# On the calculation of molecular properties of heavy element systems with ab initio approaches: from gas-phase to complex systems

André Severo Pereira Gomes

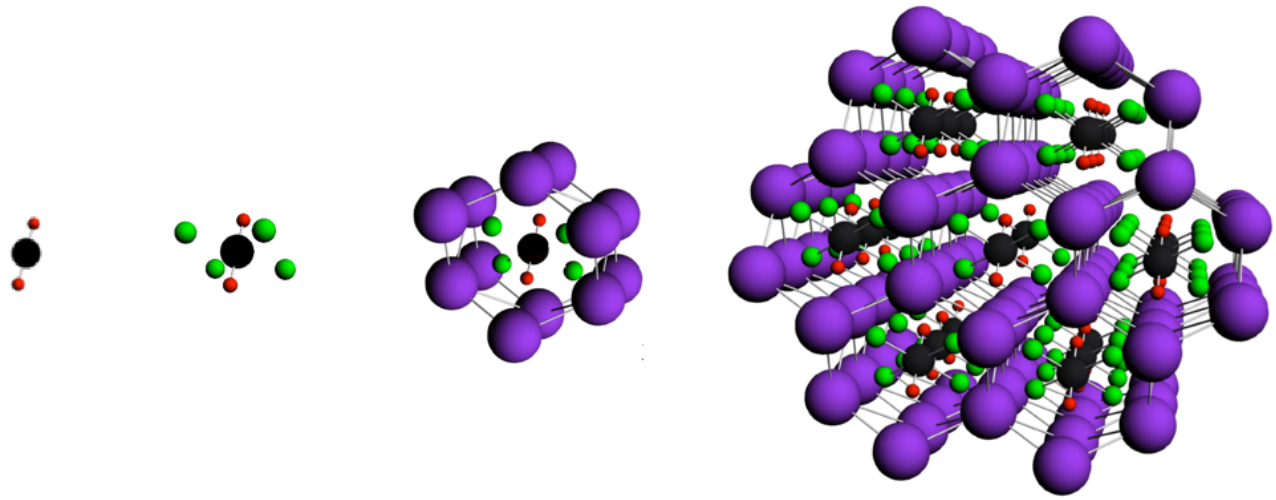
Université de Lille 1 SCIENCES ET TECHNOLOGIES

Mémoire présenté pour obtenir le diplôme d'
Habilitation à diriger des recherches – Sciences Physiques
de l'Université de Lille (Sciences et Technologies)

**ANDRÉ SEVERO PEREIRA GOMES**

Université de Lille - CNRS
Laboratoire PhLAM UMR 8523

# On the calculation of molecular properties of heavy element systems with ab initio approaches: from gas-phase to complex systems



Université de Lille 1 SCIENCES ET TECHNOLOGIES

Mémoire présenté pour obtenir le diplôme d'
Habilitation à diriger des recherches – Sciences Physiques
de l'Université de Lille (Sciences et Technologies)

**ANDRÉ SEVERO PEREIRA GOMES**

Université de Lille - CNRS
Laboratoire PhLAM UMR 8523

# On the calculation of molecular properties of heavy element systems with ab initio approaches: from gas-phase to complex systems

Soutenue publiquement le 16/12/2016,

| Membres du jury | | |
|---|---|---|
| Rapporteur | Hans Jørgen Aagaard Jensen | Assistant Professor, Department of Chemistry and Physics, University of Southern Denmark, Odense, Denmark |
| Rapporteur | Nikolas Kaltsoyannis | Professor and Head of Computational Chemistry, School of Chemistry, The University of Manchester, Manchester, UK |
| Rapporteur | Octavio Roncero | Investigador Científico, Dept. Fisica Atomica, Molecular y de Agregados, Instituto de Fisica Fundamental (C.S.I.C.), Madrid, Spain |
| Examinateur | Sylvain Cristol | Professeur, Université de Lille (Sciences et Technologies), Villeneuve d'Ascq, France |
| Garant de l'habilitation | Valérie Vallet | Directrice de recherche au CNRS, Université de Lille (Sciences et Technologies), Villeneuve d'Ascq, France |

# Contents

# Chapter 1

# Abstract

This work discusses theoretical approaches to model the electronic structure of species containing heavy elements – that is, those from the fifth row onwards on the periodic table – with a special emphasis on lanthanides and actinides, due to their importance in a number of technological issues and applications in fields as diverse as consumer electronics and nuclear energy.

The three key ingredients which should be addressed in modeling of such systems are: (i) relativistic effects, arising from the speeds close to that of the light to which inner electrons are accelerated for heavy nuclei; (ii) electron correlation effects, due to the instantaneous interactions between the electrons as well as due to quasi-degeneracies in the electronic states in heavy element species that often possess unpaired $d$ or $f$ electrons; and (iii) environment effects arising from the interaction of the heavy element species with surrounding molecules, since these are often found in the condensed phase.

We begin by briefly reviewing the approaches used to describe electron correlation and relativistic effects before turning our attention to the four-component intermediate Hamiltonian Fock-space coupled cluster (4c-IHFSCC) method in order to first establish its accuracy with respect to available experimental results and later show how served as a reference method to which more approximate ones were assessed. From this assessment it is then possible to pick the most suitable approaches to, for instance, treat treat large molecular systems which are beyond the reach of very accurate (and therefore computationally very costly) approaches.

Next we briefly review the frozen density embedding (FDE) method, a formally exact approach that myself and others use as a framework to devise computationally efficient schemes to account for environment effects on the aforementioned electronic structure approaches. Application of these approximate schemes to heavy element systems are discussed in order to show that FDE can be quite accurate describe environment effects (notably in the absence of strong interactions such as covalent bonds between subsystems), thus allowing one to use approaches such as 4c-IHFSCC to obtain electronic spectra or ionization energies.

# Chapter 2

# Sommaire

Cet ouvrage présente des approches théoriques applicables à la modélisation de la structure électronique d'espèces contenant des éléments lourds – c'est à dire, ceux qui se situent au-delà de la cinquième période de la classification périodique – avec un intérêt particulier porté aux lanthanides et actinides à cause de leur importance dans des domaines si variés que les appareils électroniques ou l'énergie nucléaire.

La modélisation de la structure électronique pour tels systèmes requiert la prise en compte de trois ingrédients : (i) des effets relativistes dû aux vitesses très élevées des électrons du cœur, causées par la forte attraction des noyaux lourds ; (ii) des effets de corrélation électronique issus non seulement de l'interaction instantanée entre électrons mais aussi de la quasi-dégénérescences souvent présentes dans des éléments lourds possédant des couches électroniques $d$ et $f$ partiellement remplies ; et (iii) des effets de l'environnement sur les propriétés des molécules contenant les éléments lourds, car celles-ci se trouvent en général en phase condensée.

Nous commençons par une brève révision des approches utilisées pour décrire la corrélation électronique et les effets relativistes avant de nous pencher sur la méthode intermediate Hamiltonian Fock-space coupled cluster à quatre composantes (4c-IHFSCC), de façon à établir sa précision par rapport à des résultats expérimentaux et ensuite montrer comment celle-ci peut être utilisée comme méthode de référence pour l'évaluation d'approches plus approximées. Ces comparaisons nous ont permis ensuite de choisir les approches les plus adéquates pour le traitement de systèmes moléculaires plus étendus, qui seraient impossibles à traiter avec des méthodes plus précises (et par conséquent plus coûteuses).

Ensuite nous présentons la méthode frozen density embedding (FDE), une approche formellement exacte que moi-même et d'autres utilisons comme point de départ pour concevoir des méthodes efficaces du point de vue computationnel afin de décrire les effets de l'environnement dans le cadre des calculs de structure électronique. Nous discutons aussi de l'application de ces méthodes computationnelles à des systèmes contenant des éléments lourds pour montrer qu'elles sont capables de très bien décrire les effets de l'environnement (notamment dans l'absence de interactions fortes telles que des liaisons chimiques entre sous-systèmes), permettant ainsi l'utilisation d'approches telles que 4c-IHFSCC pour obtenir des spectres électroniques ou des énergies d'ionisation.

# Chapter 3

# Preface

This manuscript describes my scientific activities on the calculation of molecular properties for systems containing heavy elements such as actinides, lanthanides and transition metals, carried out during my time as Postdoctoral researcher at the Free University (VU) Amsterdam under the supervision of Lucas Visscher (2005-2006, 2007-2009), at the Université de Strasbourg (2006-2007) under the supervision of Trond Saue, and at the Université de Lille 1 (2009), and following my appointment at CNRS as researcher (chargé de recherche) in 2009.

The main focus of my first year as postdoc in Amsterdam was on exploring how to execute computational chemistry codes in grid environments, but I profited from this period to start exploring with Ivan Infante the use of accurate, wavefunction-based (WFT) electronic structure methods such as Fock-space coupled cluster (FSCC) to obtain the electronic states of actinyl ions. This represented an almost complete break from my graduate training, in which I mostly worked on the development and benchmarking of basis sets for relativistic and non-relativistic electronic structure calculations, though at the time I did manage to put that experience to good use when starting a collaboration with Ken Dyall on developing energy-optimized basis sets for 5d and 4f elements.

In my year in Strasbourg I focused on calculating the energy differences between enantiomers of relatively large complexes caused by parity non conserving interactions employing density functional theory (DFT), which was at the time the only approach based on four-component wavefunctions that are sufficiently accurate and computationally feasible for such systems. There, I also managed to get a bit more involved in the development of the electronic structure code DIRAC, something that proved useful in the following years.

Upon returning for Amsterdam in 2007, I started working together with Christoph Jacob on a simple implementation in DIRAC of frozen density embedding (FDE), which we applied to construct a computationally efficient model to incorporate environment effects on the $f$-$f$ spectrum of neptunyl, obtained with FSCC. At the same time, I got involved in different projects, again using FSCC as well as other more approximate methods to study the $I_3^-$ species, the uranyl (VI) ion and some species isoelectronic to it.

These two facets of my postdoctoral activities, the development of embedding methods and the evaluation of electronic structure approaches, were continued and intensified with my arrival in Lille.

As far as methodological developments are concerned, over the past few years I focused on formulating and implementing DFT-in-DFT and WFT-in-DFT embedding approaches based on response theory in the DIRAC and DALTON codes, and contributing to the development of the PYADF scripting framework and other tools that simplify embedding calculations for complex systems.

While for the most part these efforts have focused on the response to electrical perturbations,

in these past few years I have worked on extending our response formulations to account for properties arising from magnetic perturbations such as magnetizabilities, NMR shieldings and spin-spin coupling constants, which have so far remained largely unexplored with DFT-in-DFT or WFT-in-DFT embedding approaches.

Furthermore, more recently I've been paying particular attention to the development and implementation of coupled cluster approaches in the four-component framework, notably the equation of motion approach (EOM-CC), not only due to the need to improve the capabilities of DIRAC for obtaining accurate excited states and other molecular properties, but also in order to provide me with a platform for pursuing my work on the development of WFT-in-DFT embedding approaches.

From an applications perspective, alongside with Florent Réal and Valérie Vallet, I continued to explore the reliability of Fock-space coupled cluster, to gauge the reliability of more approximated methods such as Time-dependent DFT (TDDFT) and in combination with embedding methods, to try constructing models which can describe heavy species in condensed matter with a sufficient degree of realism.

# Chapter 4

# Background and Motivation

Though most may not necessarily realize it, heavy elements – that is, those at the bottom of the periodic table such as lanthanides, actinides, and heavy ($4d$ or $5d$) transition metals, ($5p$ or $6p$) main block or alkali and alkali earth ($5s$, $6s$) elements – occupy a central position in modern societies: lanthanides, main-block and alkali (earth) elements, for instance, are intensively used as component in objects present in our daily lives such as telephones and other consumer electronics [1, 2], or as part of materials and devices that make up our communication infrastructure [3], but they also play an important role in our health care systems, via medical imaging (as the strong magnets in imaging devices or as contrast agents [4]) or nuclear medicine [5, 6].

Heavy metals are also the backbone of heavy industry: transition metals in particular are key components of homogeneous or heterogeneous catalysts [7, 8] in the chemical industry, while lanthanides and actinides are at the heart of energy production: the first for renewables as components of wind turbines [1], while the second for nuclear energy, first as fuels and later as waste [9–11].

Many of these applications hinge on the presence of partially-filled $d$ or $f$ shells, as that makes the species exhibit strong magnetism, either as single molecules or extended systems. Furthermore, different elements show different degrees to which these partially-filled orbitals are available for bonding, or affected by the environment: for instance, while they are largely unperturbed for lanthanides, making them very interesting for photonics properties exploitable in consumer (TVs, lighting) or industrial (optical fibers) applications, in the actinides these orbitals are often more available, and give attractive possibilities in catalysis that only recently have been recognized.

## 4.1 Heavy elements from the perspective of "goal-driven" research

Among the challenges faced by society due to the use of heavy elements in industry, a particularly serious one is assuring the operational safety of nuclear power plants and of operations connected to the long-term storage or reprocessing of spent nuclear fuels, as well as the readiness in case of major accidents, by continuously assessing risks and mitigating them.

Risk reduction and mitigation can take many forms, and be helped or driven by a combination of fundamental and applied research: in the case of reprocessing, by improving the processes whereby minor actinides and other fission products with shorter half-life are separated via researching more efficient complexating agents [12] and trough a better understanding of the aqueous [13] and non-aqueous chemistries [14, 15] of actinide and lanthanide species. More efficient separation processes have a direct consequence of decreasing the volume of radioactive waste that must go in long-term storage, but by reducing the amount of solvent used, they greatly reduce the risk of fires that could cause the dissemination of volatile compounds (already present or creating during such an accident)

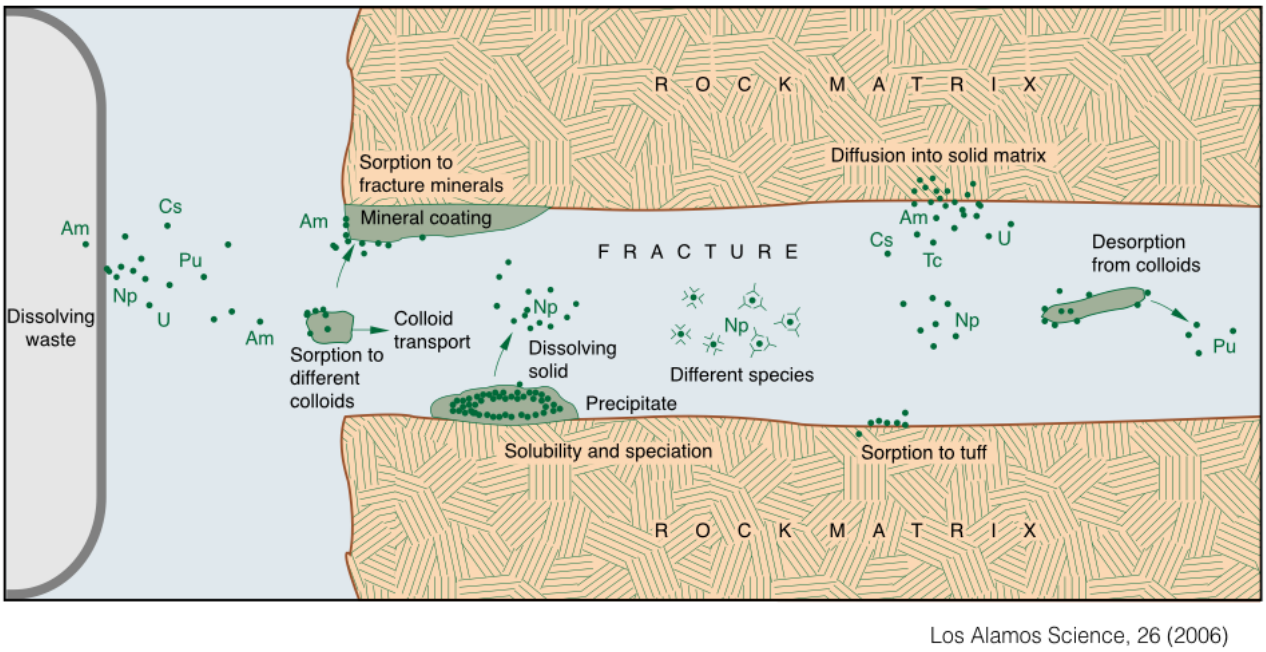



Figure 4.1: Schematic depiction of a breached waste canister and the possible pathways for the transport and retention of radionuclides in the geological environment (from [16]).

of radionuclides.

In the case of long-term storage, it means to be able to assess the risk of environmental dispersion should cannisters holding radioactive waste be breached and their contents escape to the geological environment [16]. The complexities of such a situation can be better grasped by inspecting the schematic representation in figure 4.1. From it we see that a number of concurrent processes may take place where radionuclides may react at the geological interfaces and be trapped, in solution with other radionuclides and form aggregates, or be taken up by colloidal particles and be transported away and be dispersed in the environment. Such an assessment will require, therefore, a deep knowledge of the solubility and the aqueous chemistry of the actinides [17], their interactions and reactivity at mineral surfaces [18], and their interaction with carriers such as colloids [19] with geological repository.

A similar picture can be painted for the case of a nuclear accident such as Tchernobyl or Fukushima, where a number of concurrent phenomena will take place whereby radionuclides may react and be retained at the site or be released to the environment. The difference here is that these processes will take place not in the condensed phase proper but in the gas-phase and over the surface of aerosols. Aerosol particles, shown schematically in figure 4.2, may therefore act as the colloidal particles of figure 4.1, and the radionuclides can also react with other particles adsorbed at the aerosol particle's surface. As simulations of accidents performed for instance with the ASTEC [20] code by the French radioprotection agency, will require reliable thermochemical data in order to establish which species may form non-volatile compounds (and remain on-site) and which form volatile compounds and may be dispersed in the environment.

The data that serves to guide the development of new processes or to perform the simulations, and the knowledge on the electronic structure and of the radionuclide species in different conditions may come from experimental or theoretical studies. Theoretical modeling has gained importance in recent years, as the tools of computational chemistry and physics are able to treat increasingly complex systems (a fact recognized by the 2013 Nobel prize in Chemistry) and effectively provide a virtual laboratory with which one may explore systems in ways not possible experimentally – which is frequently the case for transuranium elements, for example – thus complementing (and sometimes replacing) the latter.

It is difficulty, however, to theoretically model lanthanide and actinide species with sufficiently

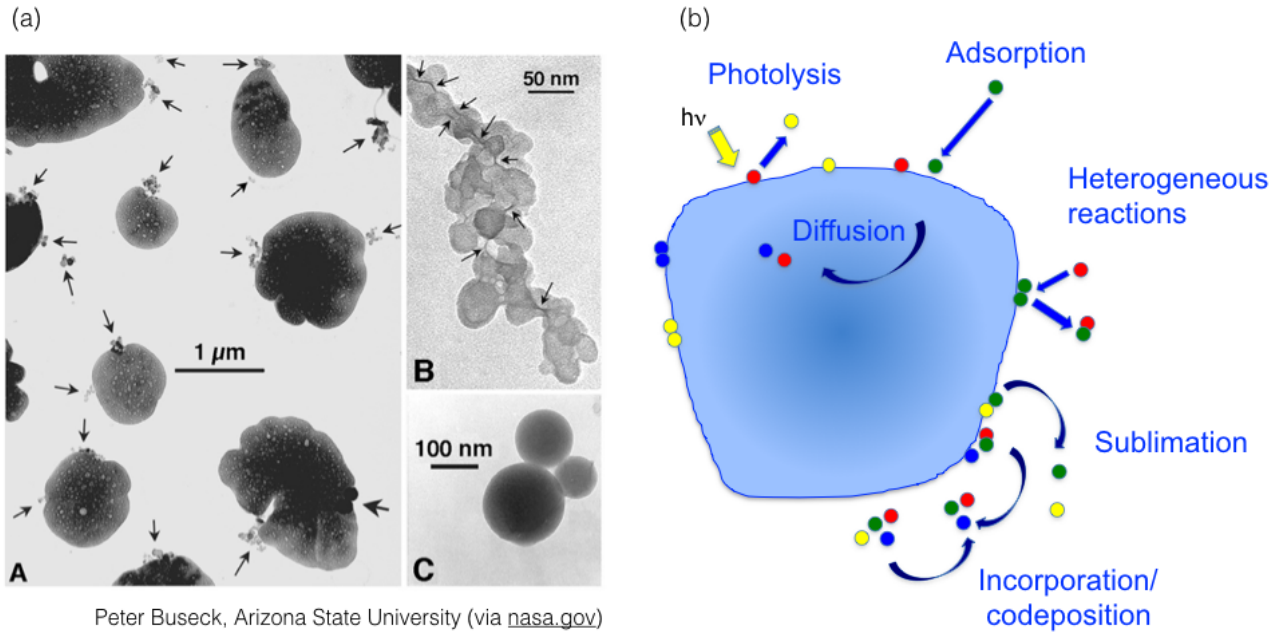



Figure 4.2: (a) Electron microscopy of sulfate particles (image A) to which carbon black particules (image B) are attached (indicated by arrows). Carbon black particles are often associated with fly ash (image B) ; (b) Schematic depiction of the processes that may take place on the surface of an aerosol particle.

high accuracy that makes them proper predictive tools. The first difficulty has to do with the near-degeneracy of the partially-filled $f$ shells in many actinide or lanthanide species, which not only requires a very fine treatment of the instantaneous interaction between electrons (dynamic electron correlation), but also the need to employ electronic structure methods that can describe the quasi-degenerate states (static electron correlation), such as those of wavefunction-based (WFT) approaches.

The second difficulty is that the few approaches that are suitable for dealing with near-degeneracies as above are computationally very costly and cannot in practice treat systems larger than 10-20 atoms, often not enough to construct models with a minimum degree of realism in comparison to the experimental systems - taking into account, for instance, finite temperature effects due to changes in the local structure of the solvent around the species of interest, as well as long-range interactions that may significantly affect molecular properties or spectra. To that end so-called embedding approaches [21–25], in which accurate WFT approaches are combined to more approximate (and cost-effective) ones such as density functional theory (DFT) or force-field based approaches (MM) - complemented by statistical sampling of geometrical configurations for the whole system through the use of molecular dynamics (MD) approaches - are a way to include the interaction between the species containing the heavy element(s) and its environment.

An emblematic case which helps to underscore the difficulties faced by theoretical approaches can be found in the plutonium atom (Pu), one of the key species in reprocessing and environmental issues: in its neutral form in the gas phase, Pu has a over a thousand of experimentally identified electronic states arising from $(7s^2 5f^6, 7s5f6d)$ configurations, of which about 200 are found within a window of 24000 $cm^{-1}$ from the ground state, corresponding to the IR/visible range. When performing correlated calculations (in LS coupling) for the same species one finds about 3000 states arising solely for the $7s^2 5f^6$ configuration, of which 250 are found within the same interval as above (24000 $cm^{-1}$). In solution the element is found either, in the +3 and +4 oxidation states, as atomic or, in the +5 and +6 oxidation states, as dioxo ($PuO_2^{n+}$) ions[1] forming aquo complexes with a varying number of water molecules (≈8-9 for atomic Pu and ≈5-6 for $PuO_2^{n+}$). The arrangement of the aquo ligands

[1] the picture of bonding for these and other actinyl species, which is of great importance to understand the interaction of these species with their environment and with perturbing electromagnetic fields in spectroscopical studies, is nicely summarized in a review by Denning [26]

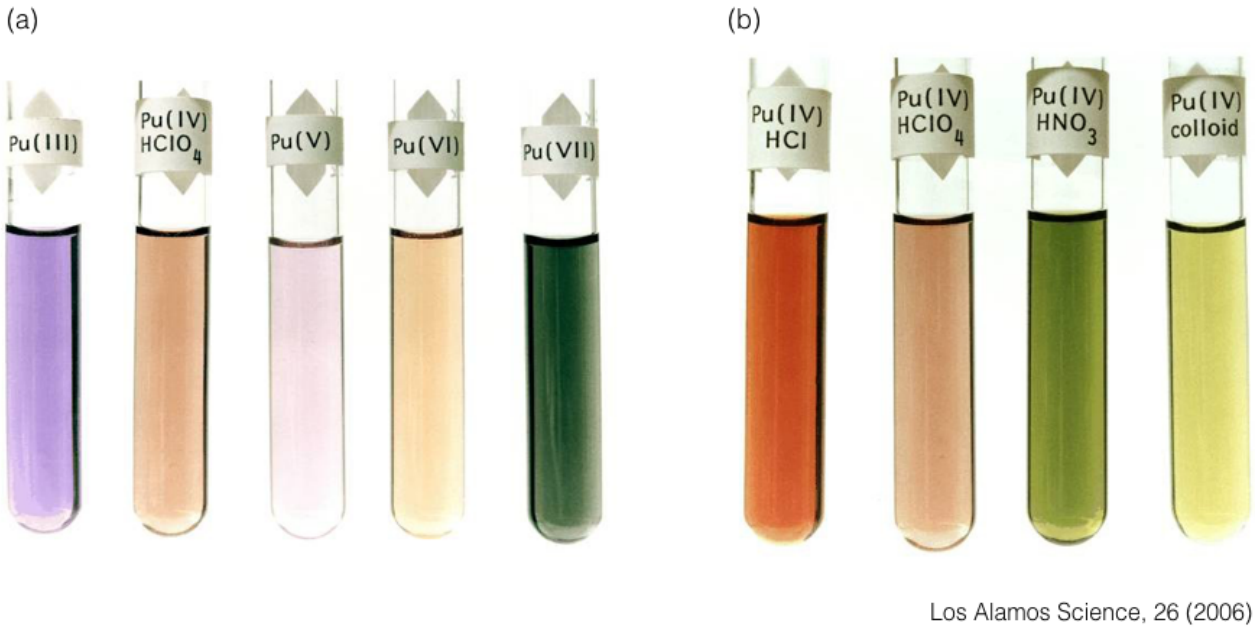



Figure 4.3: (a) Different oxidation states of Pu in perchlorate solution; (b) Pu(IV) complexes in the presence of different ligands (from [27]).

