiv l\} \\
 & +\frac{1}{4} \lambda_{cd}^{mi} t_{kl}^{cd} t_i^a \mathcal{W}_{ma}^{kl} \\
 & B_{cd}^{kl} = +\lambda_{cd}^{mi} t_i^a \mathcal{W}_{ma}^{kl}
 \end{aligned}$$



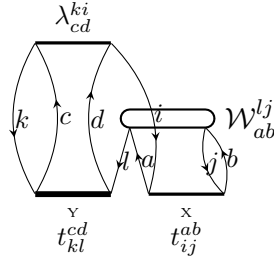
$$\begin{aligned}
 & h=3; l=2; \\
 & -\lambda_{ed}^{ki} t_{kl}^{cd} t_i^a \mathcal{W}_{ca}^{el} \\
 & B_{cd}^{kl} = \\
 & -P_{-(cd)} P_{-(kl)} \lambda_{ed}^{ki} t_i^a \mathcal{W}_{ca}^{el}
 \end{aligned}$$



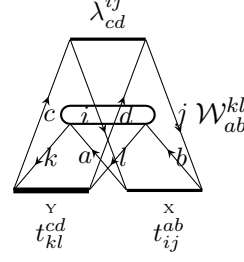
$$\begin{aligned}
 & h=3; l=2; \\
 & \{k \equiv l\}; \{c \equiv d\} \\
 & -\frac{1}{4} \lambda_{ea}^{kl} t_{kl}^{cd} t_i^a \mathcal{W}_{cd}^{ei} \\
 & B_{cd}^{kl} = -\lambda_{ea}^{kl} t_i^a \mathcal{W}_{cd}^{ei}
 \end{aligned}$$



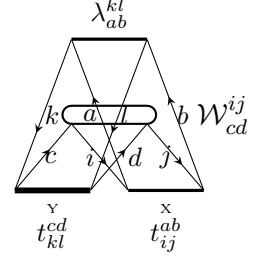
$$\begin{aligned}
 & h=4; l=3; \\
 & \{k \equiv l\} \{i \equiv j\} \\
 & -\frac{1}{4} \lambda_{ca}^{kl} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{db}^{ij} \\
 & B_{cd}^{kl} = \\
 & -\frac{1}{2} P_{-(cd)} \lambda_{ca}^{kl} t_{ij}^{ab} \mathcal{W}_{db}^{ij}
 \end{aligned}$$



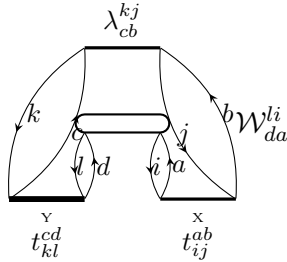
$$\begin{aligned}
 & h=4; l=3; \\
 & \{c \equiv d\} \{a \equiv b\}; \\
 & -\frac{1}{4} \lambda_{cd}^{ki} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{ab}^{lj} \\
 & B_{cd}^{kl} = \\
 & -\frac{1}{2} P_{-(kl)} \lambda_{cd}^{ki} t_{ij}^{ab} \mathcal{W}_{ab}^{lj}
 \end{aligned}$$



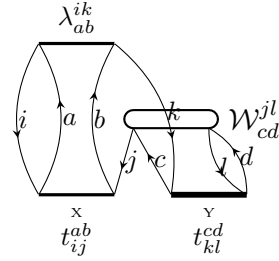
$$\begin{aligned}
 & h=4; l=2; \\
 & \{a \equiv b\} \{c \equiv d\} \\
 & \{i \equiv j\} \{k \equiv l\} \\
 & +\frac{1}{16} \lambda_{cd}^{ij} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{ab}^{kl} \\
 & B_{cd}^{kl} = +\frac{1}{4} \lambda_{cd}^{ij} t_{ij}^{ab} \mathcal{W}_{ab}^{kl}
 \end{aligned}$$