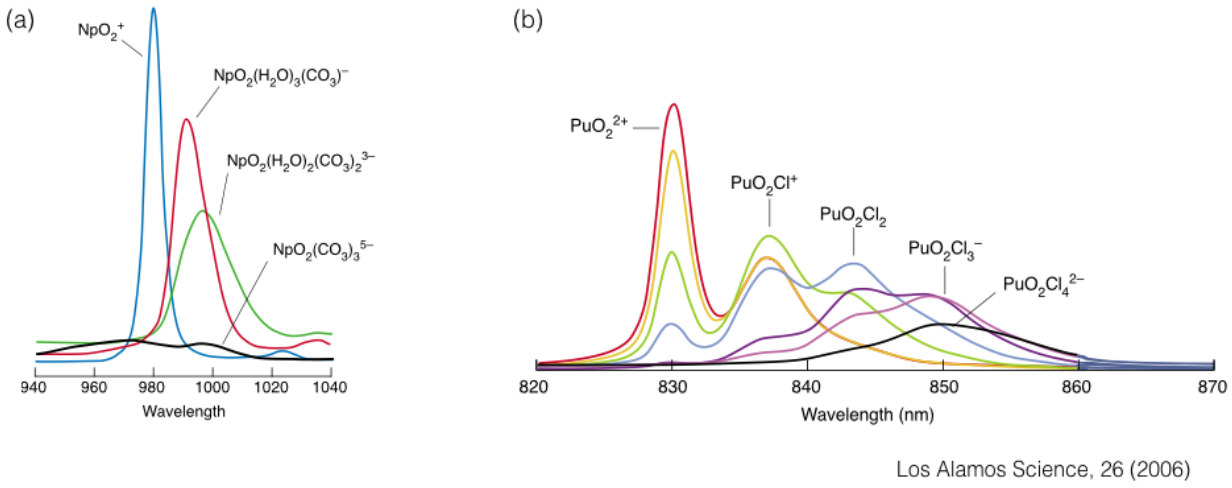



Figure 4.4: (a) UV-vis-NIR spectra for neptunyl species in aqueous solutions; (b) NIR spectra for plutonyl species in NaCl solutions (from [28]).

depends on the temperature and other experimental conditions and with the presence of other ligands such as counter-ions for a total of up to 12 bonds, and species in different oxidation states may occur simultaneously. This great sensitivity of the electronic structure to both oxidation state and the immediate environment of Pu aquo species can be seen, respectively, in figures 4.3 and 4.4. From these it is evident that simple models (gas-phase, static structures), if yielding results in agreement to experiments, will do so by chance rather than by properly capturing the physical processes taking place.

## 4.2 Heavy elements from the perspective of "curiosity-driven" research

Technological or environmental problems involving heavy elements are clear reasons for working toward understanding their properties in the gas phase or in complex environments, though it is clear that along the way experimentalists and theoreticians will focus on problems of a more fundamental nature driven more by curiosity than for a specific application goal in mind: this would be the case for instance in investigations on the interactions between actinide-containing building blocks and the formation of hybrid materials [29] or large clusters [30], and the mechanisms through which these are formed.

Heavy elements are also fascinating in their own right, not only because of the very subtle effects due to electron correlation of the near-degenerate electronic states arising from the partially-filled $f$ shells or of their interaction with the environment, but also due to their electronic structure being fundamentally altered, in comparison to lighter elements, by the increasingly strong manifestation of relativistic effects [31].

Relativistic effect arise due to the strong attraction exerted by the highly charged heavy nuclei on the innermost $(s, p)$ electrons, thereby accelerating them to speeds close to that of the light, with a resulting increase of the mass of the electron: in atomic units, the velocity of a $1s$ electron is roughly equal to the atomic number $Z$ of the atom in question, so that for an atom such as Hg, the electron would have its mass increased by about 20%. Since the mean value of its position is inversely proportional to its mass, the end result is an orbital contraction of the same amount, followed by a contraction of all other $s$ orbitals as a consequence of their orthogonality. The same process occurs for electrons in $p$ orbitals, while for electrons in $d$ and $f$ one will see an expansion as they will screened from the nuclei due to the contraction of $s$ and $p$ orbitals.

Another manifestation of relativity in atoms and molecules is the coupling of spin and angular momentum, referred to as spin-orbit coupling (SOC). Spin-orbit coupling arises from the interaction of the electron spin with the magnetic field induced by the relative movement of other charges, such as the nuclei or other electrons [32], in the framework put forward by Dirac in his attempt to reconcile quantum mechanics and special relativity [33, 34]. A brief introduction to the Dirac equation and its use in molecular calculations will be given in section 5.2. While this manuscript deals with methods and results exclusively within this framework, the reader should be aware that other more sophisticated treatments based on quantum electrodynamics (QED) and for which the Dirac equation is the starting point [35] are an active area of research.

Important consequences of the loss of spin symmetry though SOC are, first, the relaxation of certain selection rules with the practical consequence that processes such as the $f - f$ transitions shown in figures 4.3 and 4.4 are no longer strictly forbidden. Also, as discussed in section 5.3 and in some of the papers in appendices C or D, one will have the loss of degeneracies among certain (non-relativistic) orbitals and by extension the splitting of otherwise degenerate electronic states.

The changes in the electronic structure due to relativity can be seen in the radial density of the astatide ion ($At^-$), whose dynamics in solution have recently been studied by our group and compared to that of lighter ions such as iodide [36]. As other species in the $6p$ block, for this closed-shell species the effects described above (the contraction of $p$ and the loss of degeneracies) can be seen figure 4.5(b), for plots of the density of valence $p$ orbitals along the $z$ axis for iodide and astatide.

We see for the former that the maxima of the $5p_{1/2}$ and $5p_{3/2}$ spinors are very close to each other (2.06 Å and 2.11 Å respectively) and to that of the maximum for the $5p_z$ orbital (2.09 Å), whereas for astatide spatial separation between the $6p_{1/2}$ ($r_{max} = 2.00$ Å) and $6p_{3/2}$ ($r_{max} = 2.32$ Å) maxima is much more important (0.32 Å), with the maximum of the outermost ($6p_{3/2}$) spinor 0.12 Å further from the origin than that of the $6p_z$ orbital.

Relativity may also have important consequences for catalysis [37] and reactivity, as discussed recently by Demissie *et al.* [38] who showed in their investigation of mercury methylation by cobalt corrinoids that taking relativistic effects into account when modeling such an enzymatic reaction is

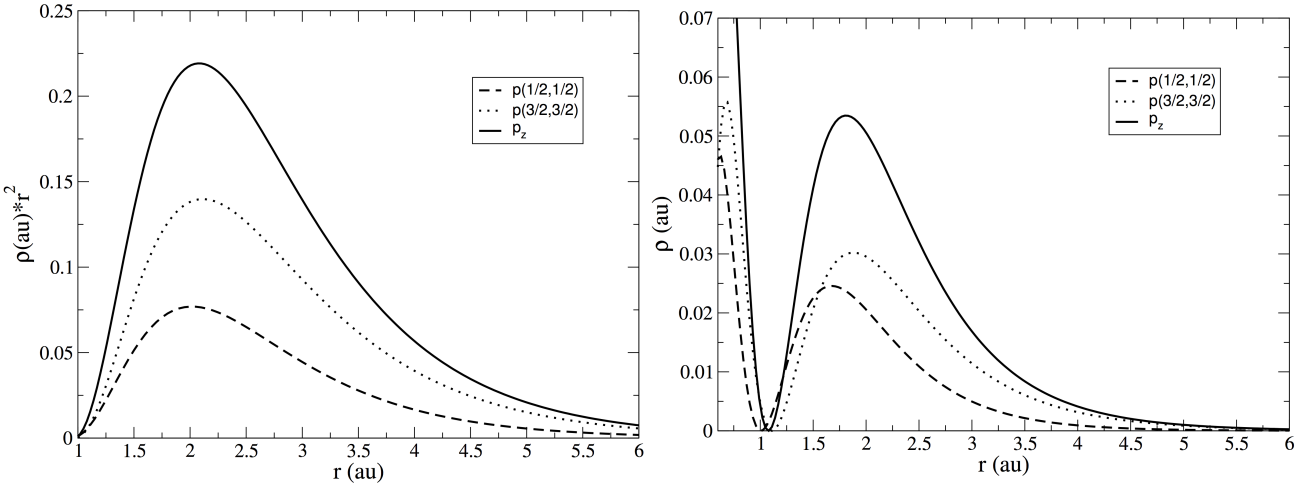



Figure 4.5: Radial densities of the valence $p$ orbital from spin-free calculations and for the $p_{(1/2,1/2)}, p_{(3/2,3/2)}$ spinors of (a) iodide ; (b) astatide.

absolutely essential for getting the mechanism right. There, relativity lowers the energy of the $6p$ spinors in such a way that they can participate in bonding. And the changes in electronic structure due to the relativity have recently been shown to drive the workings of acid-lead batteries [39].

## 4.3 This manuscript

This manuscript is organized as follows: chapter 5 outlines the basic characteristics of the electronic structure approaches used to model the electronic structure heavy elements I used in my research, followed by their application in the study of species containing actinides (Th, U, Np and Pu) and heavy elements from the main group such as halides (iodine, astatine). In chapter 6 the frozen density embedding approach is discussed, as well as its performance on the calculation of electronic spectra of species containing actinides (U, Np) and second-order magnetic properties for model systems containing main-block elements (Se, Te, Po). And finally, the future directions I intend to take in my research are presented in chapter 7.

# Chapter 5

# Electronic Structure Methods

In this chapter I review the approaches used in my activities to determine the many-electron wavefunctions for both ground and excited states, and provide an overview of the performance of these in treating the electronic states of heavy element systems based on my research, based on the papers presented in full in appendices C and D.

Since I, for the most part, have focused on the determination of electronic spectra, the discussion that follows will reflect that and properties such as structures, vibrational spectra etc. will therefore not be explicitly addressed. That said, molecular properties can be formulated as derivatives of the total energy of the system and determined through the application of analytic derivative theory or response theory [40] to the approaches described below.

## 5.1 Electronic structure methods for excited states

*This section partially reproduces the content of section 2 of* ***ASP Gomes, CR Jacob, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem., 2012, 108, 222–277***

For predicting excitation energies, quantum-chemical methods that can treat both the ground-state and excited states are necessary. Here, we will provide an overview of the most important theoretical approaches. These can be divided into two groups that tackle the problem from two formally equivalent, but conceptually very different directions.

First, *time-independent approaches* take the stationary Schrödinger equation,

$$\hat{H}|\Psi_k\rangle = E_k|\Psi_k\rangle, \tag{5.1}$$

where $\hat{H}$ is the molecular Hamiltonian and $|\Psi_k\rangle$ is the many-electron wavefunction, as their starting point. The energy eigenvalues $E_k$ and their corresponding eigenfunctions $|\Psi_k\rangle$ are the energies and wavefunctions of the different electronic states. Thus, in a time-independent picture excitation energies can be calculated by solving the stationary Schrödinger equation (Eqn. (5.1)) for the ground state and the excited states of interest. This requires an explicit construction of the excited-state wavefunctions $|\Psi_k\rangle$.

On the other hand, *time-dependent approaches* start from a stationary ground-state wavefunction $|\Psi_0\rangle$ and consider the time evolution of this initial state after switching on a time-dependent external perturbation. Here, we will only consider an oscillating electric dipole perturbation with frequency $\omega$, i.e.,

$$\hat{V}_\omega(t) = (e^{\mathrm{i}\omega t} + e^{-\mathrm{i}\omega t}) \sum_\beta \epsilon_\beta(\omega)\hat{\mu}_\beta, \tag{5.2}$$

where $\hat{\mu}_\beta$ is the $\beta$-component ($\beta = x, y, z$) of the time-independent electric dipole operator and $\epsilon_\beta(\omega)$ denotes the associated perturbation strength. The time evolution of the wavefunction could be obtained by solving the time-dependent Schrödinger equation

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = \left[\hat{H} + \hat{V}_\omega(t)\right]|\Psi(t)\rangle. \tag{5.3}$$

With the time-dependent wavefunction $|\Psi(t)\rangle$, one can obtain molecular properties by investigating the time evolution of the expectation value of, for instance, a component of the dipole moment $\hat{\mu}_\beta$, expressed as a series expansion,

$$\langle\hat{\mu}_\alpha\rangle(t) = \langle\hat{\mu}_\alpha\rangle + (e^{\mathrm{i}\omega t} + e^{-\mathrm{i}\omega t}) \sum_\beta \epsilon_\beta(\omega)\, \langle\langle\hat{\mu}_\alpha; \hat{\mu}_\beta\rangle\rangle_\omega + \ldots \tag{5.4}$$

Here, $\langle\hat{\mu}_\alpha\rangle$ is the time-independent expectation value and $\langle\langle\hat{\mu}_\alpha; \hat{\mu}_\beta\rangle\rangle_\omega$ is the linear electric dipole–electric dipole response function describing the oscillations of the $\alpha$-component of the dipole moment in response to an oscillating electric dipole field in $\beta$-direction. To first order, only oscillations with frequency $\omega$ appear in this expansion. Terms involving higher-order response functions have been omitted here, and we refer the reader to the thorough derivations of higher-order response theory available in the literature [41].

Within an exact treatment based on time-dependent perturbation theory (see, e.g., ref. [42]) the linear response function is given by the sum-over-states expression

$$\langle\langle\hat{\mu}_\alpha; \hat{\mu}_\beta\rangle\rangle_\omega = \sum_{n\neq 0} \left[ \frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|0\rangle}{\omega_k - \omega_n} + \frac{\langle 0|\hat{\mu}_\beta|n\rangle\langle n|\hat{\mu}_\alpha|0\rangle}{\omega_n - \omega_k} \right] \tag{5.5}$$

where $\omega_n = E_n - E_0$ represents the excitation energy from the ground state to $n$-th excited state. It then becomes evident that the excitation energies occur at frequencies that correspond to the poles of the linear response function. Furthermore, the transition moments for these excitations can be obtained as the corresponding residues of the linear response function.

The sum-over-states expression for the linear response function would require the solution of the time-independent Schrödinger equation for *all* excited states. However, when combined with approximate parameterizations of the many-electron wave-function it becomes possible to determine the linear response function directly by solving a linear system of equations, thus avoiding the explicit calculation of excited state wavefunctions [41, 43]. Such a response theory can, for instance, be obtained by using the quasi-energy formalism, which will be outlined in section 5.1. Subsequently, excitation energies can be determined by identifying the poles of the linear response function.

While for an exact treatment, time-independent approaches (which construct a wavefunction for each excited state explicitly) and time-dependent approaches based on response theory (which avoid the calculation of excited state wavefunctions) are equivalent, this is usually not the case anymore for approximate quantum-chemical methods. Both time-independent and time-dependent approaches are widely used for the quantum chemical calculation of excitation energies, and some of the most important methods will be highlighted in the following.

## Time-independent approaches

### Density-functional theory (DFT).

Instead of solving the stationary Schrödinger equation to determine a many-electron wavefunction $|\Psi_k\rangle$, DFT aims at calculating the corresponding electron density $\rho_k(\boldsymbol{r})$ directly. For the ground-state, the formal justification for replacing the wavefunction $|\Psi_0\rangle$ by the ground-state electron denisty $\rho_0(\boldsymbol{r})$ is given by the Hohenberg–Kohn theorem [44], which establishes the existence of a density functional $E[\rho]$ for calculating the total electronic energy,

$$E[\rho] = F_{\mathrm{HK}}[\rho] + \int \rho(\boldsymbol{r})\, v_{\mathrm{nuc}}(\boldsymbol{r})\, \mathrm{d}^3 r, \tag{5.6}$$

where $F_{\mathrm{HK}}[\rho]$ is a system-independent functional (the so-called universal Hohenberg–Kohn functional) and where $v_{\mathrm{nuc}}(\boldsymbol{r}) = \sum_A \frac{Q_A}{|\boldsymbol{r}-\boldsymbol{R}_A|}$ is the Coulomb potential of the nuclei at positions $\boldsymbol{R}_A$ and with the charges $Q_A$. Furthermore, the Hohenberg–Kohn theorem provides a variational principle for calculating the ground-state energy $E_0$ and the corresponding ground-state electron density $\rho_0$ by minimizing this energy functional, i.e., $E_0 = \min_\rho E[\rho]$.

In Kohn–Sham (KS) DFT [45], the energy functional is decomposed as,

$$E[\rho] = T_s[\rho] + J[\rho] + E_{\mathrm{xc}}[\rho] + \int \rho(\boldsymbol{r})\, v_{\mathrm{nuc}}(\boldsymbol{r})\, \mathrm{d}^3 r, \tag{5.7}$$

where $T_s[\rho]$ is the kinetic-energy of a reference system of noninteracting electrons with density $\rho$, $J[\rho] = \frac{1}{2}\iint \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}^3 r\, \mathrm{d}^3 r'$ is the classical Coulomb interaction of the electron density with itself, and the exchange–correlation functional $E_{\mathrm{xc}}[\rho]$ collects all the remaining energy terms. This energy functional is then minimized by introducing a wavefunction for the reference system of noninteracting electrons, which is given by a single Slater determinant $|\Phi_s\rangle$ built from the orbitals $\{\phi_i\}$. These Kohn–Sham orbitals can then be determined by solving the KS equations,

$$\left[-\frac{\Delta}{2} + v_{\mathrm{nuc}}(\boldsymbol{r}) + v_{\mathrm{Coul}}[\rho](\boldsymbol{r}) + v_{\mathrm{xc}}[\rho](\boldsymbol{r})\right] \phi_i(\boldsymbol{r}) = \epsilon_i\, \phi_i(\boldsymbol{r}), \tag{5.8}$$

where $v_{\mathrm{Coul}}[\rho](\boldsymbol{r}) = \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, \mathrm{d}^3 r$ is the classical Coulomb potential of the electrons and $v_{\mathrm{xc}}[\rho](\boldsymbol{r}) = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\boldsymbol{r})}$ is the exchange–correlation potential. Since these potentials both depend on the electron density, the KS equations have to be solved in self-consistent field (SCF) iterations. Even though KS-DFT provides a formally exact theory for calculating the ground-state density, the exact exchange–correlation functional is not know and approximate functionals have to be introduced. For overviews of the currently available approximate functionals, see, e.g., refs. [46–48], and for a discussion of their limitations, see, e.g., ref. [49, 50].

Even though the Hohenberg–Kohn theorem was initially formulated only for the ground-state, it can easily be extended to the lowest state in each spin or spacial symmetry [51, 52]. However, in this case the Hohenberg–Kohn functional $F_{\mathrm{HK}}[\rho]$ is no longer universal, but becomes symmetry-specific. In practice, this symmetry-specific functional is unknown, and one usually resorts to impose the symmetry constraints on the noninteracting reference system, i.e., the Slater determinant formed from the KS orbitals. Note, however, that in the exact theory neither the spin nor the spatial symmetry of the wavefunction of this reference system correspond to the one of the true wavefunction [53].

Despite the lack of a formal justification [54], excited states have been targeted in variational KS-DFT calculations by employing excited-state wavefunctions for the noninteracting reference system. In the simplest case, this corresponds to replacing an occupied KS orbital by an unoccupied one (i.e., a non-aufbau solution), but it might also be necessary to form linear combinations of different Slater determinants for specific states [55]. In particular, this ΔDFT approach has been applied to study multiplet energies in transition metal complexes with DFT [56]. This approach can be extended to a ΔSCF-DFT scheme, in which the orbitals of the excited Slater determinants constructed for the noninteracting reference system are re-optimized [57]. In this case, it has to be ensured that the SCF procedure does not collapse to the ground-state [58].

For the calculation of excitation energies with DFT, the application of such ΔSCF-DFT calculations has been rather limited. Usually, time-dependent DFT (discussed in section 5.1) provides more accurate results and avoids the rather cumbersome optimization of the KS orbitals of excited Slater determinants. There is a revived interest in the use ΔSCF-DFT and related methods for overcoming some of the limitations of TDDFT, such as for the description of Rydberg states [59] or of large conjugated organic molecules [60]. An interesting related approach is the constricted variational DFT method by Ziegler and coworkers [61, 62], that is based on the determination of stationary points of the energy functional, subject to the condition that the density difference corresponds to an electronic excitation. Both ΔSCF-DFT and TDDFT emerge from this formulation.

**Configuration Interaction-based methods.**

The starting point for most wavefunction based quantum chemical methods is the *Hartree–Fock (HF) approximation*, in which a single Slater determinant $|\Phi_{\mathrm{HF}}\rangle$ is used as ansatz for the ground-state wavefunction. Minimization of the energy expectation value then leads to the HF equations

$$\left[-\frac{\Delta}{2} + v_{\mathrm{nuc}}(\boldsymbol{r}) + v_{\mathrm{Coul}}[\rho](\boldsymbol{r}) + \hat{K}\right] \phi_i(\boldsymbol{r}) = \epsilon_i \, \phi_i(\boldsymbol{r}), \tag{5.9}$$

for determining the orbitals in this determinant. These equations are very similar to the KS equations introduced above, but instead of the exchange–correlation potential the nonlocal exchange operator $\hat{K} = \sum_j \hat{k}[\phi_j]$ with

$$\hat{k}[\phi_j] \, \phi_i(\boldsymbol{r}) = \int \frac{\phi_i(\boldsymbol{r'})\phi_j(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} \mathrm{d}^3 r' \; \phi_j(\boldsymbol{r}) \tag{5.10}$$

appears. One way of expressing the HF determinant, that will be particularly convenient later on, is an exponential parametrization [63],

$$|\Phi_{\mathrm{HF}}\rangle = \exp(\hat{\kappa})|\Phi\rangle, \qquad \hat{\kappa} = \sum_{p>q} [\kappa_{pq} \, \hat{a}_p^\dagger \hat{a}_q - \kappa_{pq}^* \, \hat{a}_q^\dagger \hat{a}_p], \tag{5.11}$$

where $|\Phi\rangle$ is a trial state (usually a single Slater determinant), and $\hat{a}_p^\dagger$ and $\hat{a}_q$ are the operators creating or annihilating an electron in orbital $p$ or $q$, respectively. The advantage of this parametrization is that it introduces only non-redundant parameters by ensuring the orthogonality of the HF orbitals. Therefore, the parameters $\kappa_{pq}$ can be determined from an unconstrained optimization of the HF energy.