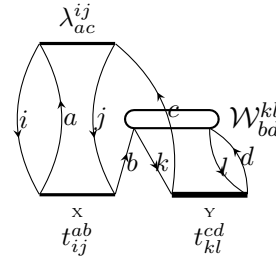
$$\begin{aligned}
 & h=4; l=0; 2 \\
 & \{a \equiv b\} \{c \equiv d\} \\
 & \{i \equiv j\} \{k \equiv l\} \\
 & +\frac{1}{16} \lambda_{ab}^{kl} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{cd}^{ij} \\
 & B_{cd}^{kl} = +\frac{1}{4} \lambda_{ab}^{kl} t_{ij}^{ab} \mathcal{W}_{cd}^{ij}
 \end{aligned}$$



$$\begin{aligned}
 & h=4; l=4; \\
 & +\lambda_{cb}^{kj} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{da}^{li} \\
 & B_{cd}^{kl} = \\
 & +P_{-(kl)} P_{-(cd)} \lambda_{cb}^{kj} t_{ij}^{ab} \mathcal{W}_{da}^{li}
 \end{aligned}$$



$$\begin{aligned}
 & h=4; l=3; \\
 & \{a \equiv b\} \{c \equiv d\} \\
 & -\frac{1}{4} \lambda_{ab}^{ik} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{cd}^{jl} \\
 & B_{cd}^{kl} = \\
 & -\frac{1}{2} P_{-(kl)} \lambda_{ab}^{ik} t_{ij}^{ab} \mathcal{W}_{cd}^{jl}
 \end{aligned}$$

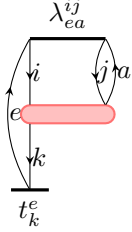


$$\begin{aligned}
 & h=4; l=3; \\
 & \{i \equiv j\} \{k \equiv l\} \\
 & -\frac{1}{4} \lambda_{ac}^{ij} t_{kl}^{cd} t_{ij}^{ab} \mathcal{W}_{bd}^{kl} \\
 & B_{cd}^{kl} = \\
 & -\frac{1}{2} P_{-(cd)} \lambda_{ac}^{ij} t_{ij}^{ab} \mathcal{W}_{bd}^{kl}
 \end{aligned}$$

Diagrammes pour la matrice densité double CCSD

• Γ_{ov}^{oo} Γ_{ka}^{ij}

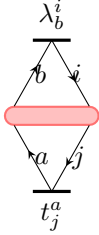
$$\Gamma_{ka}^{ij} = -\lambda_{ea}^{ij} t_k^e \quad (1)$$



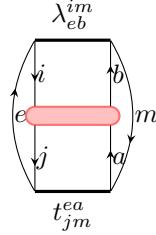
$h=3; l=2;$
 $-\lambda_{ea}^{ij} t_k^e$

• Γ_{ov}^{ov} Γ_{bj}^{ai}

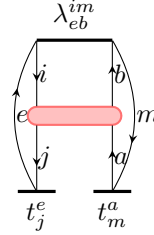
$$\Gamma_{bj}^{ai} = -\lambda_b^i t_j^a \{f6\} - \lambda_{eb}^{im} t_{jm}^{ea} - \lambda_{eb}^{im} t_e^j t_m^a \quad (2)$$



$h=2; l=2;$
 $\Gamma_{jb}^{ai} = +\lambda_b^i t_j^a$



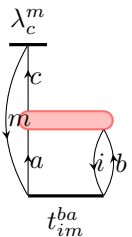
$h=3; l=2;$
 $-\lambda_{eb}^{im} t_{jm}^{ea}$



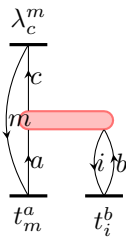
$h=3; l=2;$
 $-\lambda_{eb}^{im} t_e^j t_m^a$

• Γ_{vo}^{vv} Γ_{ci}^{ab}

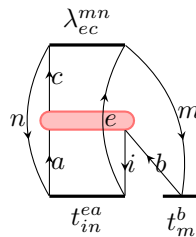
$$\begin{aligned} \Gamma_{ci}^{ab} = & +\lambda_c^m t_{im}^{ba} + P_-(ab) \lambda_c^m t_i^b t_m^a - P_-(ab) \lambda_{ec}^{mn} t_m^b t_{in}^{ea} - \frac{1}{2} \lambda_{ec}^{mn} t_i^e t_{mn}^{ba} \\ & + \frac{1}{2} P_-(ab) \lambda_{ec}^{mn} t_i^b t_{mn}^{ea} - \frac{1}{2} P_-(ab) \lambda_{ec}^{mn} t_i^e t_m^b t_n^a \end{aligned} \quad (3)$$