Electron correlation can then be included by employing a more general ansatz for the wavefunction that also includes excited Slater determinants, i.e., determinants in which one or more of the occupied HF orbitals are replaced by virtual ones. In *configuration interaction (CI)* methods [64], the many-electron wavefunction $|\Psi_k\rangle$ for a given electron state $k$ is expanded in a basis of Slater determinants as

$$|\Psi_k\rangle = (1 + \hat{C}_k)|\Phi_0\rangle, \qquad |\Phi_0\rangle \equiv |\Phi_{\mathrm{HF}}\rangle \tag{5.12}$$

where the reference state $|\Phi_0\rangle$ is usually chosen as the HF determinant. The operator $\hat{C}_k$ is in turn expressed as

$$\hat{C}_k = \hat{C}_1^{(k)} + \hat{C}_2^{(k)} + \ldots = \sum_{\mu_1} C_{\mu_1}^{(k)} \hat{\tau}_{\mu_1} + \sum_{\mu_2} C_{\mu_2}^{(k)} \hat{\tau}_{\mu_2} + \ldots, \tag{5.13}$$

that is, in terms of the product between coefficients $C_\mu^{(k)}$ and excitations operators $\hat{\tau}_\mu$ of different ranks. Here, $\mu_1$ and $\mu_2$ represent single and double excitations from occupied to virtual orbitals, and the corresponding excitation operators are $\hat{\tau}_{\mu_1} = \hat{a}_a^\dagger \hat{a}_i$ and $\hat{\tau}_{\mu_2} = \hat{a}_a^\dagger \hat{a}_i \hat{a}_b^\dagger \hat{a}_j$, where indices $i, j, k, \ldots$ refer to occupied orbitals whereas indices $a, b, c, \ldots$ label virtual orbitals. Note that the coefficients $C_\mu^{(k)}$ are different for each electronic state $k$.

The coefficients $C_\mu$ can then be determined using the variational principle, i.e., by minimizing the energy with respect to these coefficients. This leads to the Hamiltonian matrix in the basis of the HF and excited determinants,

$$H_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle \tag{5.14}$$

from which the coefficients and energies for the ground and excited states can be obtained as eigenvectors and eigenvalues, respectively. When considering expansions up to the highest possible excitation rank one arrives at the full CI solution, which is exact within a given basis set. However, the size of the corresponding Hamiltonian matrix makes its use impractical for all but the smallest molecular systems. Therefore, one is often constraint to include not more than single and double excitations (CISD). A further truncation already after single excitations results in the CIS method. While it is far from accurate because it lacks a proper inclusion of electron correlation, it nevertheless remains one of the few wavefunction methods that can be used to investigate the spectra of relatively large systems [65–67]. This accuracy can be improved [68, 69] by introducing double excitations in a perturbative fashion, resulting in the CIS(D) model [70].

While in single-reference CI methods the HF determinant is used as reference state $|\Phi_0\rangle$, this is a poor choice in situations where there are other close-lying excited states. This is often the case in open-shell systems such as transition metal complexes. One is then better served by *multi-reference configuration interaction* (MR-CI) methods, which use a reference based on a multi-configurational (MC) wavefunction,

$$|\Phi_0\rangle \equiv |\Phi_{\mathrm{MC}}(\{\mathbf{C}^{'}\})\rangle = \sum_{\nu}(1+\hat{C}^{'}_{\nu})|\Phi_{\mathrm{HF}}\rangle, \tag{5.15}$$

which is a small CI expansion where excitations are allowed only from within a set of orbitals, the so-called active space. This is comprised of a relatively small number of occupied and virtual orbitals, which is indicated above by the use of use of $\hat{C}^{'}_{\nu}$ instead of $\hat{C}_{\nu}$. There are different approaches for generating the determinants in Eqn. (5.15), the most widely used is the complete active space (CAS), where a full CI is performed within the active space.

A further improvement consists in taking the orbital coefficients $\kappa_{pq}$ as additional variational parameters,

$$|\Phi_0\rangle \equiv |\Phi_{\mathrm{MCSCF}}(\{\mathbf{C},\kappa\})\rangle = \sum_{\nu}(1+\hat{C}^{'}_{\nu})|\Phi_{\nu}(\kappa_{\nu})\rangle, \tag{5.16}$$

giving rise to the CASSCF [71] wavefunction. However, the number of configurations and thus the computational effort increases factorially with the size of the active space [63]. Thus, systems that would require large active spaces present a challenge in this respect. This is particularly the case for systems with severe quasi-degeneracies, such as (polynuclear) transition metal complexes or for early actinide compounds [72]. For alternative approaches that scale polynomially with the size of the active space and thus allow for the use of significantly larger active spaces, see refs. [73, 74].

Finally, it is also worth mentioning that — unlike full CI — all truncated CI methods (with the exception of CIS) are not size-extensive. Thus, upon extending the size of the system a spurious decrease in the correlation energy occurs. As a result, excitation energies are not size-intensive, i.e., they can spuriously change as the size of the system is increased (e.g., when studying the effect of the solvation shell on the spectrum of a solute molecule), which may even change the results qualitatively. This can be approximately corrected by using suitable correction schemes [75], but a more rigorous way for achieving size-extensivity of the correlation energy and size-intensivity of excitation energies is provided by coupled-cluster methods, that will be discussed in the following section.

**Single-reference coupled cluster methods.**

In the coupled cluster approach [63, 76], the ground-state many-electron wavefunction is obtained by the so-called exponential parameterization, acting upon the HF determinant

$$|\Psi_0\rangle = \exp(\hat{T})|\Phi_{\mathrm{HF}}\rangle \tag{5.17}$$

where the operator $\hat{T}$ is of the same form as the excitation operator $\hat{C}_k$ in CI methods, but different labels are now used for the expansion coefficients to distinguish these different methods,

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots = \sum_{\mu_1} t_{\mu_1}\hat{\tau}_{\mu_1} + \sum_{\mu_2} t_{\mu_2}\hat{\tau}_{\mu_2} + \ldots. \tag{5.18}$$

As with CI methods, this expansion is usually truncated, for instance after double excitations (resulting in the CCSD approximation). However, the exponential parametrization now generates higher excitation levels even when $\hat{T}$ is truncated, because terms such as $\hat{T}_1^2$, $\hat{T}_1\hat{T}_2$, and $\hat{T}_2^2$ now occur, which correspond to certain double, triple, and quadruple excitations, respectively. By construction, the exponential parametrization of the wavefunction ensures that the resulting energies will be size-extensive.

Since a variational determination of the coupled cluster amplitudes $t_\mu$ is not feasible, projection techniques are used instead. To this end, the Schrödinger equation is left-multiplied by $\langle\mu_i|\exp(-\hat{T})$, resulting in the ground-state energy

$$E_0 = \langle\Psi_{\mathrm{HF}}|\hat{H}\exp(\hat{T})|\Psi_{\mathrm{HF}}\rangle \tag{5.19}$$

and the amplitudes equations

$$\Omega_{\mu_i} = \langle \mu_i | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \Psi_{\mathrm{HF}} \rangle = 0 \tag{5.20}$$

where $\langle \mu_i | = \langle \Phi_{\mathrm{HF}} | \hat{\tau}^{\dagger}_{\mu_i}$, with $\mu_i$ running over all possible excited determinants of excitation rank $i$. The amplitudes will have the general form

$$t_{\mu_1} = \frac{\langle \mu_1 | \hat{A} | \Phi_{\mathrm{HF}} \rangle}{\Delta\epsilon_{\mu_1}} = \frac{\langle {}^{a}_{i} | \hat{A} | \Phi_{\mathrm{HF}} \rangle}{\epsilon_i - \epsilon_a} \tag{5.21}$$

$$t_{\mu_2} = \frac{\langle \mu_2 | \hat{B} | \Phi_{\mathrm{HF}} \rangle}{\Delta\epsilon_{\mu_2}} = \frac{\langle {}^{ab}_{ij} | \hat{B} | \Phi_{\mathrm{HF}} \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \tag{5.22}$$

where $\hat{A}$ and $\hat{B}$ are operators involving the product of $\hat{H}$ and $\hat{T}$ operators. As the expressions for the former depend on the truncation level, we refer to the literature for concrete examples [63, 76].

Because of the projective nature of the coupled cluster method, it is not possible to invoke a variational principle in order to obtain excited states. Therefore, to construct excited state wavefunctions one would need to employ different reference states. This could, for instance, be an open-shell determinant corresponding to a single excitation with respect to $|\Phi_{\mathrm{HF}}\rangle$. This approach has several disadvantages. First, the above coupled cluster equations would have to be solved for each excited state, which can become computationally rather demanding. Second, different electronic states will no longer be orthogonal, which complicates the calculation of transition moments. Finally, excited states that are not dominated by a single determinant cannot be easily described, as they would require a multi-reference coupled cluster method.

An alternative that remedies part of these difficulties is the *equation-of-motion coupled cluster* method [77] for excitation energies (EOM-EE), where one introduces the parametrization

$$|\Psi_k\rangle = \hat{C}_k \exp(\hat{T}) |\Phi_{\mathrm{HF}}\rangle \tag{5.23}$$

$$\langle \bar{\Psi}_l | = \langle \Phi_{\mathrm{HF}} | (1 + \hat{\bar{C}}_l \exp(-\hat{T})). \tag{5.24}$$

Thus, instead of the HF determinant the coupled cluster wavefunction is used as reference state of a CI-like expansion. In this EOM-EE parametrization, the bra and ket states $|\Psi_k\rangle$ and $\langle \bar{\Psi}_l|$ are neither Hermitian conjugate nor orthogonal among themselves, but rather biorthogonal, i.e., $\langle \bar{\Psi}_l | \Psi_k \rangle = \delta_{kl}$. This parametrization allows for a Hamiltonian matrix

$$H_{ij} = \langle \bar{\Psi}_i | \hat{H} | \Psi_j \rangle = \bar{\mathbf{C}}_i^T \mathbf{H} \mathbf{C}_j \tag{5.25}$$

to be formed, which can then be diagonalized to obtain the excitation energies. From the definitions of $|\Psi_k\rangle$ and $\langle \bar{\Psi}_l|$ the matrix $\mathbf{H}$ is obtained as

$$\mathbf{H} = \begin{pmatrix} E_0 & \boldsymbol{\eta}^T \\ \mathbf{0} & \mathbf{A} + E_0 \mathbf{I} \end{pmatrix}, \tag{5.26}$$

where $E_0$ is the ground-state coupled cluster energy, $\eta_\nu$ is the vector

$$\eta_\nu = \langle \Phi_{\mathrm{HF}} | [\hat{H}, \hat{\tau}_\nu] \exp(\hat{T}) | \Phi_{\mathrm{HF}} \rangle, \tag{5.27}$$

and $\mathbf{A}$ is the so-called coupled cluster Jacobian with the elements

$$A_{\mu\nu} = \langle \mu | \exp(-\hat{T}) [\hat{H}, \hat{\tau}_\nu] \exp(\hat{T}) | \Phi_{\mathrm{HF}} \rangle. \tag{5.28}$$

By rewriting the above equations in terms of the energy difference with respect to the coupled-cluster ground-state energy [63, 76], one can identify the eigenvalues of the matrix $\mathbf{A}$ in Eqn. (5.26) with the excitation energies for the system.

These excitation energies are size-intensive [78], thus presenting a significant advantage over CI approaches. The CI-like parametrization for the excited-states makes the calculation of transition

moments and other expectation values for both ground and excited states rather simple, though both left and right eigenvectors need to be calculated because **A** is not Hermitian. However, these transition moments are not size-intensive [79] and therefore in general not reliable if one wishes to study a series of systems with increased size.

Apart from electronically excited states, in the EOM formalism one can also define in a straightforward manner states which correspond to electron attachment (EOM-EA) and detachment (EOM-IP) states, while still keeping the closed-shell coupled cluster reference for the ground state [77, 80]. This is done by modifying the $\hat{\mathbf{C}}_k^T$ and $\hat{\mathbf{C}}_k$ operators used to parametrize the excited states so that instead of single and double replacements, one has a particle annihilation (creation) and single replacement plus particle annihilation (creation) for EOM-IP (EOM-EA),

$$\hat{\mathbf{C}}_k^{\text{EOM-IP}} = 1 + \sum_i c_{k;i}\{a_i\} + \sum_{a,m>i} c_{k;im}^a\{a_a^\dagger a_m a_i\} + \dots \tag{5.29}$$

$$\hat{\mathbf{C}}_k^{\text{EOM-EA}} = 1 + \sum_a c_k^a\{a_a^\dagger\} + \sum_{a>b,i} c_{k;i}^{ab}\{a_a^\dagger a_b^\dagger a_i\} + \dots \tag{5.30}$$

and so on, and by the same way it is possible to define double electron attachment (EOM-DEA) and detachment (EOM-DIP) [81, 82], or higher-ionization variants [83]. An appealing feature of these approaches is that, starting from a closed-shell reference, one can easily obtain spin-adapted eigenfunctions for open-shell states.

A related approach which aims at exploit the EOM framework but go beyond the linear parametrization for the excited states is the similarity-transformed EOM (STEOM) approach [76, 84–86], where a second similarity transformation is applied to the (T-similarity transformed) **H**,

$$\mathbf{G} = \{\exp(\hat{S})\}^{-1}\mathbf{H}\{\exp(\hat{S})\} \tag{5.31}$$

where $\hat{S} = \hat{S}_1 + \hat{S}_2$ is expressed in terms of linear expansions of single and double replacement operators. With this second transformation, by diagonalizing **G** one gets closer to the Fock-space approach described below with the advantage of avoiding the solution of nonlinear equations, at the cost of requiring prior EOM-IP or EOM-EA, depending on the case. The method has also been extended to the determination of excited state gradients [87] and time-dependent properties [88].

**Multi-reference coupled cluster methods.**

The coupled cluster-based methods discussed present significant improvements over CI-based methods, in particular because excitation energies are size-intensive. However, they are based on a single determinant reference for the ground state. As such, they might not be applicable in more complicated cases where already the ground state requires a multi-reference treatment. This limitation can be remedied by employing multi-reference coupled-cluster approaches [89] such as the Hilbert-space (HSCC) and Fock-space (FSCC) coupled cluster or other, more approximate multi-reference methods [90]. In the following we shall focus on the Fock-space approach. because it provides a direct route to excited states and has been applied in a number of molecules containing heavy elements by other authors [91–95] as well as myself (see papers C.1–C.7, D.1, D.4 and D.6).

All multi-reference methods define a reference (or model) space $P$, consisting of a set of determinants $\{|\varphi_i\rangle\}$. Usually, this is achieved by defining an active space, and all determinants contributing to a full CI expansion within this active space are considered. In MRCI methods this is usually the active space obtained from a CASSCF calculation, whereas in multi-reference coupled-cluster methods the HF orbitals are used directly without reoptimization. Of course, it is important that this model space contains all relevant configurations for the states that are of interest [76]. All other possible contributions to the exact wavefunctions that are not contained in the model space $P$ are within its orthogonal complement $Q$.

This model space then serves as the starting point for obtaining a set $\{|\Psi_k\rangle\}$ of exact solutions of the Schrödinger equation with the same dimension as $P$. The correspondence between wavefunctions

in the model space $\{|\varphi_k\rangle\}$ and the exact wavefunctions $\{|\Psi_k\rangle\}$ is then established through a projection operator $\hat{P} = \sum_i |\varphi_i\rangle\langle\varphi_i|$ so that

$$|\varphi_k\rangle = \hat{P}|\Psi_k\rangle. \tag{5.32}$$

In addition, one defines the wave operator $\hat{\Omega}$ that establishes the reverse mapping and generate the exact wavefunction from one belonging to the model space,

$$|\Psi_k\rangle = \hat{\Omega}|\varphi_k\rangle. \tag{5.33}$$

Using these definitions, one can now change from the problem of solving the exact Schrödinger equation, $\hat{H}|\Psi_m\rangle = E_m|\Psi_m\rangle$, to that of solving a Schrödinger equation for the model space

$$\hat{H}_{\text{eff}}|\varphi_m\rangle = \hat{\Omega}^{-1}\hat{H}\hat{\Omega}|\varphi_m\rangle = E_m|\varphi_m\rangle, \tag{5.34}$$

where the exact Hamiltonian $\hat{H}$ is now replaced by an effective Hamiltonian $\hat{H}_{\text{eff}}$ which has the same eigenvalues. The wave operator $\hat{\Omega}$ and, consequently, the effective Hamiltonian $\hat{H}_{\text{eff}}$ can be defined via the solution of the so-called Bloch equation. For details, we refer to ref. [96].

Once $\hat{\Omega}$ is obtained, $\hat{H}_{\text{eff}}$ in Eq. 5.34 can be constructed and its matrix representation can be diagonalized to obtain the energies $\{E_m\}$. In multi-reference coupled cluster, an exponential parametrization is used as ansatz for the wave operator $\hat{\Omega}$,

$$\hat{\Omega} = \exp(\hat{S}) \tag{5.35}$$

where the operator $\hat{S}$ contains operators exciting electrons from the $P$ to the $Q$ spaces, multiplied with the corresponding amplitudes. The form of $\hat{S}$ will depend on the details of the approach. In the case of FSCC, it is constructed by considering operators defined for different sectors of Fock-space, where each sector comprises a model space which differs from the reference situation by the addition or removal of electrons, starting from the original closed-shell determinant $|\Phi_0\rangle$. Each of these sectors is identified by the tuple $(n, m)$, where $n$ represents the number of holes and $m$ the number of electrons created on $|\Phi_0\rangle$. Thus, $\hat{S}$ is expressed as

$$\hat{S} \;=\; \hat{S}^{(0,0)} + \hat{S}^{(1,0)} + \hat{S}^{(0,1)} + \hat{S}^{(1,1)} + \ldots \tag{5.36}$$

and similar decompositions are used for the projection operators $\hat{P}$ and $\hat{Q}$. Using this decomposition, the amplitudes can be determined in a hierarchical fashion for the different sectors. For instance, in order to obtain the contributions from the $(a, b)$ sector to $\hat{S}$, one must first solve all $(m, n)$ sectors for which $m < a, n < b$ [96, 97]. For a given sector, one obtains amplitude equations of the form

$$\langle\chi_j^{(m,n)}|\hat{H}\hat{\Omega} - \hat{\Omega}\hat{H}_{\text{eff}}^{(m,n)}|\varphi_i^{(m,n)}\rangle = 0 \tag{5.37}$$

where $\{\chi_j^{(m,n)}\}$ are the determinants belonging to the complement space for the $(m, n)$ sector. For the (0,0) sector this is equivalent to single-reference amplitude equations in Eqn. (5.20).

However, in their original formulation FSCC methods are plagued by the so-called intruder state problem [98]. This occurs during the iterative solution of Eqn. (5.37). When certain low-lying states belonging to $Q^{(m,n)}$ turn out to have energies close to (or even lower than) the higher-lying states from $P^{(m,n)}$ in some amplitudes equations (which are similar to those in Eqns. (5.21) and (5.22)) might have very small energy denominators and prevent the solution of the linear system. The larger the $P^{(m,n)}$ space, the more serious this problem becomes, as accidental degeneracies with $Q$ states become increasingly likely. To alleviate this difficulty, approaches based on an intermediate Hamiltonian formulation (IHFSCC) have been introduced [99–103]. These divide the $P$ space as $P = P_m + P_i$ and $P_m$ now serves as the basis for projecting the lower exact solutions, whereas for $P_i$ this requirement is relaxed. This provides enough flexibility so that when $Q$ states interact strongly with the higher-lying $P_i$ states, the corresponding amplitude equations can be approximated as to assure convergence.

It should be noted that the equations for the $(0, 1)$ and $(1, 0)$ sectors are formally equivalent to the EOM-EA and EOM-IP approaches [76], though for the latter one avoids their iterative solution

by diagonalizing (T-)similarity transformed Hamiltonian in the space of $(0,1)+(1,2)$ or $(1,0)+(2,1)$ determinants whereas for Fock-space the diagonalization is only over the determinant space corresponding to each of the sectors, which is evidently much smaller. For the $(1,1)$ sector and beyond, however, the methods are not equivalent anymore, as it can be seen by starting from the definitions in Eqn. (5.35). There, we can write the wavefunction for state $k$ as

$$|\Psi_k(\mathrm{FS})\rangle = \hat{\Omega}^{\mathrm{FS}}|\Phi_0\rangle = \exp(\hat{S})|\Phi_0\rangle = \exp(\hat{S}-\hat{T})\exp(\hat{T})|\Phi_0\rangle. \tag{5.38}$$

Comparing Eqn. (5.38) and Eqn. (5.24), one sees that unlike EOM-EE, Fock-space coupled cluster departs from the linear parametrization for the excited states. The exponential parametrization ensures that only connected terms are taken into account, making the excited state energies different in both cases. As mentioned above, the STEOM approach was introduced to recover most of the characteristics of employing a true exponential parametrization for the excited states, though there is still no implementation of STEOM-CC that would allow us to assess its performance with respect to IHFSCC for heavy elements.

**Multireference Perturbation Theory-based methods.**

Given that coupled cluster methods can be computationally expensive, in particular for higher excitation levels such as CCSDT, other approaches have been devised that attempt to combine a reasonable accuracy in the description of electron correlation and flexibility to obtain excited states even when the reference wavefunction present a significant multi-configurational character. The perhaps most popular of such approaches is the CASPT2 method of Roos and coworkers [104, 105] where one combines a CASSCF wavefunction with second-order perturbation theory.

Being based on perturbation theory, one now starts from a partitioning of the Hamiltonian into a zeroth-order contribution $\hat{H}_0$ and a perturbing part $\hat{V}$,

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{5.39}$$

where $\hat{H}_0$ is a Fock-like operator that has the CASSCF wavefunctions $|\varphi_m\rangle$ as eigenvectors, i.e.,

$$\hat{H}_0|\varphi_m\rangle = E_{0,m}|\varphi_m\rangle. \tag{5.40}$$

As in the multi-reference coupled cluster methods above, in the multistate CASPT2 method [105] one partitions Hilbert space into different subspaces. First, the model space $P_0$ contains a subset $\{|\varphi_i\rangle\}$ of the CASSCF wavefunctions as reference states. Second, the subspace $P'$ orthogonal to $P_0$ contains the remaining $\{|\varphi_j\rangle\}$ CASSCF wavefunctions. Finally, the orthogonal complement $Q$ is made up of all other determinants not contained within the CASSCF active space. Furthermore, one defines a wave operator $\hat{\Omega}^{\mathrm{PT2}}$ that provide the mapping between the model space and the exact solutions. Instead of the exponential parametrization in coupled cluster methods, $\hat{\Omega}^{\mathrm{PT2}}$ is now expressed as a linear expansion in the orders of perturbation, and only the terms up to first order are retained,

$$\hat{\Omega}^{\mathrm{PT2}} = 1 + \hat{\Omega}^{(1)}. \tag{5.41}$$

In multistate CASPT2 [106], $\hat{\Omega}^{\mathrm{PT2}}$ can be expressed as a linear combination of state-specific [106, 107] wave operators,

$$\hat{\Omega}^{(1)} = \sum_i \hat{\Omega}_i^{(1)}|\varphi_i\rangle\langle\varphi_i|, \tag{5.42}$$

where the index runs over all the functions in the model space $P_0$.

Using such a formalism, one arrives at the equations for the individual states

$$\hat{\Omega}^{(1)}|\varphi_i\rangle = \hat{R}_i\hat{V}|\varphi_i\rangle = \sum_k \frac{|k\rangle\langle k|\hat{V}|\varphi_i\rangle}{(E_0^i - F^k)}, \tag{5.43}$$

where $k$ runs over the states belonging to the $Q$ subspace and where the actual form of the denominator involves orbital energy differences, whereas $F^k$ is a generalized Fock operator.

The main weakness of CASPT2 is its susceptibility to intruder states, that is, states that belong to $Q$ and are thus outside of active spaces, but that have energies close to $E_0^i$. These will makes the expression for $\hat{R}_i$ go to infinity causing the perturbation expansion to diverge. Instead of adopting the intermediate Hamiltonian technique as done in the coupled cluster case, Roos and coworkers [108–110] as well as others [111] devised practical solutions based on the modification of such denominators by the application of a global level shift parameter that modifies the Hamiltonian in such a way that any effects due to the quasidegeneracy of $E_0^i$ and $F^k$ can be avoided. This approach has been found to work very well in practice for the so-called "weak" intruder states that do not interact too strongly with the reference. For more strongly interacting states other solutions must be sought (e.g., including them into the active space), otherwise one should employ alternative formulations such as the NEVPT2 approach [112] that avoid intruder states by construction.

In spite of that, CASPT2 remains one of the few methods (along with MRCI approaches discussed earlier) that can accurately describe both static and dynamic correlation effects for large molecular complexes containing centers with a number of half-filled $d$ or $f$ shells such as those containing (bi)metallic centers [113–115]. It is also routinely employed to study the spectra and photochemistry of other complexes of transition metals and heavy elements [93, 116–120].