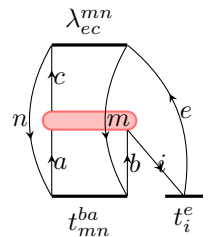
$h=2; l=2;$
 $+\lambda_c^m t_{im}^{ba}$



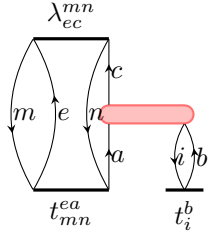
$h=2; l=2;$
 $+P_-(ab) \lambda_c^m t_i^b t_m^a$



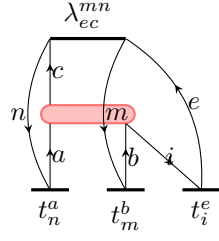
$h=3; l=2;$
 $-P_-(ab) \lambda_{ec}^{mn} t_{ni}^e t_m^b$



$h=3; l=2; \{m=n\};$
 $-\frac{1}{2} \lambda_{ec}^{mn} t_{mn}^{ba} t_i^e$



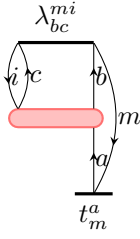
$$h=3; l=3; \{m \equiv n\} \\ + \frac{1}{2} P_{-(ab)} \lambda_{ec}^{mn} t_i^b t_{mn}^{ea};$$



$$h=3; l=2; \\ \{na \equiv mb\} \\ - \frac{1}{2} P_{-(ab)} \lambda_{ec}^{mn} t_i^e t_m^a t_n^a$$

$$\bullet \Gamma_{vv}^{vo} \quad \Gamma_{bc}^{ai}$$

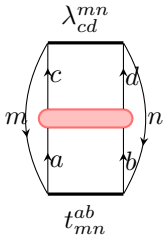
$$\Gamma_{bc}^{ai} = + \lambda_{bc}^{mi} t_m^a \quad (4)$$



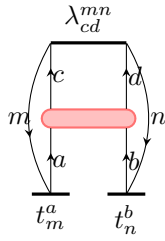
$$h=2; l=2; \\ + \lambda_{bc}^{mi} t_m^a$$

$$\bullet \Gamma_{vv}^{vv} \quad \Gamma_{cd}^{ab}$$

$$\Gamma_{cd}^{ab} = + \frac{1}{2} \lambda_{cd}^{mn} t_{mn}^{ab} + \frac{1}{2} P_{-(ab)} \lambda_{cd}^{mn} t_m^a t_n^b \quad (5)$$



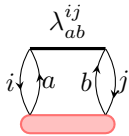
$$h=2; l=2; \{m \equiv n\} \\ + \frac{1}{2} \lambda_{cd}^{mn} t_{mn}^{ab}$$



$$h=2; l=2; \{na \equiv mb\} \\ + \frac{1}{2} P_{-(ab)} \lambda_{cd}^{mn} t_n^b t_m^a$$

$$\bullet \Gamma_{vv}^{oo} \quad \Gamma_{ab}^{ij}$$

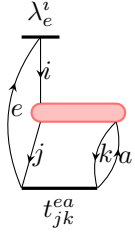
$$\Gamma_{ab}^{ij} = \lambda_{ab}^{ij} \quad (6)$$



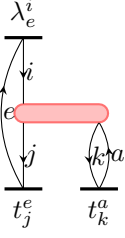
$$h=2; l=2; \\ + \lambda_{ab}^{ij}$$

• Γ_{oo}^{ov} Γ_{jk}^{ia}

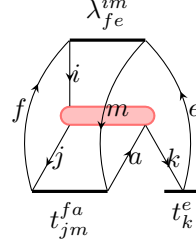
$$\begin{aligned} \Gamma_{jk}^{ia} = & -\lambda_e^i t_{jk}^{ea} - P_-(jk) \lambda_e^i t_j^e t_k^a + P_-(jk) \lambda_{fe}^{im} t_{jm}^f t_k^e + \frac{1}{2} \lambda_{fe}^{im} t_{jk}^f t_m^a \\ & - \frac{1}{2} P_-(jk) \lambda_{fe}^{im} t_k^a t_{jm}^f e + \frac{1}{2} P_-(jk) \lambda_{fe}^{im} t_m^f t_j^e t_k^e \end{aligned} \quad (7)$$