### Time-dependent approaches based on response theory

Instead of explicitly calculating the wavefunctions of the different excited states, approaches based on response theory set out to calculate the linear response function $\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega$ (see Eqn. (5.5)). A common theoretical framework for achieving this with different methods of quantum chemistry is provided by the quasienergy formalism pioneered by Christiansen and coworkers [41, 121]. It defines a quasi-energy as the time-dependent generalization of the energy,

$$Q(t) = \left\langle \tilde{\Psi}_0(t) \left| \hat{H} + \hat{V}_\omega(t) - i\frac{d}{dt} \right| \tilde{\Psi}_0(t) \right\rangle \tag{5.44}$$

with

$$|\tilde{\Psi}_0(t)\rangle = e^{-\mathrm{i}E_0 t}|\Psi_0(t)\rangle, \tag{5.45}$$

as well as its time average over one period of the perturbation $T = 2\pi/\omega$

$$\{Q(t)\}_T = \frac{1}{T}\int_{-T/2}^{T/2} Q(t)\, dt. \tag{5.46}$$

With these definitions, the (time-dependent) wavefunction $|\tilde{\Psi}_0(t)\rangle$ can be determined by making the quasi-energy stationary with respect to variations in the wavefunction,

$$\delta\{Q(t)\}_T = 0. \tag{5.47}$$

Response functions can then be determined as derivatives of the time-averaged quasi-energy with respect to the perturbation strengths. In particular, the electric dipole–electric dipole linear response function is given by the second derivative of the time-averaged quasi-energy as

$$\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega \quad = \quad \frac{d^2\{Q(t)\}_T}{d\epsilon_\alpha(\omega) d\epsilon_\beta(\omega)}. \tag{5.48}$$

Note that this is analogous to the time-independent case, where the wavefunction is determined by making the expectation value of the energy stationary, and where the corresponding static properties (e.g., the polarizability) can be determined as derivatives of the energy. The quasi-energy formalism allows one to employ the same techniques also for frequency-dependent problems.

In any quantum chemical method, a parametrization of the wavefunction is introduced, i.e., the wavefunction depends on a set of parameters $\boldsymbol{\lambda}$. In so-called variational methods these parameters are in the time-independent case determined by minimizing the energy expectation value. In the time-dependent case, the corresponding parameters can be determined by minimizing the time-averaged

quasi-energy $\{Q(t,\boldsymbol{\lambda})\}_T$. However, a number of important quantum-chemical methods, such as coupled cluster theory, are not variational, i.e., the wavefunction parameters $\boldsymbol{\lambda}$ are determined from other conditions. In this case, one can replace $Q(t,\boldsymbol{\lambda})$ by the Lagrangian

$$L(t,\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}}) = Q(t,\boldsymbol{\lambda}) + \bar{\boldsymbol{\lambda}}\,\boldsymbol{g}(\boldsymbol{\lambda}). \tag{5.49}$$

The new set of parameters $\bar{\boldsymbol{\lambda}}$ are the Lagrange multipliers and $\boldsymbol{g}(\boldsymbol{\lambda}) = 0$ is a set of auxiliary time-dependent equations. Both the parameters $\boldsymbol{\lambda}$ and the Lagrange multipliers $\bar{\boldsymbol{\lambda}}$ can then be treated as variational parameters.

The parameters and possibly also the multipliers can then be determined using variational perturbation theory by expanding the quasi-energy Lagrangian in orders of the perturbation strengths and taking the time average,

$$\{L(t,\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\}_T = L^{(0)} + \sum_\alpha \epsilon_\alpha(\omega)\, L^{(1)}_\alpha(\omega,) + \sum_{\alpha\beta} \epsilon_\alpha(\omega)\epsilon_\beta(\omega)\, L^{(2)}_{\alpha\beta}(\omega) + \ldots, \tag{5.50}$$

where the linear response function can be identified as

$$\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega = \frac{\mathrm{d}^2\{L(t,\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\}_T}{\mathrm{d}\epsilon_\alpha(\omega)\mathrm{d}\epsilon_\beta(\omega)} = L^{(2)}_{\alpha\beta}(\omega). \tag{5.51}$$

Similarly, the wavefunction parameters $\lambda$ and multipliers $\bar{\lambda}$ can be expanded as

$$\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}^{(0)} + (e^{\mathrm{i}\omega t} + e^{-\mathrm{i}\omega t}) \sum_\alpha \epsilon_\alpha(\omega)\, \boldsymbol{\lambda}^{(1)}_\alpha(\omega) + \ldots \tag{5.52}$$

$$\bar{\boldsymbol{\lambda}}(t) = \bar{\boldsymbol{\lambda}}^{(0)} + (e^{\mathrm{i}\omega t} + e^{-\mathrm{i}\omega t}) \sum_\alpha \epsilon_\alpha(\omega)\, \bar{\boldsymbol{\lambda}}^{(1)}_\alpha(\omega) + \ldots, \tag{5.53}$$

where we will use $\boldsymbol{\lambda}^{(1)}(\omega)$ and $\bar{\boldsymbol{\lambda}}^{(1)}(\omega)$ to refer to the first-order terms in these expansions. These can then be determined from the variational conditions

$$\frac{\mathrm{d}\{L(t,\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\}_T}{\mathrm{d}\boldsymbol{\lambda}} = \frac{\mathrm{d}\{L(t,\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\}_T}{\mathrm{d}\bar{\boldsymbol{\lambda}}} = 0 \tag{5.54}$$

at each perturbation order.

Taking the second derivative of $\{L(t,\boldsymbol{\lambda},\bar{\boldsymbol{\lambda}})\}_T$ with respect to the perturbation strengths $\epsilon_\alpha(\omega)$ and $\epsilon_\beta(\omega)$ while taking the implicit dependence of $\boldsymbol{\lambda}$ and $\bar{\boldsymbol{\lambda}}$ on the perturbation strengths into account via the chain rule leads to [41]

$$\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega = L^{(2)}_{\alpha\beta}(\omega) = \begin{pmatrix} \boldsymbol{\lambda}_\alpha \\ \bar{\boldsymbol{\lambda}}_\alpha \end{pmatrix}^T \begin{pmatrix} \boldsymbol{F} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{J} \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}_\beta \\ \bar{\boldsymbol{\lambda}}_\beta \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_\alpha \\ \boldsymbol{\xi}_\alpha \end{pmatrix}^T \begin{pmatrix} \boldsymbol{\lambda}_\beta \\ \bar{\boldsymbol{\lambda}}_\beta \end{pmatrix} + \begin{pmatrix} \boldsymbol{\lambda}_\alpha \\ \bar{\boldsymbol{\lambda}}_\alpha \end{pmatrix}^T \begin{pmatrix} \boldsymbol{\eta}_\beta \\ \boldsymbol{\xi}_\beta \end{pmatrix} \tag{5.55}$$

where we introduced $\boldsymbol{\lambda}_{\alpha/\beta} = \boldsymbol{\lambda}^{(1)}_{\alpha/\beta}$ and $\bar{\boldsymbol{\lambda}}_{\alpha/\beta} = \bar{\boldsymbol{\lambda}}^{(1)}_{\alpha/\beta}$ to simplify the notation and define the abbreviations for the partial derivatives with respect to the parameters and multipliers

$$\mathbf{F} = \frac{\partial^2\{L^{(2)}\}_T}{\partial\boldsymbol{\lambda}^{(1)}(\omega)\partial\boldsymbol{\lambda}^{(1)}(\omega)} \quad \mathbf{A} = \frac{\partial^2\{L^{(2)}\}_T}{\partial\boldsymbol{\lambda}^{(1)}(\omega)\partial\bar{\boldsymbol{\lambda}}^{(1)}(\omega)} \quad \mathbf{J} = \frac{\partial^2\{L^{(2)}\}_T}{\partial\bar{\boldsymbol{\lambda}}^{(1)}(\omega)\partial\bar{\boldsymbol{\lambda}}^{(1)}(\omega)} \tag{5.56}$$

and for the mixed partial derivatives with respect to parameters or multipliers and perturbation strength

$$\boldsymbol{\eta}_{\alpha/\beta} = \frac{\partial\{L^{(2)}\}_T}{\partial\boldsymbol{\lambda}^{(1)}(\omega)\partial\epsilon_{\alpha/\beta}(\omega)} \qquad \boldsymbol{\xi}_{\alpha/\beta} = \frac{\partial\{L^{(2)}\}_T}{\partial\bar{\boldsymbol{\lambda}}^{(1)}(\omega)\partial\epsilon_{\alpha/\beta}(\omega)}. \tag{5.57}$$

Note that all these partial derivatives depend on the perturbation frequency $\omega$.

The first-order parameters $\boldsymbol{\lambda}_{\alpha/\beta}$ and multipliers $\bar{\boldsymbol{\lambda}}_{\alpha/\beta}$ can be determined from the variational conditions of Eqn. (5.54). They only appear in the second-order term $L^{(2)}_{\alpha\beta}(\omega)$ and setting the derivative of Eqn. (5.55) with respect to $\boldsymbol{\lambda}_\alpha$ and $\bar{\boldsymbol{\lambda}}_\alpha$ to zero leads to the linear equations

$$\begin{pmatrix} \boldsymbol{\lambda}_\beta \\ \bar{\boldsymbol{\lambda}}_\beta \end{pmatrix} = - \begin{pmatrix} \boldsymbol{F} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{J} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\eta}_\beta \\ \boldsymbol{\xi}_\beta \end{pmatrix}. \tag{5.58}$$

Substituting this back into Eqn. (5.55) yields the linear response function

$$\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega = - \begin{pmatrix} \boldsymbol{\eta}_\beta \\ \boldsymbol{\xi}_\beta \end{pmatrix}^T \begin{pmatrix} \boldsymbol{F} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{J} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\eta}_\alpha \\ \boldsymbol{\xi}_\alpha \end{pmatrix}, \tag{5.59}$$

which has poles for frequencies that lead to zero eigenvalues of the matrix whose inverse appears in the above equation. This provides a route to the calculation of excitation energies within response theory, whereas the corresponding oscillator strengths can be calculated as the residues corresponding to these poles of the linear response function.

**Time-dependent DFT (TDDFT).**

The theoretical foundation for the time-dependent generalization of DFT was introduced by Runge and Gross [122], and the subsequently introduced TDDFT response theory [123] forms the basis for its application for the calculation of excited states. In TDDFT, the time-dependent density is obtained from the KS wavefunction of a reference system of noninteracting electrons, i.e., a single Slater determinant $|\tilde{\Phi}_s(t)\rangle$, where the KS orbitals $\tilde{\phi}(\boldsymbol{r}, t)$ now become time-dependent. This gives rise to the time-dependent density $\rho(\boldsymbol{r}, t) = \sum_i |\tilde{\phi}(\boldsymbol{r}, t)|^2$. Within the adiabatic approximation any explicit time or history dependence of the exchange-correlation contribution is neglected, and one arrives at the time-averaged quasi-energy functional [124, 125]

$$\{Q[\rho](t)\}_T = \{E[\rho](t)\}_T + \{\langle\tilde{\Phi}_s|\hat{V}(t)|\tilde{\Phi}_s\rangle\}_T - \{\langle\tilde{\Phi}_s|\mathrm{i}\tfrac{\partial}{\partial t}|\tilde{\Phi}_s\rangle\}_T, \tag{5.60}$$

where $E[\rho]$ is the energy functional of Eqn. (5.7) evaluated for the time-dependent density $\rho(\boldsymbol{r}, t)$, the second term accounts for the time-dependent perturbation (Eqn. (5.2)), and the last term arises from the time derivative in the definition of the quasi-energy (Eqn. (5.44)).

In order to obtain expressions for the response equations and the linear response function one introduces a parametrization of the time-dependent KS determinant $|\tilde{\Phi}_s\rangle$. This can be achieved by using an exponential parametrization acting upon the ground-state Kohn-Sham determinant $|\Phi_s\rangle$ (cf. Eqn. (5.11)

$$|\tilde{\Phi}_s\rangle = \exp(\hat{\kappa}(t))\,|\Phi_s\rangle, \quad \hat{\kappa}(t) = \sum_{p>q} [\kappa_{pq}(t)\hat{a}^\dagger_p\hat{a}_q - \kappa^*_{pq}(t)\hat{a}^\dagger_q\hat{a}_p], \tag{5.61}$$

i.e., the time dependence is contained in the operator $\hat{\kappa}(t)$. This parametrization automatically ensures the orthogonality of the time-dependent KS orbitals without introducing additional Lagrange multipliers. Using this parametrization, the linear response function can be determined by differentiating the time-averaged quasi-energy with respect to the parameters $\kappa_{pq}(t)$.

Because these parameters can be determined variationally, no additional Lagrange multipliers are needed and the linear response function is given by

$$\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega = -\boldsymbol{\eta}_\beta \mathbf{F}(\omega)^{-1} \boldsymbol{\eta}_\alpha \tag{5.62}$$

where for the $\mathbf{F}$ matrix one obtains

$$\mathbf{F}(\omega) = \frac{\partial^2 \{Q^{(2)}\}_T}{\partial \boldsymbol{\kappa}^{(1)} \partial \boldsymbol{\kappa}^{(1)}} = \boldsymbol{E}^{[2]} - \omega \boldsymbol{S}^{[2]} \tag{5.63}$$

where $\mathbf{S}^{[2]}$ is a diagonal matrix arising from the last term in Eqn. (5.60) and $\mathbf{E}^{[2]}$ is the electronic Hessian which has the form

$$\mathbf{E}^{[2]} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix}. \tag{5.64}$$

The subblocks of this matrix are given by

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + 2(ia|bj) + (ia|f_{\mathrm{xc}}|bj), \tag{5.65}$$

$$B_{ia,jb} = 2(ia|jb) + (ia|f_{\mathrm{xc}}|jb), \tag{5.66}$$

where $\varepsilon_p$ are the KS orbital energies, $(pq|rs)$ are two-electron integrals in the Mulliken (charge cloud) notation and

$$(ia|f_{\mathrm{xc}}|jb) = \iint \phi_i(\boldsymbol{r})\phi_a(\boldsymbol{r}) \left[\frac{\delta^2 E_{\mathrm{xc}}[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right] \phi_j(\boldsymbol{r}')\phi_b(\boldsymbol{r}')\ \mathrm{d}^3r\ \mathrm{d}^3r' \tag{5.67}$$

are the integrals over the exchange-correlation kernel. Thus, excitation energies can be calculated by determining the eigenvalues of $\mathbf{E}^{[2]}$. Since the size of this matrix is generally rather large, iterative diagonalization procedures which only require the calculation of matrix–vector products are usually employed (see e.g. the discussion in ref. [126]).

Furthermore, there are a number of approximations to Eqn. (5.64) to reduce the overall computational cost, in addition to approximations for the calculation of the two-electron integrals or those over the exchange-correlation kernel described before (density fitting etc). The first is the so-called Tamm-Dancoff approximation, in which the $\mathbf{B}$ subblock of $\mathbf{E}^{[2]}$ is disregarded. This has as effect the reduction of the eigenvalue problem to

$$(\mathbf{A} - \omega\mathbf{I})\mathbf{X} = 0 \tag{5.68}$$

One can go further and invoke the "single-pole approximation" (SPA), where the off-diagonal elements of $\mathbf{A}$ are disregarded. This is in principle a much more severe approximation since the flexibility to describe the excited states in a CI-like expansion is lost and one moves towards the ΔSCF-DFT picture, albeit with the advantage with respect to the latter in that all states are orthogonal. In practice, however, one notes that SPA can still yield a semiquantitatively correct picture, since one often sees (for certain molecules) states that are dominated by one singly excited determinant.

The expressions above apply to the case of pure exchange–correlation functionals (i.e., not containing Hartree–Fock exchange). With hybrid functionals, the subblocks of $\mathbf{E}^{[2]}$ become

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a^{KS} - \varepsilon_i^{KS}) + [ia||bj]^{\gamma} + (1-\gamma)(ia|f_{\mathrm{xc}}|bj)\,, \tag{5.69}$$

$$B_{ia,jb} = [ia||jb]^{\gamma} + (1-\gamma)(ia|f_{\mathrm{xc}}|jb)\,. \tag{5.70}$$

where

$$[pq||rs]^{\gamma} = 2(pq|rs) - \gamma(ps|rq)\,, \tag{5.71}$$

and where $\gamma$ is the fraction of Hartree-Fock exchange.

TDDFT has been widely used for studying excited states in medium to large molecules [127–129]. Its main advantage is its excellent cost–accuracy ratio. For a broad range of chemical applications, TDDFT is capable of providing results which are qualitatively correct, and often come close to the accuracy of more sophisticated wavefunction-based methods such as CASPT2. This has been demonstrated in a number of recent benchmarking studies on light [130–136] and heavy-element containing [137] molecules. However, the accuracy often depends on the proper choice of the exchange–correlation functional and kernel, and choosing these appropriately for a particular application is based on experience and often requires careful comparison with more accurate wavefunction based calculations.

However, with the currently available approximations, TDDFT also has some severe shortcomings [138]. First, with standard non-hybrid exchange–correlation functionals, it does not provide a correct description of Rydberg states. This is caused by the wrong asymptotic form of the exchange–correlation potential [139]. This problem can be addressed by constructing approximations for the exchange–correlation potential that enforce the correct asymptotic behavior, for instance by using orbital-dependent model potentials [140]. Alternatively, range-separated hybrid functionals can be used that also result in asymptotically correct exchange–correlation potentials [50]. Second, charge-transfer excitations are not described correctly. For detailed discussions of this problem a possible

solutions, we refer to refs. [141–143]. Finally, within the adiabatic approximation TDDFT does not include double excitations. This could in principle be addressed by using a frequency-dependent exchange–correlation kernel [144, 145], but such approximations are not suitable for general applications yet.

**Linear-response coupled cluster.**

While conventional coupled-cluster theory does not provide a direct route to excited states, it can still be used as starting point for a response theory. Starting from the exponential parametrization of the (time-dependent) wavefunction,

$$|\tilde{\Psi}\rangle = \exp(\hat{T}(t))|\Phi_{\mathrm{HF}}\rangle, \tag{5.72}$$

and performing a projection of the time-dependent Schrödinger equation onto the the ground-state reference $\langle\Phi_0|$ as well as onto the set of excited determinants $\langle\mu_i|$ one arrives at the time-dependent analogs of Eqn. (5.19) and (5.20),

$$Q(\mathbf{t};t) = \langle\Phi_{\mathrm{HF}}|\big(\hat{H} + \hat{V}_\omega(t)\big)\exp(\hat{T}(t))|\Phi_{\mathrm{HF}}\rangle \tag{5.73}$$

$$0 = \Omega_{\mu_i}(\mathbf{t};t) = \langle\mu_i|\exp(-\hat{T}(t))\Big(\hat{H} + \hat{V}_\omega(t) - i\frac{\partial}{\partial t}\Big)\exp(\hat{T}(t))|\Phi_{\mathrm{HF}}\rangle \tag{5.74}$$

for the coupled cluster quasi-energy and time-dependent amplitude equations, respectively. One should note that, unlike methods such as time-dependent MP2, orbital relaxation is generally not included explicitly in the parametrization of Eqn. (5.72), as such effects can be incorporated via the $T_1$ amplitudes (hence the denomination "orbital-unrelaxed"). In what follows the "orbital-unrelaxed" approach is assumed.

As in the time-independent case, the coupled-cluster wavefunction is not determined variationally. Therefore, the coupled-cluster Lagrangian [41, 121]

$$L_{CC}(\mathbf{t},\bar{\mathbf{t}};t) = Q(\mathbf{t};t) + \sum_\mu \bar{t}_\mu \Omega_\mu(\mathbf{t};t), \tag{5.75}$$

is used as starting point for a response theory, where the multipliers $\bar{\mathbf{t}}$ have been introduced and the amplitude equations $\Omega_{\mu_i}(\mathbf{t};t)$ serve as constraints. The linear response function can then be determined by taking the derivative of the time-averaged Lagrangian $\{L_{CC}(\mathbf{t},\bar{\mathbf{t}};t)\}_T$, yielding the linear response function (cf. Eqn. (5.59))

$$\langle\langle\hat{\mu}_\alpha;\hat{\mu}_\beta\rangle\rangle_\omega = -\begin{pmatrix}\boldsymbol{\eta}_\beta\\ \boldsymbol{\xi}_\beta\end{pmatrix}^T\begin{pmatrix}\boldsymbol{F} & \boldsymbol{A}\\ \boldsymbol{A}^T & \mathbf{0}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\eta}_\alpha\\ \boldsymbol{\xi}_\alpha\end{pmatrix}, \tag{5.76}$$

where the matrix $\boldsymbol{J}$ is zero because the multipliers $\bar{\boldsymbol{\lambda}}$ only appear linearly in the Lagrangian. This response function can be rewritten as

$$\langle\langle\hat{\mu}_\alpha;\hat{\mu}_\beta\rangle\rangle_\omega = -\begin{pmatrix}\boldsymbol{\eta}_\beta\\ \boldsymbol{\xi}_\beta\end{pmatrix}^T\begin{pmatrix}\mathbf{0} & -(\boldsymbol{A}^T)^{-1}\\ -\boldsymbol{A}^{-1} & \boldsymbol{A}^{-1}\boldsymbol{F}(\boldsymbol{A}^T)^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{\eta}_\alpha\\ \boldsymbol{\xi}_\alpha\end{pmatrix}, \tag{5.77}$$

which reveals that the poles of the response function correspond to zero eigenvalues of the matrix $\boldsymbol{A}$, which is given by

$$A_{\mu\nu} = \frac{\partial^2\{L^{(2)}_{CC}(\mathbf{t},\bar{\mathbf{t}})\}_T}{\partial\bar{t}^{(1)}_\mu\partial t^{(1)}_\nu} = A_{\mu\nu} - \omega\,\delta_{\mu\nu} \tag{5.78}$$

where $A_{\mu\nu}$ are the elements of the Jacobian introduced in Eqn. (5.28). Therefore, excitation energies can be obtained as eigenvalues of this matrix. We note that the Jacobian can also be expressed as the derivative of the amplitude equations $\Omega_\mu$, with respect to $t_\nu$, because $\Omega_\mu = (\partial L_{\mathrm{CC}}/\partial\bar{t}_\mu)$ and $A_{\mu\nu} = (\partial^2 L_{\mathrm{CC}}/\partial\bar{t}_\mu\partial t_\nu)$. For details and for the the form of the matrix elements for the other quantities $(\mathbf{F},\boldsymbol{\xi}^Y,\ldots)$, appearing in coupled-cluster response theory we refer to the original literature [41, 121]. The approach outlined above is applicable to the different levels of the coupled cluster hierarchy [146],

starting with CCS (equivalent to CIS for excitation energies) and proceeding to CCSD, CCSDT and so on.

While feasible for small molecules, the relatively large computational cost of calculating and diagonalizing the CCSD Jacobian have motivated the development of more approximate coupled cluster methods based on perturbative approaches that yield energies correct to second order or higher. The first of these is the CC2 method [147], where the main idea is to retain the singles equation $\Omega_{\mu_1}$ as in CCSD but to approximate the doubles equation $\Omega_{\mu_2}$. This results in a Jacobian in which the doubles–doubles block is diagonal, thus resulting in significant computational gains. To make these linear response coupled-cluster methods, in particular CC2, applicable to truly large molecular systems, there has been significant work in recent years to combine it with efficient computational techniques [148, 149].

## 5.2 The molecular Hamiltonian in the relativistic framework

In the discussion above it was convenient to represent the molecular Hamiltonian in the Born-Oppenheimer approximation in its most general form,

$$\hat{H} = \sum_i \hat{h}_i + \frac{1}{2}\sum_{i\neq j} \hat{g}_{ij} + V_{NN}, \tag{5.79}$$

with $V_{NN}$ denoting a classical repulsion potential of clamped nuclei, since the structure of electronic structure approaches, in particularly WFT-based ones, remains essentially the same in a relativistic or non-relativistic frameworks[33, 34, 76]. In what follows we shall go a bit more in detail over the forms of the one and two-electron operators $\hat{h}_i$ and $\hat{g}_{ij}$, since these will determine whether we work in a non-relativistic or relativistic framework – and for the latter whether it corresponds to a two- or four-component approach[32, 35]. But before discussing many-electron wavefunctions we introduce the basic features of the Dirac equation for a free particle and simple atomic systems.