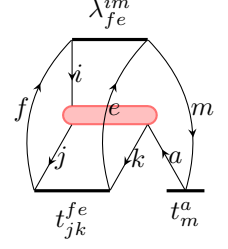
$h=3; l=2;$
 $-\lambda_e^i t_{jk}^{ea}$



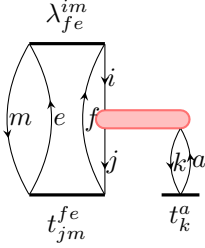
$h=3; l=2;$
 $-P_-(jk) \lambda_e^i t_j^e t_k^a$



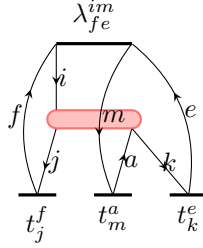
$h=4; l=2;$
 $+P_-(jk) \lambda_{fe}^{im} t_{jm}^f t_k^e$



$h=4; l=2; \{e \equiv f\}$
 $+\frac{1}{2} \lambda_{fe}^{im} t_{jk}^f t_m^a$



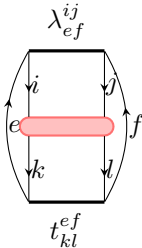
$h=4; l=3; \{e \equiv f\}$
 $-\frac{1}{2} P_-(jk) \lambda_{fe}^{im} t_k^a t_{jm}^f e$



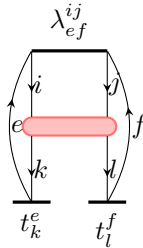
$h=4; l=2; \{ke \equiv jf\}$
 $+\frac{1}{2} P_-(jk) \lambda_{fe}^{im} t_m^a t_j^f t_k^e$

• Γ_{oo}^{oo} Γ_{kl}^{ij}

$$\Gamma_{kl}^{ij} = +\frac{1}{2} \lambda_{ef}^{ij} t_{kl}^{ef} + \frac{1}{2} P_-(kl) \lambda_{ef}^{ij} t_k^e t_l^f \quad (8)$$



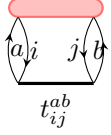
$h=4; l=2; \{e \equiv f\}$
 $+\frac{1}{2} \lambda_{ef}^{ij} t_{kl}^{ef}$



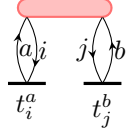
$h=4; l=2; \{ke \equiv lf\}$
 $+\frac{1}{2} P_-(kl) \lambda_{ef}^{ij} t_k^e t_l^f$

• Γ_{oo}^{vv} Γ_{ij}^{ab}

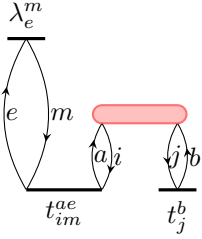
$$\begin{aligned}
\Gamma_{ij}^{ab} = & +t_{ij}^{ab} + \frac{1}{2}P_{-}(ij)P_{-}(ab)t_i^a t_j^b + P_{-}(ij)P_{-}(ab)\lambda_e^m t_{im}^a t_j^b - P_{-}(ij)P_{-}(ab)\lambda_e^m t_i^a t_m^b \\
& - P_{-}(ij)\lambda_e^m t_i^e t_m^{ab} - \frac{1}{2}P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{mn}^e t_i^a t_j^b - \frac{1}{2}P_{-}(ab)\lambda_{ef}^{mn} t_{mn}^e t_i^a t_j^b \\
& - \frac{1}{2}P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{im}^e t_n^a t_j^b - \frac{1}{2}P_{-}(ij)\lambda_{ef}^{mn} t_{im}^e t_n^a t_j^b - P_{-}(ab)\lambda_e^m t_m^a t_{ij}^{eb} \\
& + \frac{1}{4}\lambda_{ef}^{mn} t_{ij}^e t_{mn}^{ab} + \frac{1}{4}P_{-}(ab)\lambda_{ef}^{mn} t_{ij}^e t_m^a t_n^b + \frac{1}{4}P_{-}(ij)\lambda_{ef}^{mn} t_{mn}^e t_i^a t_j^b \\
& + P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{im}^e t_{jn}^{bf} - P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{im}^e t_j^a t_n^b + \frac{1}{4}P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_i^e t_m^a t_n^b t_j^f \quad (9)
\end{aligned}$$



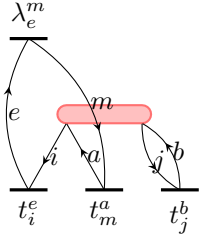
$$h=2; l=2; \\ +t_{ij}^{ab}$$



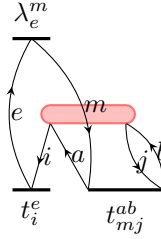
$$h=2; l=2; \{ai \equiv bj\} \\ +\frac{1}{2}P_{-}(ab)P_{-}(ij)t_i^a t_j^b$$



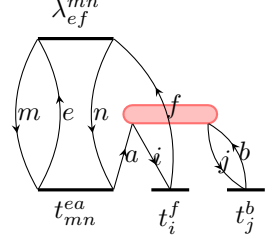
$$h=3; l=3; \\ +P_{-}(ij)P_{-}(ab)\lambda_e^m t_{im}^{ae} t_j^b$$



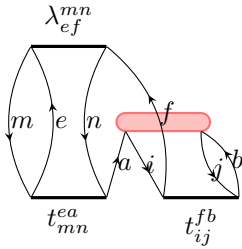
$$h=3; l=2; \\ -P_{-}(ij)P_{-}(ab)\lambda_e^m t_i^e t_m^a t_j^b$$



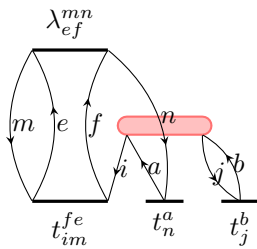
$$h=3; l=2; \\ -P_{-}(ij)\lambda_e^m t_i^e t_{mj}^{ab}$$



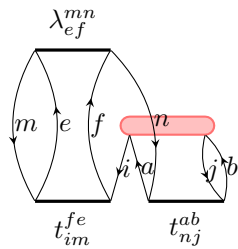
$$h=4; l=3; \{m \equiv n\} \\ -\frac{1}{2}P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{mn}^e t_i^a t_j^b$$



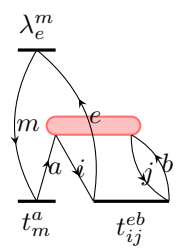
$$h=4; l=3; \{m \equiv n\} \\ -\frac{1}{2}P_{-}(ab)\lambda_{ef}^{mn} t_{mn}^e t_{ij}^{fb}$$



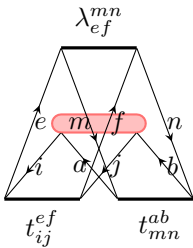
$$h=4; l=3; \{e \equiv f\} \\ -\frac{1}{2}P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{im}^e t_n^a t_j^b$$



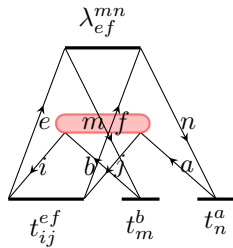
$$h=4; l=3; \{e \equiv f\} \\ -\frac{1}{2}P_{-}(ij)\lambda_{ef}^{mn} t_{im}^e t_{nj}^{ab}$$



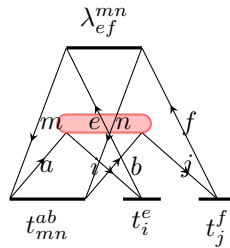
$$h=3; l=2; \\ -P_{-}(ab)\lambda_e^m t_m^a t_{ij}^{eb}$$