### The Dirac equation for free particles and hydrogen-like atoms

In order to treat time and momentum $\mathbf{p} = -i\hbar(\nabla_x, \nabla_y, \nabla_z)$ on the same footing and arrive at an equation that was invariant to Lorentz transformations, Dirac started out by proposing

$$\hat{H} = c\boldsymbol{\alpha}\cdot\mathbf{p} + \boldsymbol{\beta}mc^2 \tag{5.80}$$

as the Hamiltonian for the time evolution of a free particle,

$$\hat{H}\Psi = i\frac{\partial}{\partial t}\Psi \tag{5.81}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the $4\times 4$ matrices

$$\boldsymbol{\alpha} = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & \mathbf{0} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix} \tag{5.82}$$

and $\boldsymbol{\sigma}$ a vector composed of the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.83}$$

The form of the Hamiltonian implies a system of four coupled differential equations, instead of a single one in the case of the Schrödinger equation, and in the time-independent case one can verify that the Dirac equation possesses two branches of solutions separated by a large ($2mc^2$) energy gap: $-\infty < E < -mc^2$ and $mc^2 < E < +\infty$, which are also referred to as negative and positive energy solutions, respectively. From the solutions

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \tag{5.84}$$

one can also identify the charge ($\rho$) and current densities ($\mathbf{j}$) – which obey the continuity equation – as

$$\rho = \Psi^{\dagger}\Psi \tag{5.85}$$
$$\mathbf{j} = \Psi^{\dagger} c\boldsymbol{\alpha}\Psi \tag{5.86}$$

after identifying $c\boldsymbol{\alpha}$ with the velocity operator[33, 150].

External electric of magnetic fields can be introduced in Eq. 5.81 via the minimal coupling relation

$$E \rightarrow E + e\phi \tag{5.87}$$
$$\boldsymbol{\pi} \rightarrow \mathbf{p} + e\mathbf{A} \tag{5.88}$$

for the electron (charge $-e$), scalar potential $\phi$ and vector potential $\mathbf{A}$ respectively. Considering now the case of an electron in the presence of a static nucleus (so that $e\phi \equiv V = -Z/r$ and, in the reference frame of the nucleus, $\mathbf{A} = 0$) we can write the time-independent four-component Dirac equation – after rescaling $\boldsymbol{\beta}^{'} \rightarrow \boldsymbol{\beta} - \mathbf{I}_{4\times4}$ in order to align the relativistic and non-relativistic energy scales – as

$$\begin{pmatrix} V-E & c(\boldsymbol{\sigma}\cdot\mathbf{p}) \\ c(\boldsymbol{\sigma}\cdot\mathbf{p}) & V-E-2mc^2 \end{pmatrix} \begin{pmatrix} \psi^{\mathbf{L}} \\ \psi^{\mathbf{S}} \end{pmatrix} = 0 \tag{5.89}$$

where now we divided the four-component solution into the so-called large ($\psi^{\mathbf{L}}$) and small ($\psi^{\mathbf{S}}$) components, where "large" and "small" refer to the fact that solutions belonging to the positive energy solutions have the largest amplitudes for the large component and negative solutions have the largest amplitudes for the small component. In order to solve Eq. 5.89 the two-components $\psi^{\mathbf{L}}$ and $\phi^{\mathbf{S}}$ are expanded in terms of products of radial functions and two-component angular functions $\chi_l^{j,m_j}$ which are expressed in a basis of products of a spherical harmonic ($Y_l^{m_l}(\rho,\theta)$) and a Pauli spinor ($\xi(m_s)$),

$$\psi^{\mathbf{L}} = R^L(r)\chi_{j,m_j}(\rho,\theta) = R^L(r)\sum_{m_l,m_s}\langle lm_l sm_s|jm_j\rangle Y_l^{m_l}(\rho,\theta)\xi(m_s) \tag{5.90}$$

with $\langle lm_l sm_s|jm_j\rangle$ a Clebsch-Gordan coefficient.

Due to the fact that $\boldsymbol{\sigma}\cdot\mathbf{p}$ and consequently the Hamiltonian does not commute individually with the angular momentum ($\hat{l}$) or spin ($\hat{s}$) but does with the total angular momentum operator

$$\hat{j} = \hat{l} + \hat{s}, \tag{5.91}$$

only the latter will be a proper constant of motion in the relativistic framework. Another constant of motion here is the Dirac quantum number $\kappa$,

$$\kappa = \begin{cases} -(l+1) & \text{if} \quad j = l + 1/2 \\ l & \text{if} \quad j = l - 1/2 \end{cases} \tag{5.92}$$

of the angular functions $\chi_l^{j,m_j}$ for the angular momentum operator $\hat{K} = (\sigma\cdot\hat{l} + 1)$. Equation 5.91 is the manifestation of spin-orbit coupling that can be understood[32, 150] by considering the reference frame of the moving electron, which will now be subjected not only the Coulomb but also to the vector potential due to the (in the new reference frame, moving) nucleus.

Equation 5.89 turns out to have, apart from the two branches previously discussed, a number of discrete solutions in the $[-2mc^2, 0]$ interval with energies

$$E(n,\kappa,Z) \simeq -\frac{Z^2}{2n^2}\left[1 + \frac{1}{n}\left(\frac{Z^2}{c^2}\right)\left(\frac{1}{|\kappa|} - \frac{3}{4n}\right)\right] \tag{5.93}$$

where $n$ is the non-relativistic principal quantum number. From this expression we see that due to their dependence on $j$ through $\kappa$, the energy levels with $l \neq 0$ will no longer be strictly degenerate

Table 5.1: Quantum number for relativistic atomic orbitals

| | Labels | | | | | | |
|---|---|---|---|---|---|---|---|
| | $s_{1/2}$ | $p_{1/2}$ | $p_{3/2}$ | $d_{3/2}$ | $d_{5/2}$ | $f_{5/2}$ | $f_{7/2}$ |
| $l$ | 0 | 1 | 1 | 2 | 2 | 3 | 3 |
| $j$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{3}{2}$ | $\frac{3}{2}$ | $\frac{5}{2}$ | $\frac{5}{2}$ | $\frac{7}{2}$ |
| $\lvert m_j\rvert$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}, \frac{3}{2}$ | $\frac{1}{2}, \frac{3}{2}$ | $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$ | $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$ | $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$ |
| $\kappa$ | -1 | +1 | -2 | +2 | -3 | +3 | -4 |

and the spinors will have significantly different spatial extents with increased $Z$ as can be seen in figure 4.5. Furthermore, spinors with same $j$ (see table 5.1) will show the same angular dependence ($s_{1/2}$ and $p_{1/2}$ are both spherical, $p_{3/2}$ and $d_{3/2}$ look like non-relativistic $p$ orbitals and so on) and will be nodeless for all but the highest $m_j$ values.

Equation 5.89 can be expressed in terms of $\psi^{\mathbf{L}}$ only, by isolating $\psi^{\mathbf{S}}$ in the second equation and substituting it back on the first, yielding the two-component expression

$$\left[(V-E)+\frac{c^2(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p})}{E+2mc^2-V}\right]\psi^{\mathbf{L}}=0 \tag{5.94}$$

that can be refactored as

$$\left[(T+V-E)+\left(\frac{(V-E)}{E+2mc^2-V}T\right)+\left(\frac{c^2}{(E+2mc^2-V)^2}\times\{(\mathbf{p}V)\cdot\mathbf{p}+i\boldsymbol{\sigma}\cdot(\mathbf{p}V)\times\mathbf{p}\}\right)\right]\psi^{\mathbf{L}}=0 \tag{5.95}$$

so that one can identify the Schrödinger equation (first term), plus scalar (second term) and spin-orbit (third term) corrections to it, and see than in the non-relativistic limit ($c\rightarrow\infty$) it reduces to the Schrödinger equation. Here $T=(\boldsymbol{\sigma}\cdot\mathbf{p})(\boldsymbol{\sigma}\cdot\mathbf{p})/(2m)$ is the non-relativistic kinetic energy operator.

### The Dirac equation for many-electron systems

The extension of the formalism above for the many-electron case is relatively straightforward for the one-electron terms $\hat{h}_i$, defined as

$$\hat{h}_i\equiv\hat{h}_D(i)=c\boldsymbol{\alpha}\cdot\mathbf{p}(i)+\boldsymbol{\beta'}mc^2+v_{nuc} \tag{5.96}$$

where $v_{nuc}=v_{nuc}\mathbf{I}_{4\times4}$ corresponds to the usual nuclear potential from an atom or the molecular framework.

A fully Lorentz invariant description of the electron-electron interaction, on the other hand, is not a simple task as one would have to take into account the fact that the interaction between electrons is not instantaneous but rather shows a retardation as it propagates at the speed of light. This can be done in the QED framework, via perturbation theory on the basis of the exchange of virtual photons (for a more detailed discussion see [33] and references therein), where one arrives at the expression

$$\hat{g}_{ij}=\frac{1}{r_{ij}}-\frac{\boldsymbol{\alpha}_i\cdot\boldsymbol{\alpha}_j}{r_{ij}}-\frac{(\boldsymbol{\alpha}_i\cdot\boldsymbol{\nabla}_i)(\boldsymbol{\alpha}_j\cdot\boldsymbol{\nabla}_j)\cos(\omega r_{ij}-1)}{\omega^2 r_{ij}}+\ldots \tag{5.97}$$

for the exchange of a virtual photon of frequency $\omega=|\epsilon_i-\epsilon_j|/c$ ($\epsilon_i,\epsilon_j$ the energies for spinors $i$ and $j$). In the case of molecular or atomic systems, since for the most part the differences in spinor energies will not be very large (within the valence), or if they are large (between valence and core) the electrons will be close to each other, retardation effects will not be very large and one can therefore use the following expression for the two-electron interaction

$$\hat{g}_{ij}=\left[\frac{\mathbf{I}_{4\times4}}{r_{ij}}\right]-\left[\frac{\boldsymbol{\alpha}_i\cdot\boldsymbol{\alpha}_j}{2r_{ij}}\right]-\left[\frac{(\boldsymbol{\alpha}_i\cdot\mathbf{r_{ij}})(\boldsymbol{\alpha}_j\cdot\mathbf{r_{ij}})}{r_{ij}^2}\right]=\hat{g}_{ij}^{\text{Coulomb}}+\hat{g}_{ij}^{\text{Gaunt}}+\hat{g}_{ij}^{\text{gauge}}=\hat{g}_{ij}^{\text{Coulomb}}+\hat{g}_{ij}^{\text{Breit}} \tag{5.98}$$

where $\hat{g}_{ij}^{\text{Coulomb}}$ and $\hat{g}_{ij}^{\text{Breit}} = \hat{g}_{ij}^{\text{Gaunt}} + \hat{g}_{ij}^{\text{gauge}}$ represent the electrostatic and magnetic interactions between electrons, respectively. One can understand the latter by recalling the identification of $c\boldsymbol{\alpha}$ with the velocity operator, so that the two-electron integral over $\hat{g}_{ij}^{\text{Gaunt}}$ can be viewed as representing current-current interactions, which will then shown a gauge dependence. As it may be cumbersome to implement the $\hat{g}_{ij}^{\text{gauge}}$ term, in practice electronic structure calculations will use either

$$\hat{g}_{ij} = \hat{g}_{ij}^{\text{Coulomb}} \tag{5.99}$$

defining the Dirac-Coulomb (DC) Hamiltonian, or

$$\hat{g}_{ij} = \hat{g}_{ij}^{\text{Coulomb}} + \hat{g}_{ij}^{\text{Gaunt}} \tag{5.100}$$

defining the Dirac-Coulomb-Gaunt (DCG) Hamiltonian.

Once defined the one and two-electron parts of the Hamiltonian, one can proceed with the formulation and implementation of electronic structure approaches. Their point of departure, as in non-relativistic approaches, is the construction of a Slater determinant on the basis of molecular spinors (MS) described in terms of atom-centered basis sets, with its subsequent use evaluating the matrix elements of the Hamiltonian,

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \sum_{i=1}^{n} (\phi_i | \hat{h}_D | \phi_i) + \frac{1}{2} \sum_{i,j=1}^{n} [(\phi_i \phi_i | \phi_j \phi_j) - (\phi_i \phi_j | \phi_j \phi_i)]. \tag{5.101}$$

Due to the presence of negative energy solutions, in determining the ground-state wavefunction one does not proceed via a simple minimization of Eq. 5.101 but rather by a minimization with respect to positive energy solutions and maximization with respect to negative solutions [151].

Another significant difference is that, by virtue of the one-electron operators, the spinors and therefore the wavefunction will be complex quantities, something that along with the loss of spin symmetry will have consequences such as loss of permutational symmetry of two-electron integrals, and the size of the matrices representing the operators is therefore much larger than their non-relativistic counterparts, though one can, in the absence of vector fields, redefine the one and two-electron operators in such a way as to exploit time-reversal symmetry[33].

An elegant framework which can exploit time-reversal symmetry while at the same time reduce the size of matrix representations was proposed by Saue *et al.* [151], and involved recasting the SCF problem in terms of quaternion algebra. The exploitation of time-reversal symmetry together with the exploitation of point group symmetry[152] effectively allows, in the case of point groups such as $C_2$, $C_{2v}$ and $D_{2h}$, the use of real diagonalization. Another appealing feature that it allows for expanding the molecular spinors in terms of scalar basis functions such as those commonly used in non-relativistic codes and therefore adapt and continue to use non-relativistic molecular integral codes. This way, a spinor can be written in terms of a large ($\varphi_i^L(\mathbf{r})$) and a small ($\varphi_j^S(\mathbf{r})$) component basis sets

$$\phi_k = \begin{pmatrix} \boldsymbol{\Phi}_k^L \\ \boldsymbol{\Phi}_k^S \end{pmatrix} = \begin{pmatrix} \phi_k^{L\alpha} \\ \phi_k^{L\beta} \\ \phi_k^{S\alpha} \\ \phi_k^{S\beta} \end{pmatrix} = \begin{pmatrix} \sum_i^{M_L} C_{ik}^{L\alpha} \varphi_i^L(\mathbf{r}) \\ \sum_i^{M_L} C_{ik}^{L\beta} \varphi_i^L(\mathbf{r}) \\ \sum_j^{M_S} C_{jk}^{S\alpha} \varphi_j^S(\mathbf{r}) \\ \sum_j^{M_S} C_{jk}^{S\beta} \varphi_j^S(\mathbf{r}) \end{pmatrix} \tag{5.102}$$

as opposed to expanding the spinors in terms of two-component functions[153, 154]. A drawback of using scalar basis sets, on the other hand, is that one must ensure that the kinetic balance condition[155, 156]

$$\boldsymbol{\phi}_k^S = \left[ \frac{c(\boldsymbol{\sigma}\cdot\mathbf{p})}{E + 2mc^2 - V} \right] \phi^L \simeq \frac{1}{2mc} (\boldsymbol{\sigma}\cdot\mathbf{p}) \phi^L \tag{5.103}$$

is respected in the matrix representation. In the absence of vector fields, this means that the large and small component basis respect the relationship

$$\{\varphi_j^S(\mathbf{r})\} \supseteq \{(\boldsymbol{\sigma}\cdot\mathbf{p})\varphi_j^L(\mathbf{r})\} \tag{5.104}$$

Because of kinetic balance, the total basis set size in a four-component calculation will be rather large and contain higher angular momentum functions than the original large-component set as the operation of $(\boldsymbol{\sigma}\cdot\mathbf{p})$ on the latter yields

$$\frac{\partial}{\partial x}x^n e^{-\gamma r^2} = (nx^{n-1} - 2\gamma x^{n+1})e^{-\gamma r^2}, \tag{5.105}$$

Substituting Eq. 5.102 in the expression of the two-electron integrals in Eq. 5.101 yields different classes of integrals corresponding to electrostatic interaction between charge densities for the large and small components $((LL|LL), (LL|SS), (SS|SS))$ and those corresponding to the Gaunt interaction $((LS|LS))$, whenever used. Among these, the $(SS|SS)$ ones are the most numerous but, at the same time, are extremely localized around the atoms since it is there the small component amplitudes will be largest. The recognition of this locality has led to an approximation[157] where calculating the $(SS|SS)$ integrals is completely avoided and their contribution to the energy obtained as a classical repulsion between atomic charges.

### The treatment of electron correlation and of molecular properties

Thorough surveys of correlated approaches in the relativistic domain can be found in [158, 159]. We note here (a non-exhaustive list of) papers discussing the implementation of coupled cluster [101–103, 160, 161], CI[162–169], and ADC[170–172] in connection to DIRAC.

The most important point concerning the formulation of wavefunction-based methods for electron correlation is that these are virtually always based on the so-called "no-pair" approximations[35], where the two-electron part of the Hamiltonian is projected onto the positively energy solutions in order to allow the creation of excited determinants only involving electrons. In this case, the approaches outlined above for the calculation of excitation energies and properties due to electric perturbations are by and large transposable to the relativistic domain, especially when formulated in a spin-orbital formalism. That said, one should carefully account for the fact that relativistic wavefunctions are complex – see for instance the thorough discussion by Shee on this matter for the implementation of CCSD first-order properties [173] – as well as to the contribution of negative-energy solutions to orbital relaxation effects (with respect to the SCF wavefunction), as is the case of MCSCF wavefunction optimizations [174] and the determination of properties for non-variational wavefunctions [175].

Concerning the calculation of molecular properties by response theory, in DIRAC both linear [125, 176–178] and quadratic response [179–181] properties can be calculated for mean-field wavefunctions such as Hartree-Fock and DFT, and taking advantage of the efficiency of the quaternion-based framework, as do the four-index transformation [182] and coupled-perturbed Hartree-Fock code used to obtain MP2 first-order properties [175]. In the case of DFT, however, one should note that as a practical realization of a fully relativistic formulation of DFT [183] requires the availability of suitable density functionals and these are not generally available. In DIRAC and in other codes one therefore resorts to using non-relativistic functionals [184] as well as to working within the spin-density functional theory (SDFT) [178] for response theory.

For magnetic properties, on the other hand, the linearity of the Dirac Hamiltonian with respect to the inclusion of vector fields is a feature that sets it apart from the non-relativistic world, as it greatly simplifies their formulation [33, 185]. Here, both the paramagnetic and diamagnetic terms from nonrelativistic formalisms are obtained from a single operator in the relativistic domain [186], the diamagnetic terms having been shown to arise from the contributions of negative energy states to the response of the wavefunction to the perturbations.

With respect to the discussion above, the presence of a vector field means the restricted kinetic balance relation must be modified accordingly, yielding the so-called restricted magnetic balance (RMB) condition [186–190], which has been fully realized in practice for the first time in [191]. If one considers the Zeeman ($\hat{h}_B$) and hyperfine ($\hat{h}_{m_K}$) operators central to describing NMR spectra,

$$\hat{h}_B = \frac{1}{2}(\vec{r}_G \times c\vec{\alpha}), \qquad \hat{h}_{m_K} = \frac{1}{c^2}\frac{\vec{r}_K \times c\vec{\alpha}}{r_K^3}, \tag{5.106}$$

where $\vec{r}_X = \vec{r} - \vec{R}_X$ with an arbitrary gauge origin $\vec{R}_G$ and the center of nucleus $K$ in $\vec{R}_K$, we have for RMB the expression the expression

$$\{\varphi_j^{S;[\mathbf{m^K},\mathbf{B}]}(\mathbf{r})\} \supseteq \{\left(\boldsymbol{\sigma}\cdot\mathbf{p} + \frac{1}{c}\,\boldsymbol{\sigma}\cdot\,\hat{h}_B + \frac{1}{c}\,\boldsymbol{\sigma}\cdot\,\hat{h}_{m_K}\right)\varphi_j^L(\mathbf{r})\} \tag{5.107}$$

for the small component while the large component basis remains unchanged. These are subsequently used as a basis for the response treatment whereby the effect on perturbed wavefunctions are obtained,

$$\phi_k^L(B;m_K) = C_k^L(B;m_K)\varphi_j^L(\mathbf{r}) \tag{5.108}$$

$$\phi_k^S(B;m_K) = C_k^S(B;m_K)\varphi_j^{S;[\mathbf{m^K},\mathbf{B}]}(\mathbf{r}) \tag{5.109}$$

In DIRAC a simpler approach was taken in the use of the unrestricted kinetic balance (UKB) [151, 192–194].

In the presence of external magnetic fields a gauge dependence on results of the corresponding magnetic properties is introduced, which is only removed in the limit of complete basis sets. Thus, it is common practice in such cases to use perturbation-dependent basis sets such as the so-called London orbitals (LAOs)

$$\omega_\mu^K(\vec{r}) = exp\left\{-\frac{i}{2}\vec{B}\times(\vec{R}_K - \vec{R}_G)\cdot\vec{r}\right\}\varphi_\mu^K(\vec{r}), \tag{5.110}$$

which guarantee the gauge-origin invariance of results in finite basis approximation. LAOs are also appealing when used in combination with RMB as they make the magnetic balance atomic and easy to handle by the simple scheme (sMB).[194]

### Approximate Hamiltonians

Whatever the computational efficiency of four-component implementations, it remains the case that they are computationally expensive approaches, and if one wishes to apply relativistic approaches to large systems it is preferable to devise approximations that can reproduce the four-component results as faithfully as possible. One way of doing so is to proceed in a similar vein to what was done in Eqs. 5.94 and 5.95, and obtain a transformed equation which is only defined for the positive energy branch but that still reproduces the spectrum of the original Hamiltonian.

This can be achieved by applying a unitary block diagonalization to the (usually the one-electron) Hamiltonian [195]

$$\mathbf{U}\,\dagger \begin{bmatrix} h_{D;LL} & h_{D;LS} \\ h_{D;SL} & h_{D;SS} \end{bmatrix}\mathbf{U} = \begin{bmatrix} h_{++} & \mathbf{0} \\ \mathbf{0} & h_{--} \end{bmatrix} \tag{5.111}$$

where $\mathbf{U}$ is defined via the so-called decoupling operators $\mathbf{R}$ unique to each approximate method,

$$\mathbf{U} = \begin{bmatrix} \mathbf{1} & -\mathbf{R}^\dagger \\ \mathbf{R} & \mathbf{1} \end{bmatrix}\begin{bmatrix} (\mathbf{1}+\mathbf{R}^\dagger\mathbf{R})^{-1/2} & \mathbf{0} \\ \mathbf{0} & (\mathbf{1}+\mathbf{R}\mathbf{R}^\dagger)^{1/2} \end{bmatrix} \tag{5.112}$$

and we have that an exact decoupling operator is

$$\mathbf{R} = \frac{c(\boldsymbol{\sigma}\cdot\mathbf{p})}{E + 2mc^2 - v_{nuc}} \tag{5.113}$$

Examples for approximate approaches can be found in [195]. In the discussion that follows the ZORA [196, 197] and the exact two-component (X2C) [198–201] Hamiltonians are used, as well as the X2C approach based on molecular mean-field (X2Cmmf) solutions [202]. The X2C approach has gained popularity in recent years since it is defined in a purely matrix form, not requiring extensive derivations such as the Douglas-Kroll-Hess scheme [203].

A difference between the X2Cmmf [202] scheme and the others is that in it one solves the Dirac equation for the molecular system, using either the DC or DCG Hamiltonians (thus including spin-same orbit (SSO) and spin-other orbit (SOO, for DCG) interactions exactly) and at the end of the SCF

performs the transformation above, instead of starting from a free particle Hamiltonian and adding spin-orbit interactions in an approximate manner, with atomic mean-field integrals for instance [204]. The X2Cmmf makes no sense if one wants to accelerate mean-field calculations (Hartree-Fock, DFT), but is particularly interesting for correlated calculations as it will retain the accuracy the original calculations (as discussed in paper C.5) while greatly reducing the computational cost of the four-index transformation step.

Another important point related to the transformation above is that not only the Hamiltonian but also properties have to be transformed, otherwise so-called "picture-change" errors [195, 205] will be introduced and produce unreliable results.

The reader is referred to a recent review[206] for a discussion of approaches available to obtain molecular properties with emphasis on approximate Hamiltonians such as X2C.

## 5.3 A comparison of IHFSCC to experiment for heavy elements

In this section I summarize my main results in the application of IHFSCC in a two- or four-component framework, focusing on systems for which experimental data is readily available, and refer the reader to the manuscripts in appendices C and D to a more complete discussion and the background for each study.

Before discussing any results, it is useful to briefly discuss the strategies used in the calculations. As mentioned above in (IH)FSCC one starts by defining a reference Slater determinant, on which the wave operator will act in order to create the target electronic states for a given sector of Fock space. The constraint of starting from a single determinant means that, depending on the states we wish to treat, we might not be able to start from the species we are actually interested and need, instead, to start from an auxilary species (hereafter referred to as M) and add or remove electrons from it until we arrive at the species and states we want, as shown in figure 5.1.

(a) $\mathrm{M}(0,0) \rightarrow \mathrm{M}^{+}(1,0) \rightarrow \mathrm{M}^{-}(0,1) \rightarrow \mathrm{M}^{(*)}(1,1)$

(b) $\mathrm{M}^{2+}(0,0) \rightarrow \mathrm{M}^{+}(0,1) \rightarrow \mathrm{M}^{(*)}(0,2)$

(c) $\mathrm{M}^{2-}(0,0) \rightarrow \mathrm{M}^{-}(1,0) \rightarrow \mathrm{M}^{(*)}(2,0)$

Figure 5.1: Schematic representation of the possible ways obtaining electronic states for a target species $\mathrm{M}^{(*)}$ starting from auxiliary species (M, $\mathrm{M}^{2+}$ or $\mathrm{M}^{2-}$) well-described by a single Slater determinant. It should be noted that M might itself be a neutral, anionic or cationic species

A first example of this process can be found in the study of the $f-f$ spectrum of the bare actinyl species $\mathrm{NpO}_2^{+}$, $\mathrm{NpO}_2^{2+}$, $\mathrm{PuO}_2^{2+}$ (paper C.1) where we studied the $f^1$ and $f^2$ configurations of these species, as well as for the $\mathrm{NpO}_2\mathrm{Cl}_4^{2-}$ where we were interested in the $f^1$ electronic states (paper D.1). Since in these cases the ground states correspond to open-shell configurations, we had to start with auxiliary systems with one (for $\mathrm{NpO}_2^{2+}$ and $\mathrm{NpO}_2\mathrm{Cl}_4^{2-}$) and two (for $\mathrm{NpO}_2^{+}$ and $\mathrm{PuO}_2^{2+}$) less electrons and proceed to add electrons to the auxilary species $\mathrm{NpO}_2^{3+}$, $\mathrm{NpO}_2\mathrm{Cl}_4^{-}$ and $\mathrm{PuO}_2^{4+}$ via the single and double electron attachement sector (case b in figure 5.1).

The results from these calculations are found in table 5.2. Due to the constraints in the computational resources at the time, all these calculations were performed with double-zeta basis and, while for the bare actinyls the DC Hamiltonian was use, for neptunyl chloride we used the X2C Hamiltonian[207] with two-electron spin same-orbit (SSO) and spin-other orbit (SOO) contributions were included via atomic mean-field integrals obtained with the AMFI[208, 209] code. In spite of these limitations, we observe that the IHFSCC results show a very good agreement with experimental ones – even though one should keep in mind that these have been measured not in gas phase but in solution – and tend to perform much better than more approximate methods such as MRCI and CASPT2 with a perturbative treatment of spin-orbit coupling, either in a rather approximate manner

as in the restricted active space state interaction (RASSI) approach [210] as used in the Molcas [211] codes, or via a more sophisticated spin-orbit CI approach as implemented in the EPCISO code [212].

An appealing feature of IHFSCC was that, not only the energies for the different states were in better agreement with experiment, but also their symmetry classification followed the experimental assignments, whereas for the CI and CASPT2 states there were some changes in ordering. At the time paper C.1 was published, we were not sure of the extent to which these differences between methods were mostly due to the difference in the correlation treatment or also had to do with the perturbative treatment of spin-orbit coupling. Here it is important to have in mind that in the two-step approach it is essential to calculate a sufficiently large number of spin-free states of different symmetries, otherwise the description of spin-orbit coupling will be deficient and can contribute to a uneven quality in the final electronic states (irrespective of a good quality in the description of the spin-free states).

In this sense, calculations which treat spin-orbit coupling variationally and construct the many-electron wavefunctions in terms of molecular spinors end up requiring a bit less effort in the multi-reference description. That said, the drawback of the DIRAC implementation which only allows the use of the Fock space sectors presented in figure 5.1 is that we are effectively restricted to studying the spin-orbit coupled states arising from at most triplet configurations.

As this was, to the best of my knowledge, the first application of IHFSCC approach to such systems, a particularly time-consuming part in this first study was to properly set-up the model ($P_m$) and intermediate ($P_i$) spaces for the calculations: in the minimal configuration tested, the $P$ space contained 25 spinors, of which the six lowest-lying unoccupied $5f$ spinors were placed in $P_m$ (for both Pu and Np) and the remaining spinors, consisting of roughly the $7p$ and the next $f$ shell were were placed in $P_i$. In spite of using such extended model spaces, we observed a rather slow convergence of the CC iterations but managed to properly converge the calculations.

Table 5.2: Low-lying $f$-$f$ excited state energies (in $cm^{-1}$) for $NpO_2^{2+}$ and $NpO_2Cl_4^{2-}$, using (a) the gas-phase geometry from ref.'[213], $r_{NpO} = 1.675$ Å; (from the calculations of Ref. [213] (b) the X-ray structure from ref. [214], $r_{NpO} = 1.775$ and $r_{NpCl} = 2.653$ Å respectively; (c) 170.1 pm (d) 176.6 pm For comparison the experimental values due to Denning and coworkers[215] are shown.

| | Excited Electronic States | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Species | I | II | III | IV | V | VI | VII | VIII |
| $NpO_2^{2+}$ | $\Delta_{3/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ | | | |
| IHFSCC(a) | 3221 | 8565 | 7225 | 30877 | 34947 | | | |
| IHFSCC(b) | 4297 | 9661 | 7229 | 29021 | 32379 | | | |
| SO-RASPT2(c) [216] | 3575 | 7798 | 6108 | – | – | | | |
| $NpO_2Cl_4^{2-}$ | $\Delta_{3/2u} + \Phi_{5/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ | | | |
| IHFSCC | 886 | 7679 | 9262 | 20018 | 22445 | | | |
| SO-RASPT2 [216] | 1055 | 5767 | 6658 | – | – | | | |
| Exp.[215] | 900-1050 | 6880 | 7890 | 17241 | 20081 | | | |
| $NpO_2^{+}$ | $\Sigma_{0g}$ | $\Pi_{1g}$ | $^3H_{5g}$ | $^3\Pi_{0g}$ | $^3\Sigma 1g$ | $^3\Pi_{0g}$ | $^3H_{6g}$ | $^3\Pi_{2g}$ |
| IHFSCC | 2527 | 4102 | 5379 | 8628 | 8929 | 9378 | 9690 | 10056 |
| SO-CI [217] | 3366 | 4938 | 4721 | 9537 | 9076 | 9708 | 8867 | 11187 |
| Exp. [218] | – | – | 6173 | 8953 | 9146 | 9780 | – | 10208 |
| $PuO_2^{2+}$ | $\Sigma_{0g}$ | $\Pi_{1g}$ | $^3H_{5g}$ | $^3\Pi_{0g}$ | $^3\Sigma 1g$ | $^3\Pi_{0g}$ | $^3H_{6g}$ | $^3\Pi_{2g}$ |
| IHFSCC | 2530 | 4870 | 6700 | 10334 | 10983 | 11225 | 11651 | – |
| SO-CI [219] | 4295 | 7044 | 6593 | 7393 | 12874 | 9415 | 7848 | 14169 |
| SO-CASPT2 [220] | 4190 | 6065 | 8034 | 12874 | 12906 | 14606 | 13326 | 14910 |
| Exp. [218] | – | – | – | 10185 | 10500 | 10700 | – | 12037 |

As was the case for the bare actinyls, in spite of the modest basis sets the results for neptunyl chloride also compare very favourably to experiment, with rather good agreement for the three lowest-lying excited states (discrepancies in the order of 10-15%) but less so for the two higher-lying Π states. It is interesting to note that more recent SO-RASPT2 calculations [216] haven't improved on the accuracy of IHFSCC.

Paper C.2 presents a second example of the usage of IHFSCC, in the study of the potential energy surfaces for the ground and optically active excited states of $I_3^-$. In this case we use case (a) in figure 5.1, since we are looking for singly excited states. In table 5.3 the results for IHFSCC, SO-CASPT2 (using Molcas) and (four-component) MRCI calculations is presented for the two optically active excited states, and from that we clearly see that the IHFSCC results are very close to the experimental values for both transitions, whereas CASPT2 overestimates both by about 2400 $cm^{-1}$ and MRCI is in good agreement with experiment for the first but overestimate the second state by nearly the same amount as CASPT2.

The difference between the (a) and (b) IHFSCC (or MRCI) results is the extent to which we truncate the basis set and correlation space and, therefore, approximate the electron correlation treatment: in (a) we use an augmented triple zeta basis and correlate spinors between -1 and 4 hartree (including the occupied orbital space the $\sigma, \pi$ bonding system and three other $\sigma$-type orbitals arising from the $5s5p$ orbitals of iodine), whereas in (b) augmented core-valence triple zeta basis are used, and spinors between -3 and 12 hartree are correlated (adding the occupied $4d$ electrons of iodine and many more virtuals). In both cases the $P_m$ space contained 8 occupied (one $\sigma_g$, one $\pi_g$, two $\pi_u$ and one $\sigma_u$ and 11 virtual spinors (two $\sigma_g$, one $\pi_g$, three $\sigma_u$ and two $\pi_u$ orbitals), with a full $P$ space containing 11 occupied (5 of $g$ and 6 of $u$ parity) and 22 virtual (12 of $g$ and 10 of $u$ parity) spinors.

The truncation of the correlation space ($Q$) is a strategy that had widespread use in the DIRAC user and developer community for pragmatic reasons as it greatly decreases the cost of the step of integral transformation step. But one must be careful in using to avoid introducing artifacts in the calculations by, for instance, removing part of the high angular momentum functions that are essential to the treatment of electron correlation: one should view the basis sets as an equally important component in the treatment of electron correlation as the many body methods themselves, since no reliable results can come from the use of a very sophisticated correlation method and a basis set poorly designed for correlation. I discuss this point further in reference [221], which outlines the development of basis sets for correlated calculations on lanthanides.

Table 5.3: Vertical ($T_v$) excitation energies (in $cm^{-1}$) calculated with the MRCI (a), CASPT2 and IHFSCC(a and b) methods for $I_3^-$ calculated at the optimum bond length for each method for the optically active $0_u^+$ states

| State | MRCI(a) | CASPT2 | IHFSCC(a) | IHFSCC(b) | Experimental |
|---|---|---|---|---|---|
| $0_u^+$ a | 27986 | 30002 | 27744 | 27098 | 27421 |
| $0_u^+$ b | 37180 | 37664 | 33631 | 34438 | 34680 |

The results in paper C.2 also show that the difference between IHFSCC and SO-CASPT2 or MRCI for the excited state energies is not constant but varies greatly from state to state, as can be seen in figure 5.2. Unfortunately, it was not possible for us to avoid problems with intruder states in the IHFSCC calculations as the I-I bonds were stretched beyond 3.58 Å, as it can be seen in figure 2 of paper C.2, in spite of the relatively large model space. This same figure shows that the SO-CASPT2 method, on the other hand, was perfectly capable of dissociating whem employing a similar CASSCF active space (consisting of 16 electrons in the 9 orbitals arising from the $5p$ orbitals of the three iodine atoms) but we nevertheless observe spikes in the potential curves at about 3.5 Åand beyond, signaling a weak intruder state problem in the (spin-free) CASPT2 treatment that could not be remedied in spite of us applying real [108] and imaginary [109] level shifts up to 0.20 $E_h$.

Because of the need to go through the $(1, 0)$ sector of Fock-space, the study of $I_3^-$ also yielded, at no additional cost, a number of electronic states for the $I_3$ radical shown in table 5.4 as well as its

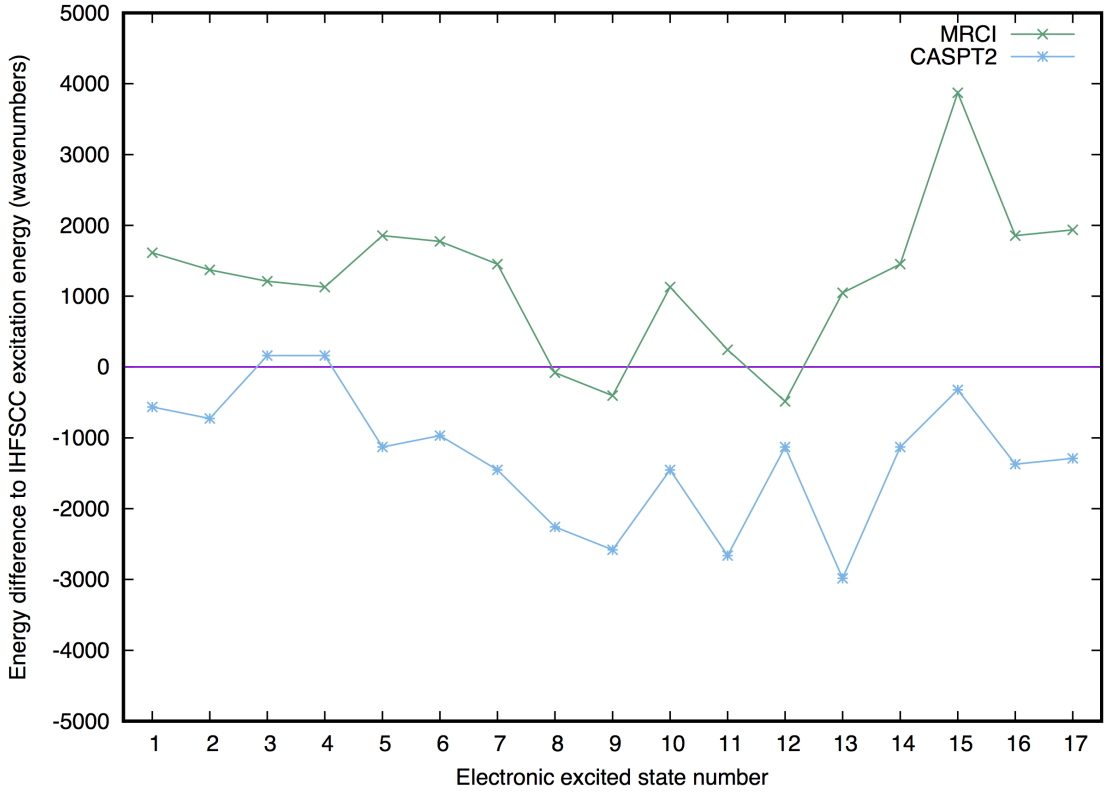



Figure 5.2: Comparison of methods at r=2.93. IHFSCC(a) is the reference

electron affinities. From these results, we see that the electron affinities obtained with IHFSCC are within the experimental error bounds. Furthermore, we see that the first two $\Omega = 1/2$ excited states are in very good agreement with experiment. The agreement with experiment for the third $\Omega = 3/2$ state, on the other hand, is particularly poor if one compares the vertical ($T_v$) transition energy but improves somewhat for the adiabatic ($T_e$) value. While there were no experimental bond lengths available for comparison at the time, we see that the ground-state IHFSCC harmonic vibrational frequency is significantly higher than the experimentally derived on. As we shall see, this appears to be a systematic problem for IHFSCC calculations.

Table 5.4: IHFSCC (b) excitation energies $T_v$ (in $cm^{-1}$) and electron affinities (in eV) for the $I_3$ radical

| | Experimental | | | IHFSCC (b) | | | | |
|---|---|---|---|---|---|---|---|---|
| State ($\Omega$) | $\omega_e$ | $T_e$ | EA[1] | $R_e$ | $\omega_e$ | $T_v$ | $T_e$ | EA[2] |
| X $(3/2_u)$ | 115±5 | 0 | 4.15±0.12 | 2.828 | 132 | 0 | 0 | 4.29 |
| A $(1/2_g)$ | | 2258 | | 2.884 | 113 | 2016 | 1855 | |
| B $(1/2_u)$ | | 5000 | | 2.837 | 129 | 5424 | 5242 | |
| C $(3/2_g)$ | | 5484 | | 2.951 | 115 | 7097 | 6291 | |

[1] Ref.[222]; vertical detachment energy of $I_3^-$: 4.25 eV (ref.[223])

[2] Value for the adiabatic EA. Vertical processes: IP($I_3^-$) = 4.39 eV, EA($I_3$) = 4.20 eV.

Another radical species whose electronic states have been studies (in paper C.6) with IHFSCC is IO, a species of interest to atmospheric sciences whose spectroscopic constants, excitation energies, ionization potentials and electron affinities (presented in table 5.5) were, much like the $I_3$ radical, a by-product of our interest in studying the $IO^+$ and $AtO^+$ species, due to the interest of the latter in the field of nuclear medicine[5, 224–227].

Since at the outset we were not sure that the $IO^+$ and $AtO^+$ species could be investigated with single-reference approaches as their low-lying states originate from the $p$–$p$ manifold, our strategy was

to start from the $IO^-$ (and $AtO^-$) species and remove electrons following case (c) in figure 5.1. In these calculations we used the DC Hamiltonian and augmented triple zeta basis and correlated all electrons between -10.0 and 100.0 au, corresponding to 32 correlated electrons and 248 virtual spinors for the systems containing iodine (and 46 correlated electrons and 340 virtual spinors for the systems containing astatine).

For both species, the model ($P_m$) space used contained the valence $\sigma_{1/2}, \pi_{3/2}, \pi_{1/2}, \pi^*_{3/2}, \pi^*_{1/2}$ spinors with the exception of the $\sigma^*_{1/2}$, which is unoccupied in the anion reference, and means that electronic states with important contributions from Slater determinants in which $\sigma^*_{1/2}$ is occupied will not be properly described. Also, since we proceed via the ionization sectors, the spinors lower in energy that are not part of the correlation space belong to $P_i$.

We see from our results that IHFSCC again tend to overestimate the harmonic frequencies in a systematic manner, but bond lengths and excitation energies show an extremely good agreement with experiment. If one compares these results to the single-reference CCSD(T) calculations of Peterson [228], one observes an equally good agreement for bond lengths and an improvement on the vibrational frequencies that indicates that the missing ingredient in the IHFSCC calculations is a proper treatment of triples (and higher) excitations in the correlated treatment.

For ionization potentials IHFSCC is again in very good agreement with both experimental values available, but underestimates the electron affinities. This may be an effect of both less than ideal starting orbitals, since those from $IO^-$ are used, but may also be due to a lack of triples in the correlation treatment.

Table 5.5: Bond lengths ($R_e$ in Å), harmonic frequencies ($\omega_e$, in $cm^{-1}$), vertical ($T_v$, in $cm^{-1}$) and adiabatic ($T_e$, in $cm^{-1}$) excitation energies, ionization potentials (IP, in eV) and electron affinitues (EA, in eV) calculated with the DC-IHFSCC method for IO.

| Method | Ω | $R_e$ (Å) | $\omega_e$ ($cm^{-1}$) | $T_e$ (eV) | $T_v$ (eV) | EA | IP |
|---|---|---|---|---|---|---|---|
| IHFSCC | X 3/2 | 1.875 | 722 | 0.00 | 0.00 | 2.23 (2.28) | 9.78 (9.72) |
| | a 1/2 | 1.887 | 702 | 2258 | 2258 | | |
| | a 3/2 | 2.095 | 492 | 22259 | 25969 | | |
| CCSD(T) [228] | X 3/2 | 1.872 | 684 | 0.00 | | | |
| | a 1/2 | 1.889 | 651 | 1855 | | | |
| Exp. [229] | X 3/2 | 1.8677 | 681.6 | 0.00 | | | |
| | a 1/2 | 1.887 | 658 | 2091 | | | |
| Exp. [230] | X 3/2 | 1.86762 | 681.7 | 0.00 | | | |
| | a 1/2 | 1.88468 | 645.3 | 2091 | | | |
| | a 3/2 | 2.072 | 514.5 | 21557 | | | |
| Exp. [229, 231] | X 3/2 | | | | | 2.378 ± 0.005 | |
| Exp. [232] | X 3/2 | | | | | | 9.66 ± 0.10 |
| Exp. [233] | X 3/2 | | | | | | 9.735 ± 0.017 |

Following up on halide oxide radical systems, in table 5.6 one can find results for the calculated splitting in the doublet ground state of these species due to spin-orbit coupling from our recent EOM-IP implementation in the four-component framework in DIRAC [1], along with IHFSCC numbers and those obtained by Epifanovsky *et al* [234] for EOM-IP including spin-orbit coupling with the Breit-Pauli Hamiltonian. Here, instead of using the DC Hamiltonian as done for IO, the molecular mean-field X2C (X2Cmmf) Hamiltonian [202], as Our assessment of it in paper C.5 has shown it to yield results nearly indistinguishable from those obtained with the DC Hamiltonian.

[1] at the time of writing, the code can solve for the right eigenvectors of the (T-)similarity transformed Hamiltonian and provide excitation energies for EOM-IP/EE/EA models. We are finalizing the parts necessary to obtain the left eigenvectors for the same models and, with those, calculation transition moments and expectation values for excited states

In table 5.6 we see that the SO splitting greatly increases along the series, arriving at a staggering value of 12000 $cm^{-1}$ for TsO. The importance of the Gaunt interaction decreases along the series, and is very important for the lightest of the molecules (ClO and BrO) in order to provide a good agreement with experiment. As expected, the EOM-IP and IHFSCC yield nearly the same values, and the more marked differences for TsO arise from the fact that we were not able to have as large a model space as for the other molecules without suffering from difficulties with convergence for IHFSCC (the $\sigma$ spinor had to be moved to $P_i$). This observation underscores the appeal of the EOM approaches. Furthermore, we observe that even for light systems we seem to outperform the approximate SO-EOM-IP of [234], due to the use of a more suitable Hamiltonian as well as the basis set limit calculations of Peterson *et al* [228] using single-reference CCSD(T) calculations.

Table 5.6: EOM-IP and IHFSCC Spin-Orbit splitting for the ground state (in $cm^{-1}$) for XO Radicals at their equilibrium geometry. Quadruple zeta basis set weres used for all the calculations. X2Cmmf results from the DCG Hamiltonian are shown in parenthesis.

| | EOM-IP | | IHFSCC | | 2-step EOM-IP [234] | CCSD(T) [228] | Expt. |
|---|---|---|---|---|---|---|---|
| Species | X2Cmmf | | X2Cmmf | | Breit-Pauli | Scalar+SO | |
| ClO | 336 | (315) | 336 | (315) | 306 | 313 | 322 |
| BrO | 1061 | (1034) | 1061 | (1034) | 904 | 856 | 975 |
| IO | 2430 | (2430) | 2395 | (2396) | | 1775 | 2091 |
| AtO | 6392 | (6352) | 6396 | (6356) | | | |
| TsO | 11921 | (11893) | 12112 | (12091) | | | |

In the cases discussed above we have mostly employed a single basis set, usually of triple zeta quality. As it is commonly known in the literature, using such sets (or even quadruple zeta quality ones) means we are still far from having converged our calculations on the one-electron basis dimension. If the properties we are interested in, such as the energies of electronic states, are relatively well separated, these shortcomings usually have no serious consequences. In situations where there are quasi degeneracies, on the other hand, any missing electron correlation can dramatically change our results and with that our interpretation.

One situation where I encoutered this in the study of the electronic structure of $ThF^+$ (see paper C.7), a species that is gaining considerable attention in recent times by being a possible candidate for attempted measurements of the electric dipole moment of the electron [235]. In such experiments polar, charged, diatomic systems are appealing since they offer a much larger enhancement of the expected energy shifts than atomic systems. Furthermore, the species considered so far have in common is an energetically low-lying $^3\Delta$ electronic ground state (in $\Lambda$-$S$ coupling picture). In the fluorides and oxides this state is deeply bound which is an experimental advantage. The magnetic moment in the $\Omega = 1$ component of this term is approximately zero which helps reduce the vulnerability of the experiment to decoherence and systematic errors [236].