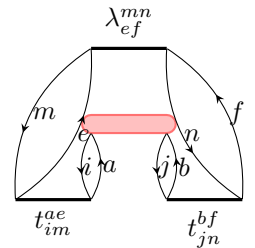
$$h=4; l=2; \{e \equiv f\} \\ * \{m \equiv n\} \\ +\frac{1}{4}\lambda_{ef}^{mn} t_{ij}^e t_{mn}^{ab}$$



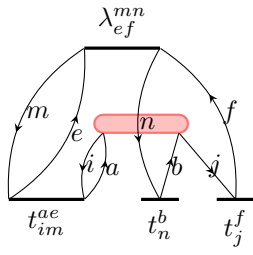
$$h=4; l=2; \{e \equiv f\} * \\ \{m \equiv n\} \\ +\frac{1}{4}P_{-}(ab)\lambda_{ef}^{mn} t_{ij}^e t_m^b t_n^a$$



$$h=4; l=2; \{m \equiv n\} \\ * \{i \equiv j\} \\ +\frac{1}{4}P_{-}(ij)\lambda_{ef}^{mn} t_{mn}^e t_i^e t_j^f$$

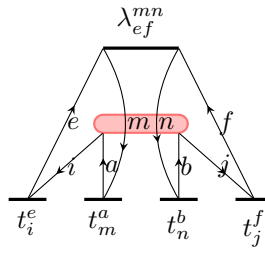


$$h=4; l=4 \\ +P_{-}(ij)P_{-}(ab)\lambda_{ef}^{mn} t_{im}^e t_{jn}^{bf}$$



$$h=4; l=3$$

$$-P_{-(ij)}P_{-(ab)}\lambda_{ef}^{mn}t_{im}^{ae}t_j^ft_n^b$$



$$h=4; l=2; \{ma \equiv nb\} * \{ie \equiv jf\}$$

$$+\frac{1}{4}P_{-(ij)}P_{-(ab)}\lambda_{ef}^{mn}t_i^et_m^at_n^bt_j^f$$

Table des matières

Résumé	xv
Remerciements	xvii
Sommaire	xix
Liste des tableaux	xxiii
Table des figures	xxv
Introduction	1
1 Méthodes	5
1.1 Méthode HARTREE-FOCK	5
1.1.1 Hamiltonien non-relativiste	5
1.1.2 Fonction d'onde	7
1.1.3 Énergie HARTREE-FOCK	8
1.1.4 Expansion en fonction de base, Basis Set	9
1.1.5 Size-Consistent Size-Extensive	11
1.2 La Méthode full Configuration Interaction fCI	11
1.2.1 Énergie de corrélation	12
1.2.2 Méthodes multi-références (MR)	13
1.3 Théorie perturbative	15
1.3.1 RAYLEIGH-SCHRÖDINGER <i>Perturbation Theory</i>	15
1.3.2 MÖLLER-PLESSET <i>Perturbation Theory</i>	16
1.4 Méthode Coupled-Cluster	17
1.4.1 Énergie et Amplitudes	17
1.4.2 Approche diagrammatique	20
1.5 Equation of Motion Coupled-Cluster	22
1.5.1 Description générale	23