Our calculations follow scheme (c) in figure 5.1 and are discussed in detail in paper C.7. There, we have explored a number of active space definitions and found that there was a very marked stabilisation of the $\Omega = 0^+$ state with respect to the $\Omega = 1$ state whenever the $5d$ spinors of thorium were included in the correlation space ($Q$) while the other states were largely unaffected. We also found that the inclusion of $6d, 5f, 7s$ spinors in $P_m$ was necessary to ensure a good balance between two limiting bonding situations in this species: one of mostly ionic bonding, where states coming from the $6d5f$ $Th^{2+}$ configurations would be important, and another for a mostly covalent bonding, where the $6d^27s$ manifold of atomic $Th^+$ would be important for the molecular electronic state. On top of that, we went as far as having the $7p7d8s8p6f$ spinors in $P_i$, in order to try to achieve the accuracy while avoiding intruder states in the calculations.

An important factor for the accuracy in IHFSCC calculations is linked to both the dimension of the model space $P$ (it has been argued [103] that large P spaces may alleviate the need of considering triple or higher excitations in the dynamical correlation treatment due to the inclusion of corresponding excited determinants in the effective Hamiltonian) as well as of the intermediate space $P_m$ (states

with their largest components in $P_m$ are described more accurately than those for which the largest components are in $P_i$ [237]).

We present our results with the most sophisticated model space in table 5.7, along with prior theoretical results from Barker *et al.* [238] and Heaven *et al.* [239]. An important point from our calculations is the importance of using a suitable basis set, as discussed above, in order to capture as much as possible of subtle electron correlation effects: from our results we see that employing a modified quadruple zeta basis set, from which higher angular momentum primitives were removed, results in a complete reversal in the relative order of the $\Omega = 0^+$ and $\Omega = 1$ states. In addition to that, we see that the triple zeta basis set has a hard time capturing these subtle effects and because of that we opted to perform an extrapolation to the basis set limit in order to compare it to experiment, and after such a procedure our best theoretical estimate is in extremely good agreement with the most recent experiment [240] in placing the $^1\Sigma_0^+$ as the ground state and the $^3\Delta_1$ as the first excited state. We note this is in qualitative agreement with CCSD(T) and CCSDT single-reference calculations but at odds with supposedly more accurate CCSDT(Q) single-reference calculations and the prior experimental results [238] which had (incorrectly) assigned the $^3\Delta_1$ as the ground state.

Table 5.7: Electronic spectra of $ThF^+$ obtained with IHFSCC for model space III ($Q$: 4f5s5p5d6s6p ; $P_m$: 5f6d7s: $P_i$: 7p7d8s8p6f) at $R = 1.981$ [Å], obtained with the X2Cmmf Hamiltonian and including SOO interactions. Energies are given in $cm^{-1}$, and E($\infty$) denotes results extrapolated to the basis set limit via a two-point extrapolation formula[241]

| | Electronic state energy | | | | |
|---|---|---|---|---|---|
| Method | $^1\Sigma^+_{0^+}$ | $^3\Delta_1$ | $^3\Delta_2$ | $^3\Delta_3$ | $^3\Pi_{0^-}$ |
| IHFSCC† | 15.25 | 0.00 | 1062.22 | 3149.47 | 4510.50 |
| IHFSCC‡ | 190.85 | 0.00 | 1048.27 | 3156.71 | 4123.14 |
| IHFSCC§ | 0.00 | 108.26 | 1157.05 | 3235.93 | 4415.96 |
| IHFSCC($\infty$) | **318.99** | **0.00** | **1038.94** | **3161.99** | **3841.17** |
| CCSD(T)+SO [238] | 500.7 | 0.0 | 889.5 | 2156.8 | |
| CCSDT+SO [238] | 143.3 | 0.0 | 889.7 | 2157.1 | |
| CCSDT(Q)+SO [238] | 0.0 | 65.5 | 955.3 | 2222.9 | |
| MRCI+Q/SO [239] | 0.0 | 202 | 1047 | 2163 | |
| Exp. [238] | 0.00 | 315.0(5) | 1052.5(5) | 3150(15) | 3395(15) |
| Exp. [240] | 314.0(2) | 0.00 | 1052.5(1,0) | 3149(30) | |

† TZ basis set for Th, [33$s$29$p$20$d$15$f$5$g$1$h$]
‡ QZ basis set, [37$s$34$p$26$d$23$f$9$g$5$h$1$i$].
§ QZ basis set for Th ([37$s$34$p$26$d$23$f$5$g$1$h$]).

## 5.4 Benchmaking approximate methods for small systems

The comparison to experiment above underscores the role of reference method IHFSCC has played over the years in my research. Besides that, and as briefly discussed when presenting the correlated WFT methodologies, the question of how IHFSCC would compare to the single-reference based treatments such as LRCC/EOM-EE is of great importance since the latter are more straightforward to use by avoiding possible convergence issues due to intruder states and do not require the definition of model spaces.

Such a comparison is presented in paper C.3, but since at the time we did not possess a four-component implementation of LRCC/EOM-EE, we used instead [242] a LRCC implementation including scalar relativistic effects, using the spin-free states as a basis for the SO–CI calculation to obtain spin-orbit coupled states. In table 5.8 we summarize our IHFSCC results including spin-orbit, the SO-LRCC of [242] and the SO-CASPT2 results of van Besien [243], corresponding to table 4 in paper C.3.

We note that in our IHFSCC calculations we explored different model and correlation spaces, but in the minimal configuration consisted, respectively, of at least 12 electrons in 24 spinors for $P_m$ (6 occupied, comprising the valance $\sigma, \pi$ system and 12 virtuals, including the $f_\phi, f_\delta, \sigma^*, \pi^*$ spinors), and a $Q$ with a total of 24 electrons correlated (freezing the $5d$).

What we observe in this comparison is that the two coupled cluster approaches are, in effect qualitatively very similar: not only the equilibrium bond lenghts for the electronic states are quite close, but the electronic states below 28000 $cm^{-1}$ are of the same symmetry and show very similar spacings between excited state ($\Delta T_e$) – though the origin of the transitions by nearly 4000 $cm^{-1}$, with LRCC excitations being shifted upwards with respect to the IHFSCC ones. This difference is not really due to the SO–CI treatment of spin-orbit coupling in this case, since as one can see in table 2 of paper C.3, the same behavior is seen for spin-free calculations. Rather, we attributed these differences to the difference in the parametrization of the excited states, since in the $(1,1)$ sector IHFSCC contain terms such as $S^{(0,1)}S^{(1,0)}$[76, 97], which cancel out the disconnected terms which arise from the linear parametrization in LRCC to third order or higher.[76, 97, 244, 245].

In a comparison to the coupled cluster methods, the SO-CASPT2 shows significant differences in both bond lengths and in the relative position of the excited states, though its origin is closer to that of IHFSCC. Since CASPT2 is also size-extensive, the greater similarity of its spectra's origin and that of IHFSCC goes to some length at supporting the attribution of the differences for LRCC and IHFSCC as coming from disconnected terms from LRCC's linear parametrization.

Table 5.8: Equilibrium geometries ($R_e$, in Å) and adiabatic ($\Delta T_e^0$, in $cm^{-1}$) spectrum of the lowest fine structure excited states of $UO_2^{2+}$, computed at the SO-IHFSCC, SO-LR-CCSD[242] and SO-CASPT2.[243] Here $\Delta T_e^1$ (in $cm^{-1}$) denote vertical and adiabatic excitations where the origin of the spectrum is taken to be the first excited state. The minima of the SO-IHFSCC calculations were obtained by interpolating the symmetrical stretching mode by second-order polynomials. Changes in the ordering of the states are marked in italics.

| | SO-IHFSCC | | | SO-LR-CCSD | | | SO-CASPT2 | | |
|---|---|---|---|---|---|---|---|---|---|
| $\Omega$ | $R_e$ | $\Delta T_e^0$ | $\Delta T_e^1$ | $R_e$ | $\Delta T_e^0$ | $\Delta T_e^1$ | $R_e$ | $\Delta T_e^0$ | $\Delta T_e^1$ |
| $0_g^+$ | 1.683 | 0 | -17557 | 1.679 | 0 | -21338 | 1.708 | 0 | -18888 |
| $1_g$ | 1.724 | 17557 | 0 | 1.732 | 21338 | 0 | 1.765 | *18888* | 0 |
| $2_g$ | 1.719 | 17834 | 277 | 1.743 | 21826 | 488 | 1.782 | *17227* | -1661 |
| $3_g$ | 1.725 | 18627 | 1070 | 1.743 | 22361 | 1023 | 1.783 | 18293 | -595 |
| $2_g$ | 1.722 | 20082 | 2525 | 1.736 | 24027 | 2689 | 1.769 | 20911 | 2023 |
| $3_g$ | 1.720 | 23073 | 5516 | 1.735 | 26723 | 5385 | 1.769 | 24026 | 5138 |
| $4_g$ | 1.727 | 23857 | 6300 | 1.743 | 27923 | 6585 | 1.784 | 24190 | 5302 |
| $3_g$ | 1.730 | 26679 | 9122 | 1.750 | 30833 | 9495 | 1.796 | *26446* | 7558 |
| $2_g$ | 1.731 | 28757 | 11200 | 1.749 | *32912* | 11574 | 1.848 | *26500* | 7612 |
| $4_g$ | 1.772 | 29991 | 12434 | 1.798 | *32113* | 10775 | 1.848 | *26259* | 7371 |
| $1_g$ | 1.778 | 30680 | 13123 | 1.795 | *32815* | 11477 | 1.833 | 27923 | 9035 |

The differences between IHFSCC and LRCC presented in 5.8 appear to be in line with the differences between FSCC and EOM-EE that Musial and coworkers[246–248] find when comparing them for light systems: a non-negligible difference in excitation energies, with EOM-EE being higher, but with rather similar spacings between states. Now with the availability of our EOM-EE code, we will be able to revisit this point in a more controlled manner, using the same bases and Hamiltonians.

The performance of CASPT2 relative to methods such as LRCC/EOM-EE or IHFSCC has been evaluated in a number of instances other than those already presented here[92, 94, 249, 250], and these results point to the same overall trends as those discussed here, with CASPT2 yields excitation energies rather close to those obtained with methods that are formally more accurate at a fraction of the computational cost. However, their relative position and the corresponding symmetry classification are often less reliable, which can be attributed to the relatively low accuracy in which dynamical correlation is taken into account.

Another interesting comparison of IHFSCC is with DFT-based approaches employing different density functional approximations (DFAs) for excited states, as done in paper C.4 in the spin-free case and in paper C.5 including spin-orbit for the $UO_2^{2+}$, $NUO^+$ and NUN molecules, which are all isoelectronic. For a detailed discussion of the differences in electronic structure of $NUO^+$ and NUN with respect to $UO_2^2$ the reader should consult section 3.1 of paper C.4.

Before discussing excited states though, it is instructive to look at how the methods describe the system's orbitals. For the occupied a straightforward way to do so is by comparing how the ionization potentials taken simply as the negative of the orbital energies for the DFAs compare to the corresponding IHFSCC results[251]. We note that, while this comparison is strictly valid only for HOMO,[252] our results (selected results from paper C.4 is shown here in table 5.9) indicate that this is a good approximation as found by others[253]. Our results, if quite insensitive to the Hamiltonian used (thus attesting the suitability of the ZORA Hamiltonian for valence states), show significant differences for the different functionals: the best agreement with IHFSCC is found for the SAOP[254] model potential, followed by CAM-B3LYP, B3LYP, BLYP and LDA functionals, with discrepancies on the order of 0.5, 1, 3 and 4 eV, respectively. This behavior can be explained by the ability (or lack thereof) of each of these DFAs to properly mimic the discontinuities in the energy and exchange-correlation (xc) potential that the exact potential should exhibit with the change in particle number, and it is not suprising then that SAOP, constructed to present the correct asymptotic (long-range) behavior, indeed provides the best agreement with the IHFSCC values.

Doing the same for the virtuals, on the other hand, is not as straightfoward due to the very different meaning of the virtual orbital energies[255, 256] when pure (e.g. LDA or (meta)GGAs) or hybrid functionals are employed: for pure functionals the virtual orbital energies are good approximations to the ionization potentials of excited states, whereas in Hartree-Fock they represent approximations to electron affinities. For hybrids they are thus somewhere in between these two values making it difficult to compare these values with the IHFSCC values (that strictly represent electron affinities). A consequence of the difference between Hartree-Fock and pure DFT is that one finds, for the GGA functionals employed here, the low-lying virtuals to be uranium-centered $f_\phi$ and $f_\delta$ orbitals (to which we will observe the transitions from the occupied $\sigma, \pi$ orbitals discussed above), whereas for the hybrids these are often found higher in energy than other orbitals such as the $\sigma$ and $\pi$ antibonding orbitals.

Table 5.9: Comparison of DFT and IHFSCC for the first three ionization potentials (IPs) for $UO_2^{2+}$, NUO+ and NUN (in eV). As these ionized states in the IHFSCC are dominated by contributions from a single orbital and the DFT values are approximated by the negative of the orbital energies, we identify the IPs with the respective orbitals (which range from HOMO-2 to HOMO for DFT). ZORA and DC are Zero Order Regular Approximated and Dirac-Coulomb Hamiltonians, respectively.

| | | NUN | | | $NUO^+$ | | | $UO_2^{2+}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\pi_u$ | $\sigma_g$ | $\sigma_u$ | $\sigma$(U-O) | $\pi$ | $\sigma$(U-N) | $\pi_u$ | $\sigma_g$ | $\sigma_u$ |
| LDA | ZORA | 6.58 | 6.08 | 5.50 | 14.57 | 13.78 | 12.59 | 23.45 | 22.95 | 22.22 |
| | DC | 6.53 | 6.05 | 5.46 | 14.51 | 13.63 | 12.47 | 23.39 | 22.87 | 22.16 |
| BLYP | ZORA | 6.18 | 5.76 | 5.15 | 14.22 | 13.36 | 12.24 | 23.02 | 22.59 | 21.84 |
| | DC | 6.16 | 5.77 | 5.17 | 14.18 | 13.22 | 12.16 | 22.95 | 22.52 | 21.81 |
| B3LYP | ZORA | 7.26 | 6.85 | 6.50 | 15.63 | 14.45 | 13.59 | 24.36 | 24.03 | 23.39 |
| | DC | 7.24 | 6.83 | 6.50 | 15.61 | 14.40 | 13.50 | 24.32 | 23.98 | 23.38 |
| CAM-B3LYP | DC | 9.02 | 8.48 | 8.44 | 17.49 | 16.22 | 15.37 | 26.23 | 25.78 | 25.40 |
| SAOP | ZORA | 10.16 | 9.92 | 9.28 | 18.73 | 17.69 | 16.69 | 27.82 | 27.64 | 26.72 |
| | DC | 10.08 | 9.83 | 9.10 | 18.62 | 17.52 | 16.49 | 27.76 | 27.55 | 26.57 |
| IHFSCC | DC | 10.15 | 9.45 | 9.43 | 18.66 | 17.76 | 16.74 | 27.76 | 27.15 | 27.08 |

By this discussion, and the analysis of the relative errors on the excitation energies for the different DFAs considered in our papers and shown in figure 5.3, we indeed observe that accurately describing the asymptotic behavior of the occupied orbitals is not sufficient to assure a superior quality for excitation energies: our analysis shows that the CAM-B3LYP turns out to be the most reliable

of the functionals considered (small standard deviation and mean close to zero) followed closely by PBE0 and B3LYP, while all GGAs strongly underestimate the excitation energies. The SAOP model potential shows mixed results, with a mean error close to zero but rather large standard deviation on the errors. Though when assessing the effect of spin-orbit coupling we did not consider the same dataset, our results do point to the same conclusions.

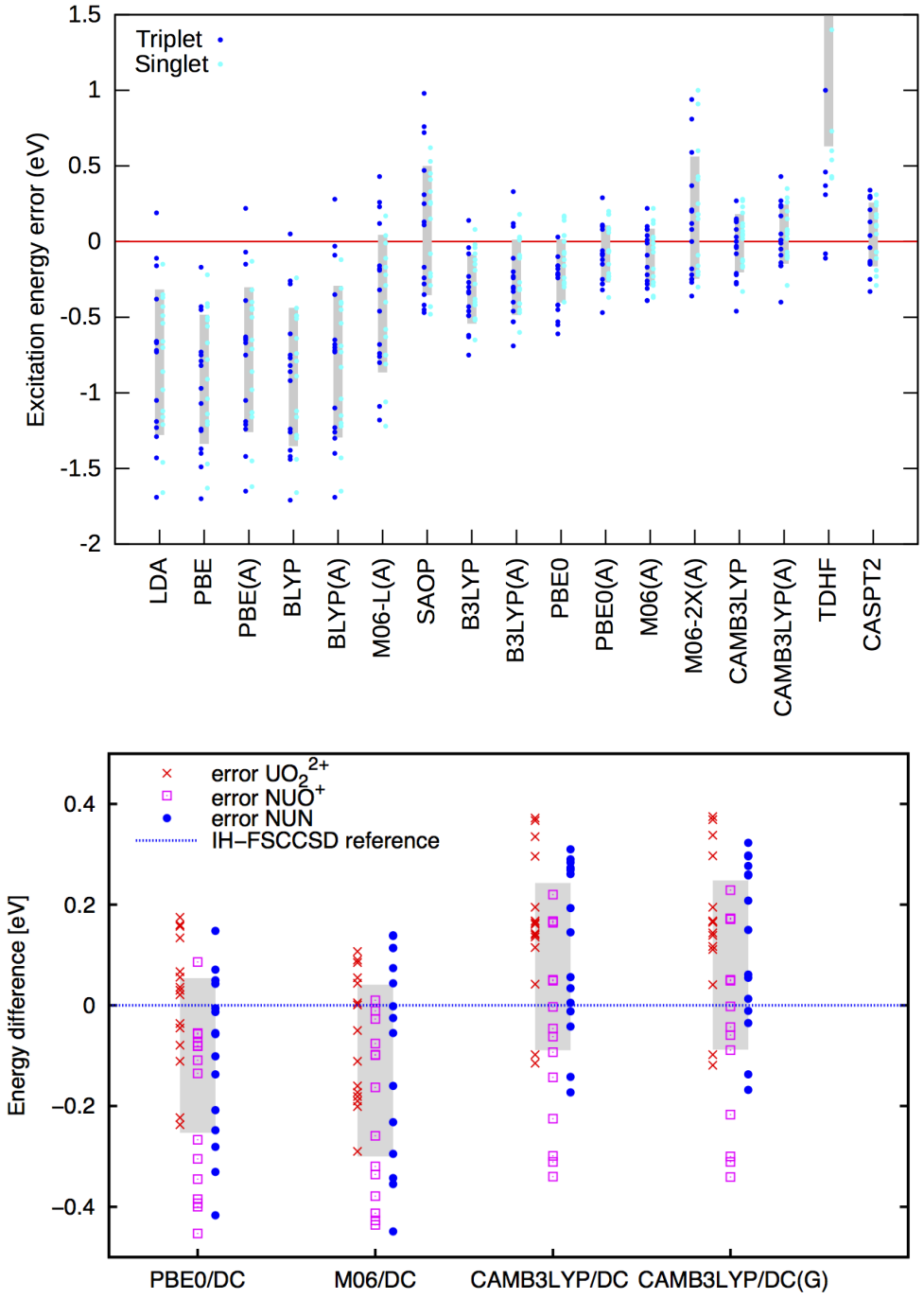



Figure 5.3: Errors with respect to IHFSCC for all excitations and all molecules for spin-free (top) and spin-orbit (bottom) calculations. The gray boxes enclose a range of one sample standard deviation above and below the average error. Dots show individual errors for each excitation energy. The two highest (DFT) states for NUO+ has been left out of the analysis. (A) Evaluated using the ALDA approximation.

Another interesting molecule that is isoelectronic to uranyl is CUO, which was investigated in paper D.4 in order try to better understand its interaction with noble gas matrices, and whether these could induce a change in the ground state from $^1\Sigma^+$ to $^3\Phi$ observed from changes in the vibrational spectra [257] as prior CASPT2 investigations [258] had suggested.

Because of the relatively good performance of hybrid functionals such as PBE0 in TDDFT calculations on the other species isoelectronics to uranyl discussed above, we attempted to use it in order to provide a simple orbital picture of this species and of its intereaction with the noble gas atoms, starting in a spin-free framework, as a complement to IHFSCC calculations. To our surprise these TDDFT results, shown in table 5.10, suffer from triplet instabilities that places the $^1\Phi$ excited state below the $^3\Sigma^+$ state, the latter being the more stabilized the larger the fraction of Hartree-Fock exchange, a result clearly at odds with all correlated WFT calculations.

For the WFT methods, however, we need to make a distinction between CASPT2[258] and coupled cluster approaches (both single reference[259] and IHFSCC), since the former places the spin-free $^3\Phi$ much closer to the $^1\Sigma^+$ than the latter. If we arrange these by the degree of expected combined accuracy for describing dynamical and nondynamical correlation, in the order CASPT2–SRCC–IHFSCC, we would in effect observe a trend of larger $^1\Sigma^+$–$^3\Phi$ separation with increased accuracy.

These TDDFT and WFT results underscore yet again the difficulties in finding a proper description of the balance between exchange and correlation (in a DFT picture) and static and dynamic correlation (in a WFT picture) for actinides, as the both need to provide an equally good description of the strong angular correlation found in the radially localized $\phi$-orbitals as well as for the qualitatively different correlation in the $\sigma$ bonding orbital.

Table 5.10: Spin–free vertical excited-states of the CUO molecule with respect to the $^1\Sigma^+$ ground-state (in eV): a comparison of different methods.

| Symmetry | TD-DFT(PBE) | | | TD-HF | CASPT2[a] | CCSD[b] | CCSD(T) | | IHFSCC |
|---|---|---|---|---|---|---|---|---|---|
| | HF=0% | HF=0.1% | HF=25% | | | | Ref. b | Ref. c | |
| $^3\Sigma^+$ | 6452 | 6452 | 8307 | 13549 | - | - | - | - | 7097 |
| $^3\Phi$ | - | -1048 | -1613 | -4436 | 726 | 6210 | 6694 | 5484 | 10807 |
| $^3\Delta$ | 3307 | 3307 | 3710 | 18953 | 4355 | - | - | - | 10968 |
| $^1\Sigma^+$ | 11856 | 11936 | 14275 | 22985 | - | - | - | - | 12098 |
| $^1\Delta$ | 5726 | 5726 | 7904 | 22824 | 5807 | - | - | - | 12339 |
| $^1\Phi$ | 807 | 807 | 4033 | 11372 | 4758 | - | - | - | 13469 |

*a*) Ref. [258]; *b*) Ref. [259]; *c*) Ref. [260]

The inclusion of spin-orbit coupling brings about a stabilization of about 3000 cm$^{-1}$ for the $\Omega = 2$ component of the $^3\Phi$ spin-free state with respect to the $^3\Sigma^+$ state for all methods. In the case of coupled cluster approaches, electron correlation has placed the $^3\Phi$ state sufficiently higher than the $^3\Sigma^+$ so that no change in the ground state can occur. On the other hand, for CASPT2 which places the $^3\Phi$ higher than the $^3\Sigma^+$ by less than a 1000 cm$^{-1}$, the inclusion of spin-orbit coupling does make the change in ground state possible. That said, taking into consideration the assesments of different methods for systems with reliable experimental date in the gas-phase, it is hard to be confident that CASPT2 got the relative energies of these two states right and the coupled cluster approaches wrong.

A comparison of IHFSCC, second-order multiconfigurational approaches and TDDFT in the study of $IO^+$ and $AtO^+$, shown in paper C.6, also uncovered the presence of triplet instabilities for not only DFAs including fractions of Hartree-Fock exchange but also for the SAOP model potential and the M06-L metaGGA. These species, which in a spin-free framework would have a $^3\Sigma^+$ ground state (with two unpaired electrons in the antibonding $\pi$ orbitals, hae been found in the presence of spin-orbit coupling to have the $\Omega = 0$ component of the triplet sufficiently stabilized over the $\Omega = \pm 1$ components so that both can be considered to be relativistically closed shell species as can be seen from the IHFSCC results in table 5.11. SO-NEVPT2 calculations confirm this finding, while at the same time underscoring the importance of including relaxation effects on the SO–CI treatment $AtO^+$:

the approach that includes such effects (uc–SOCI) shows a very good agreement with the IHFSCC results with the exception of the b $0^-$ state which exhibits some contribution from the first $^3\Pi$ state that is absent from the IHFSCCSD calculation. The c–SOCI approach, on the other hand, shows and increased deterioration for the states above 10000 $cm^{-1}$.