1.5.2	Énergie d'excitation, Potentiel d'ionisation, Affinité électronique	24
1.5.3	Matrice EOM-CCSD	25
1.5.4	EOM-CC et l'algorithme de Davidson	26
1.6	Relativité restreinte	28
1.6.1	Solution pour un système mono-électronique	29
1.6.2	Solution pour un système poly-électronique	31
1.6.3	Hamiltonien de BREIT - Électrodynamique Quantique	32
1.6.4	Approximation à 2-composantes de l'hamiltonien du Dirac-Coulomb : ${}^2\text{DC}$ - ${}^2\text{DC}^{\text{M}}$	33
2	Les différentes problématiques	35
	PROBLÉMATIQUE 1 : OBTENIR LES IP ET LES EE DE CŒUR	35
2.1	<i>X-ray Photoelectron Spectroscopy (XPS)</i>	36
2.1.1	Présentation d'une expérience XPS	36
2.1.2	Expérimentalement	37
2.2	Chimie quantique et potentiels d'ionisation des électrons de cœur	40
2.2.1	Lien entre les calculs et les expériences XPS	40
2.2.2	Classes de méthode	41
2.2.3	Vers le chapitre 3	44
	PROBLÉMATIQUE 2 : QUELLES APPROXIMATIONS RÉALISER?	45
2.3	Approximations	45
2.3.1	Approximations dans l'origine des intégrales	45
2.3.2	Approximations perturbatives des termes de la matrice \mathcal{H}_{eff}	47
2.3.3	Vers le chapitre 4	49
	PROBLÉMATIQUE 3 : EXPLOITER DE FAÇON EFFICACE DES PLATEFORMES HPC POUR DES CALCULS CORRÉLÉS	50
2.4	High-Performance Computing (HPC)	50
2.4.1	Hardware	51
2.4.2	Software	53
2.4.3	Vers le chapitre 5	55
3	EOM-(IP; EE; EA) - CVS	57
3.1	Utilisation de CVS avec EOM-CC	58
3.2	Options pour CVS : <i>Frozen Core</i> (FC) et <i>No-Double</i> (ND)	60
3.3	Influence de l'hamiltonien	64
3.4	Conclusion	67
4	{P}-EOM-CCSD-{MBPT(2)}	69
4.1	Les différentes approximations	70
4.2	Détails computationnels et présentation des résultats	72

Table des matières	207
4.3 Potentiels d'ionisation	73
4.3.1 Triiodure I_3^-	73
4.3.2 Les ions hypohalogénites XO^- , $X \in \{Cl; Br; I; At; Ts\}$. . .	74
4.4 Énergie d'excitation et affinité électronique	86
4.5 Conclusion	89
4.6 Remarques sur les coûts computationnels	91
5 ExaCorr	93
5.1 Une nouvelle bibliothèque : ExaTensor	94
5.2 Exemple d'expression informatique pour les calculs tensoriels .	95
5.3 Résultats sur Summit	96
5.3.1 La symétrie	96
5.3.2 Complexe d'uranyl : le trinitrate d'uranyl $[UO_2(NO_3)_3]^-$.	97
5.4 CC2	99
5.4.1 Coût computationnel	100
5.4.2 Résultats préliminaires	101
5.4.3 Travaux en cours de réalisation et perspectives vers ExaTensor	104
5.5 LR-CCSD - α	107
5.5.1 Polarisabilité dans le cadre de CC	108
5.5.2 Expressions programmables	111
5.6 Conclusion	114
Conclusion	115
Bibliographie	117
A Équations	143
A.1 Équations CC2	143
A.2 Méthodes approximées	144
A.2.1 EOM-MBPT(2)	144
A.2.2 Partitioned-EOM	145
B Annexes du Chapitre 3	147
B.1 Courbes d'énergie potentielle	147
B.1.1 Courbes	147
B.1.2 Tables : Position des minimums - Énergies de transition .	148
B.2 Géométrie des molécules	150
C Article 1 : Relativistic EOM-CCSD and CVS	151
D Article 2 : ExaCorr	169

E Diagrammes	191
Table des matières	205

LA MÉTHODE EQUATION OF MOTION COUPLED CLUSTER POUR LA MODÉLISATION DES ÉTATS
EXCITÉS ET PROPRIÉTÉS DES MOLÉCULES CONTENANT DES ÉLÉMENTS LOURDS

Résumé

Dans cette thèse, nous cherchons à obtenir certaines propriétés moléculaires pour des espèces contenant des éléments lourds ou présentant des intérêts atmosphériques. Pour cela, nous utilisons des techniques permettant de caractériser les électrons de cœur, avec les potentiels d'ionisation (IP) ou avec les énergies d'excitation (EE), offrant la possibilité par exemple d'interpréter respectivement les expériences *X-ray Photoelectron Spectroscopy* (XPS) et *X-ray Absorption Spectroscopy* (XAS). Nous cherchons également à caractériser les électrons de valence au travers de la polarisabilité qui est utilisée par exemple pour développer des champs de force. Quand nous travaillons avec des éléments lourds ou avec des électrons de cœur, il faut prendre en compte les effets relativistes. Nous avons donc employé l'hamiltonien de Dirac-Coulomb(-Gaunt). De plus, pour comparer nos résultats aux expériences, il nous faut des méthodes précises. Ainsi, nous travaillerons avec la méthode *Coupled-Cluster* (CC) et pour obtenir les (IP), les (EE) et également les affinités électroniques (EA), nous utiliserons *Equation of Motion Coupled-Cluster* (EOM-CC). Cependant, ces deux éléments (hamiltoniens à 4-composantes et méthodes post-Hartree-Fock) impliquent des coûts de calcul considérables, nécessitant les ressources de plateformes *High Performance Computing* (HPC).

Cette thèse se présentera donc selon les éléments décrits précédemment. Premièrement, nous étudierons la méthode *Core-Valence Separation* (CVS) qui nous permettra, à partir de EOM-CC, d'atteindre les propriétés des électrons de cœur (IP et EE). Comme ces électrons sont proches du noyau où les effets relativistes sont les plus importants, nous étudierons différents hamiltoniens, notamment *exact two-component molecular mean field Hamiltonian*. Deuxièmement, nous nous intéresserons aux approximations perturbatives (*Partitioned* et *Many Body Perturbation Theory 2d order* (MBPT(2)) à appliquer à la matrice EOM-CC pour limiter les coûts computationnels. Enfin, nous présenterons des travaux réalisés sur ExaCorr, une nouvelle implémentation de Coupled-Cluster relativiste pour les architectures hybrides et massivement parallèles. Nous terminerons en décrivant le formalisme et les équations de travail de la méthode *Linear Response Coupled-Cluster* (LR-CC), grâce à laquelle des polarisabilités moléculaires analytiques (dépendantes de la fréquence) peuvent être obtenues.

Mots clés : chimie quantique, éléments lourds, structure électronique, théorie à n corps

Laboratoire PhLAM

CNRS UMR 8523 – Université de Lille, – F-59655 Villeneuve d'Ascq Cedex – France

THE EQUATION OF MOTION COUPLED CLUSTER METHOD FOR MODELING EXCITED STATES AND PROPERTIES OF MOLECULES CONTAINING HEAVY ELEMENTS

Abstract

In this thesis, we seek to obtain certain molecular properties for species containing heavy elements or presenting atmospheric interests. For this, we use techniques to characterize the core electrons, with ionization potentials (IP) or with excitation energies (EE), allowing for example to respectively interpret *X-ray Photoelectron Spectroscopy* (XPS) and *X-ray Absorption Spectroscopy* (XAS). We also seek to characterize valence electrons through the polarizability, which is used for example to develop force fields. When we work with heavy elements or with core electrons, we must take relativistic effects into account. We therefore used the Dirac-Coulomb(-Gaunt) Hamiltonian. Furthermore, to compare our results with experiments, we need precise methods. Thus, we will work with the *Coupled-Cluster* (CC) method, and will use the *Equation of Motion Coupled-Cluster* (EOM-CC) method to obtain the IPs, EEs and electron affinities (EA). However, these two elements (4-component Hamiltonians and post-Hartree-Fock methods) imply considerable computational costs, requiring the resources of *High Performance Computing* (HPC) platforms.

This thesis presents a study of the *Core-Valence Separation* (CVS) method, which will allow us to reach the properties of core electrons (IP and EE) with EOM-CC. We provide a detailed investigation of the performance of different Hamiltonians, in particular the *exact two-component molecular mean field* Hamiltonian. Second, we will focus on the perturbative approximations (*Partitioned* and *Many Body Perturbation Theory 2d order* (MBPT(2))) to be applied to the EOM-CC matrix to limit computational costs, including for core processes. Finally, we present the work carried out in ExaCorr, a new relativistic coupled cluster implementation for hybrid and massively parallel architectures. We will finish by outlining the formalism and working equations for the *Linear Response Coupled-Cluster* (LR-CC) method, through which analytical (frequency-dependent) molecular polarizabilities can be obtained.

Keywords: quantum chemistry, heavy elements, electronic structure, many-body theory