The TDDFT calculations, however, show strongly stabilized $\Omega = \pm 1$ components with respect to the $^3\Sigma^+$ state, with a splitting between these states which ranges from 4000 $cm^{-1}$ for SAOP and M06-L (no Hartree-Fock exchange) to over 6000 $cm^{-1}$ for M06-2X (which contains over 50% of Hartree-Fock exchange), and show an even more catastrophic failure for the latter if the other excited states are considered. The pragmatic approach of invoking the Tamm-Dancoff approximation greatly improves the situation for SAOP and M06-L, and we recover results which are quite in line with the correlated calculations' results for $IO^+$ and $AtO^+$ – with the exception of the $\Omega = \pm 1$ states of $IO^+$ which are still found to be lower than the $\Omega = 0$ state and signal a persistent problem in the description of unpaired spins for these DFAs. For hybrid functionals, on the other hand, agreement with correlated calculations is still poor. The same trends were found for a complex $AtO^+$-water, but I have not pursued the matter further.

All in all, this study has served as a warning for those who interested in employing (TD)DFT to investigate astatine in solution [224–227], by underscoring the importance of a proper benchmark of approximate methods and the identification of pathological behavior for the DFAs before using them in calculations for complex systems.

## 5.5 Towards large systems with TDDFT

In 2015 I became involved in a collaboration with Valérie Vallet and Richard Wilson, an experimentalist from Argonne National Laboratories, aiming at understanding the electronic structure of the $UO_2Cl_2(phen)_2$ complex. The $UO_2Cl_2(phen)_2$ complex is an interesting one since it is one of the few molecular systems containing the uranyl moiety presenting a bent structure, and the questions our investigation aimed at addressing were, first, what was the nature of the low-lying excited states and, second, to what extend the bending of the O-U-O was due steric effects from the bulky phenantroline ligands (structure shown in figure 5.4).

Due to its size, it is clear that a wavefunction treatment of the complex is not (yet) feasible, and we resorted to DFT calculations. Here, the benchmarks above were quite useful in weeding out DFAs that were not sufficiently reliable for describing excited states of species containing actinyls, and we selected the PBE0 functional, which provides an accuracy close to that of CAM-B3LYP but at a smaller computational cost, for the SO-ZORA TDDFT calculations with the ADF [261] code. Given the difficulties with triplet instabilities encountered for other systems we opted here to invoke the TDA approximation.

We started out by an investigation of the ground state, and found that the $UO_2Cl_2$ subunit takes up a bent structure even in the absence of the phenantroline ligands, in line with the findings of Su *et al* [262] for $UO_2Cl_2$ in an argon matrix. For the excited states, on the other hand, there is a marked effect due to the phenantroline ligands since the virtual spinors located on one of the phenantroline ligands can now mix in with the spinors related to the $\phi$ and $\delta$ in the idealized (linear) uranyl subunit – which, since the O-U-O bond is bent, can themselves mix with the chlorine spinors.

In table 5.12, we present the calculated excitation energies. We observe the low-lying transitions for $UO_2Cl_2(phen)_2$ are mostly between the HOMO and LUMO+2 and LUMO+3 spinors shown in figure 5.4, and they show a marked shift upwards with respect to the spectrum of the $UO_2Cl_2$ subunit, which is essentially made up of transitions from the occupied spinor that greatly resembles the HOMO of figure 5.4 to the spinors related to the $\phi$ and $\delta$ ones. Both in the case of $UO_2Cl_2$ and $UO_2Cl_2(phen)_2$, we see that the first and second excited states are made up of transitions with strong $\Phi$ character. However, we also see that these states are much too close for us to be able to inequivocally affirm which is the lowest, given the limits in accuract of (TD)DFT calculations.

Table 5.11: Electronic excitation energies (in $cm^{-1}$) for different correlation and SOC treatments for $IO^+$ and $AtO^+$ calculated at R($IO^+$) = 1.806 Å and R($AtO^+$) = 1.903 Å respectively. Here (u)c–SOCI denote (un)contracted SOCI, and the TDDFT calculations both with and without (results in parenthesis) employing the Tamm-Dancoff approximation (TDA).

| | | | QD-NEVPT2 | | TDDFT | | | |
|---|---|---|---|---|---|---|---|---|
| Species | Ω | IHFSCC | c–SOCI | uc–SOCI | SAOP | M06-L | M06 | M06-2X |
| $IO^+$ | X $0^+$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | a 1 | 1049 | 726 | 726 | -1532 | -1694 | -2662 | -3629 |
| | | | | | (-4839) | (-4920) | (-5807) | (-6936) |
| | a 2 | 5807 | 5485 | 5565 | 4759 | 4597 | 4113 | 3871 |
| | | | | | (4194) | (4113) | (-6291) | (-8308) |
| | a $0^+$ | 11211 | 10808 | 10888 | 30407[1] | 31859[1] | 29601[1] | 30649[1] |
| | | | | | (30165[1]) | (31698[1]) | (29036[1]) | (29439[1]) |
| | b $0^-$ | 19519 | 20486 | 20325 | 21777 | 23309 | 15647 | 10163 |
| | | | | | (21051) | (22664) | (8791) | (-17825) |
| | a 3 | 20003 | 21132 | 20728 | 21858 | 23390 | 15566 | 9921 |
| | | | | | (21212) | (22745) | (8308) | (-18712) |
| $AtO^+$ | X $0^+$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | a 1 | 5162 | 4355 | 3952 | 3791 | 3307 | 2339 | 2016 |
| | | | | | (-4436) | (-3871) | (-6130) | (-8388) |
| | a 2 | 8469 | 7985 | 7501 | 9033 | 8791 | 8146 | 8227 |
| | | | | | (8630) | (8469) | (3710) | (-5565) |
| | a $0^+$ | 16373 | 17905 | 15970 | 30488[1] | 31456[1] | 28552[1] | 29842[1] |
| | | | | | (30165[1]) | (31214[1]) | (28068[1]) | (27342[1]) |
| | a 3 | 16696 | 19922 | 16131 | 19438 | 20486 | 13389 | 7098 |
| | | | | | (18793) | (19841) | (5162) | (-19035) |
| | b $0^-$ | 22342 | 18712 | 19761 | 19357 | 20406 | 13389 | 7259 |
| | | | | | (18712) | (19761) | (5162) | (-18712) |

1. States with singly excited character, whereas WFT methods indicate doubly excited character.

Table 5.12: SO-TDDFT/TDA vertical transition energies ($\Delta E$, in $cm^{-1}$) for the $UO_2Cl_2$, and $UO_2Cl_2(phen)_2$ computed at the crystal geometries. Here $\sigma_u, \pi_u, \phi, \delta$ represent spinors related to those of linear uranyl Cl the valence spinors of the chloride ligands that combine with those of uranyl, and $\pi^*(L)$ the *pi* system on the phenantroline ligand roughly on the same plane as the O-U-O bond.

| Species | $\Delta E$ | nature of transition |
|---|---|---|
| UO2Cl2 | 14842 | $\sigma_u(19\%) + \mathrm{Cl}(72\%) \rightarrow \phi$ |
| | 14993 | $\sigma_u(19\%) + \mathrm{Cl}(74\%) \rightarrow \phi$ |
| | 15492 | $\sigma_u(15\%) + \mathrm{Cl}(78\%) \rightarrow \phi$ |
| | 15533 | $\sigma_u(16\%) + \mathrm{Cl}(73\%) \rightarrow \phi$ |
| | 16463 | $\sigma_u(20\%) + \mathrm{Cl}(63\%) \rightarrow \delta$ |
| | 16530 | $\sigma_u(28\%) + \mathrm{Cl}(63\%) \rightarrow \delta$ |
| UO2Cl2phen2 | 18349 | $\sigma_u + \mathrm{Cl} \rightarrow (\phi, \pi^*(L))$ |
| | 18448 | $\sigma_u + \mathrm{Cl} \rightarrow (\phi, \pi^*(L))$ |
| | 19026 | $\sigma_u + \mathrm{Cl} \rightarrow (\delta + \pi^*(L))(27\%) + (\phi + \pi^*(L))(30\%)$ |
| | 19180 | $\sigma_u + \mathrm{Cl} \rightarrow (\delta + \pi^*(L))(42\%) + (\phi + \pi^*(L))(14\%)$ |
| | 19443 | $\sigma_u + \mathrm{Cl} \rightarrow (\phi + \pi^*(L))(45\%) + (\delta + \pi^*(L))(15\%)$ |
| | 19475 | $\sigma_u + \mathrm{Cl} \rightarrow (\phi + \pi^*(L))(59\%) + (\delta + \pi^*(L))(7\%)$ |

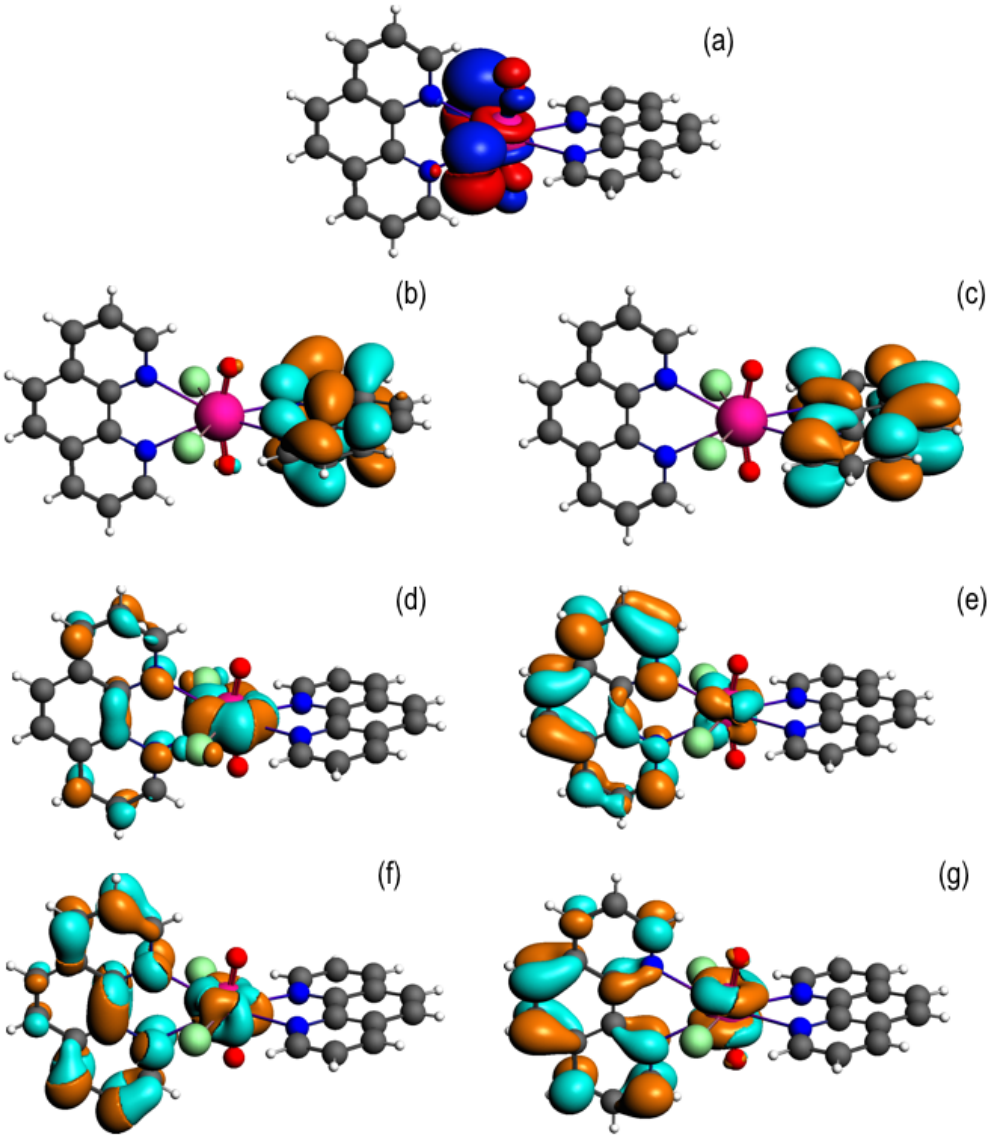



Figure 5.4: $UO_2Cl_2(phen)_2$ Molecular spinors: (a) HOMO; (b) LUMO; (c) LUMO+1; (d) LUMO+2; (e) LUMO+3; (f) LUMO+4; (g) LUMO+5

# Chapter 6

# Frozen Density Embedding Methods

Having discussed the essentials of electronic structure methods in chapter 5 in this chapter we discuss the formal and practical aspects of the frozen density embedding approach, some of my contributions to its development and its application to heavy elements, based on the papers presented in full in appendix D.

## 6.1 Basic Ideas and Exact Theory

*This section partially reproduces the contents of section 3 of* ***ASP Gomes, CR Jacob, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem., 2012, 108, 222–277***

All quantum-chemical methods for excited states discussed in the previous section show a rather steep increase of the computational effort with the size of the system. This is particularly the case for wavefunction based methods where there are well-defined hierarchies that allow for a systematic improvement of the calculation. Thus, their applicability is limited to comparably small molecules and a treatment of electronic excitations in complex chemical systems remains a challenge. However, if only local excitations are of interest, additional simplifications can be introduced.

The simplest possibility for exploiting this locality is the truncation of the full system to a smaller model. Such a model contains only the part of the system where the local excitations of interest take place and of its environment. For the case of a chromophore molecule in solution the definition of such cluster model is rather straightforward. In the simplest case, one starts from an isolated chromophore, which is then progressively surrounded by solvent molecules up to some limit e.g. to include a complete solvation shell. Similar constructions can be used for chromophores in protein environments by only including specific amino acid residues that are close to the chromophore. Also for treating impurities in solids truncated cluster models can be set up by including only atoms within a certain distance from the impurity. Even though such a truncation of the full system may be a rather crude approximation, the results obtained with such cluster models will eventually converge towards a full calculation if the size of the model system is systematically enlarged. However, this convergence can be rather slow and, therefore, cluster models that provide sufficiently accurate results will often be rather large and contain hundreds or thousands of atoms. Thus, a full quantum-chemical treatment of sufficiently large truncated models is rarely possible with accurate quantum-chemical methods.

Embedding methods follow an intermediate strategy between a full quantum-chemical treatment and the use of small truncated model systems. They still restrict the accurate treatment to a small subsystem of interest, but instead of neglecting the environment of this model system, it is included in a more approximate manner. As with truncated model systems, the results obtained with such embedding schemes will eventually converge towards those of a full treatment when enlarging the

size of the explicitly treated subsystem. However, since the environment is always included, this convergence should be much faster, which makes it possible to restrict the size of the active subsystem considerably if only local excitations or other local spectroscopic properties are of interest.

Formally, such embedding approaches can be formulated as an exact theory as outlined in the following subsection. From this exact embedding theory one always obtains the same results as with a full quantum-chemical treatment, independent of the chosen size of the active subsystem. However, such a treatment is not suitable for practical applications since its computational cost would be comparable to (or even larger than) the one of a full quantum-chemical treatment. Therefore, approximations have to be applied for the description of the environment. The quality of these approximations determines how fast the results of embedding calculations converge with the size of the explicitly treated subsystem. More accurate embedding methodologies will allow for the use of a smaller active subsystem, which in turn makes it possible to apply more sophisticated and more accurate quantum-chemical methods for the description of the local excitations of interest. It should also be noted that embedding approaches also simplify the interpretation of the computational results significantly, since information from the parts of the system not treated explicitly will, in effect, be filtered away by construction.

## Frozen-Density Embedding Theory

Embedding approaches start from a partitioning of the total system into a subsystem of interest (subsystem I in the following) and its environment (subsystem II). The frozen-density embedding (FDE) theory formulated by Wesolowski and Warshel[263] — following earlier work of Senatore and Subbaswamy[264, 265] and of Cortona[266] — provides a formally exact theoretical framework for introducing such a partitioning. It is based on the formally exact DFT (i.e., considering exact density functionals) and uses the electron density of the total system $\rho_{\mathrm{tot}}(\boldsymbol{r})$ as its starting point. This total density is partitioned into the electron densities of the active subsystem, $\rho_{\mathrm{I}}(\boldsymbol{r})$, and of the environment, $\rho_{\mathrm{II}}(\boldsymbol{r})$, i.e.,

$$\rho_{\mathrm{tot}}(\boldsymbol{r}) = \rho_{\mathrm{I}}(\boldsymbol{r}) + \rho_{\mathrm{II}}(\boldsymbol{r}). \tag{6.1}$$

These subsystem densities are allowed to overlap. In the following, we will always assume that the subsystem densities $\rho_{\mathrm{I}}(\boldsymbol{r})$ and $\rho_{\mathrm{II}}(\boldsymbol{r})$ integrate to an integer number of electrons. However, the theory can be generalized to subsystems with fractional electron numbers.[267] In addition to the electron density, the nuclear charges are also partitioned. These divisions of the density and the nuclei define the two subsystems I and II. The environment density $\rho_{\mathrm{II}}$ could be further partitioned into an arbitrary number of subsystems[266, 268, 269] as

$$\rho_{\mathrm{II}}(\boldsymbol{r}) = \sum_n \rho_{\mathrm{II}}^{(n)}. \tag{6.2}$$

This is particularly useful for formulating subsystem approaches, in which a large system is partitioned into many smaller subsystems, that are then treated on an equal footing. For a recent review of such fragment-based methods in quantum chemistry, see ref. [270]. As our focus here will be on methods that single out a specific subsystem of interest, the discussion in the following will be restricted to two subsystems, i.e., the densities of all but one subsystem will be collected into a single environment density $\rho_{\mathrm{II}}$.

### Interaction energy.

Using this partitioning into subsystems, the DFT total energy can be expressed as a functional of the two subsystem densities $\rho_{\mathrm{I}}$ and $\rho_{\mathrm{II}}$,

$$\begin{aligned} E_{\mathrm{tot}} &= E[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] \\ &= \int \big(\rho_{\mathrm{I}}(\boldsymbol{r}) + \rho_{\mathrm{II}}(\boldsymbol{r})\big)\big(v_{\mathrm{I}}^{\mathrm{nuc}}(\boldsymbol{r}) + v_{\mathrm{II}}^{\mathrm{nuc}}(\boldsymbol{r})\big)\,\mathrm{d}^3r + \frac{1}{2}\int \frac{\big(\rho_{\mathrm{I}}(\boldsymbol{r}) + \rho_{\mathrm{II}}(\boldsymbol{r})\big)\big(\rho_{\mathrm{I}}(\boldsymbol{r'}) + \rho_{\mathrm{II}}(\boldsymbol{r'})\big)}{|\boldsymbol{r} - \boldsymbol{r'}|}\,\mathrm{d}^3r\,\mathrm{d}^3r' \\ &+ E_{\mathrm{xc}}[\rho_{\mathrm{I}}] + E_{\mathrm{xc}}[\rho_{\mathrm{II}}] + E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] + T_s[\rho_{\mathrm{I}}] + T_s[\rho_{\mathrm{II}}] + T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] + E_{\mathrm{NN}}, \end{aligned} \tag{6.3}$$

where $E_{\mathrm{NN}}$ is the nuclear repulsion energy, $v_{\mathrm{I}}^{\mathrm{nuc}}$ and $v_{\mathrm{II}}^{\mathrm{nuc}}$ are the electrostatic potentials of the nuclei in subsystems I and II, respectively, $E_{xc}[\rho]$ is the exchange–correlation energy functional, and $T_s[\rho]$ is the kinetic energy of a reference system of noninteracting electrons with density $\rho$. The nonadditive exchange–correlation and kinetic energies are defined as

$$E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}] = E_{\mathrm{xc}}[\rho_{\mathrm{I}}+\rho_{\mathrm{II}}] - E_{\mathrm{xc}}[\rho_{\mathrm{I}}] - E_{\mathrm{xc}}[\rho_{\mathrm{II}}] \tag{6.4}$$

and

$$T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}] = T_s[\rho_{\mathrm{I}}+\rho_{\mathrm{II}}] - T_s[\rho_{\mathrm{I}}] - T_s[\rho_{\mathrm{II}}], \tag{6.5}$$

respectively.

The total energy given in Eqn. (6.3) can be partitioned as,

$$E_{\mathrm{tot}} = E_{\mathrm{I}} + E_{\mathrm{II}} + E_{\mathrm{int}}, \tag{6.6}$$

into the energies of the two individual subsystems ($n = \mathrm{I}, \mathrm{II}$)

$$\begin{aligned} E_{\mathrm{n}} = E[\rho_{\mathrm{n}}] = E_{\mathrm{NN}}^{(n)} &+ \int \rho_n(\boldsymbol{r})\, v_{\mathrm{nuc}}^{(n)}(\boldsymbol{r})\, \mathrm{d}^3r + \frac{1}{2}\int \frac{\rho_n(\boldsymbol{r})\rho_n(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}\, \mathrm{d}^3r\, \mathrm{d}^3r' \\ &+ E_{\mathrm{xc}}[\rho_n] + T_s[\rho_n], \end{aligned} \tag{6.7}$$

and the interaction energy

$$\begin{aligned} E_{\mathrm{int}} = E_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}] = E_{\mathrm{NN}}^{(\mathrm{int})} &+ \int \rho_{\mathrm{I}}(\boldsymbol{r}) v_{\mathrm{nuc}}^{\mathrm{II}}(\boldsymbol{r})\, \mathrm{d}^3r + \int \rho_{\mathrm{II}}(\boldsymbol{r}) v_{\mathrm{nuc}}^{\mathrm{I}}(\boldsymbol{r})\, \mathrm{d}^3r \\ &+ \int \frac{\rho_{\mathrm{I}}(\boldsymbol{r})\rho_{\mathrm{II}}(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}\, \mathrm{d}^3r\, \mathrm{d}^3r' + E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}] + T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}], \end{aligned} \tag{6.8}$$

where the nuclear repulsion energy is partitioned into the repulsion among nuclei in the same subsystem $E_{\mathrm{NN}}^{(n)}$ and between those in different subsystems $E_{\mathrm{NN}}^{(\mathrm{int})}$.

Eqn. (6.8) provides an exact expression for the interaction energy between two subsystems with fixed electron densities $\rho_{\mathrm{I}}$ and $\rho_{\mathrm{II}}$. The first four terms add up to the classical electrostatic interaction energy between the nuclei and electron densities of the two subsystems. In addition, the nonadditive exchange–correlation energy $E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$ and the nonadditive kinetic energy $T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$ account for the non-classical contributions to the interaction energy. While the classical terms can be calculated directly for any two subsystems once the nuclear charges and positions as well as the subsystem electron densities are known, the evaluation of the non-classical contributions requires the knowledge of the exchange–correlation and kinetic-energy functionals, $E_{\mathrm{xc}}[\rho]$ and $T_s[\rho]$. Even though these are not know, Eqn. (6.8) provides a useful starting point for the development of embedding methods. We note that its applicability is not limited to DFT calculations since an electron density can always be defined within any theoretical framework and then used to evaluate an interaction energy within FDE theory.

**Embedding potential.**

So far, the electron densities of the two subsystems were kept fixed. However, the total electron density of two *interacting* subsystems will not be equal to the sum of the densities of the *isolated* subsystems. Therefore, the subsystem electron densities change when the two subsystems interact and the presence of an environment, $\rho_{\mathrm{II}}(\boldsymbol{r})$, modifies the electron density of the active subsystem, $\rho_{\mathrm{I}}(\boldsymbol{r})$. To account for this, the environment has to be included in the quantum-chemical description of the active subsystem. This is possible both for a description of the active subsystem with KS-DFT and for a wavefunction based treatment.

For the case of a KS-DFT description, the density of the active subsystem I can be obtained from the KS orbitals $\{\phi_i^{\mathrm{I}}\}$ as $\rho_{\mathrm{I}}(\boldsymbol{r}) = \sum_i |\phi_i^{\mathrm{I}}(\boldsymbol{r})|^2$. Note that in this case the noninteracting kinetic energy $T_s[\rho_{\mathrm{I}}]$ can also be calculated directly from the KS orbitals. For a given frozen electron density

$\rho_{\mathrm{II}}(\boldsymbol{r})$ in subsystem II, the KS orbitals (and the electron density) of the active subsystem I can then be determined by minimizing the total energy given in Eqn. (6.3) with respect to $\rho_{\mathrm{I}}$, while keeping $\rho_{\mathrm{II}}$ frozen. Performing this minimization under the constraint that the number of electrons $N_{\mathrm{I}}$ in subsystem I is conserved leads to a set of equations for the KS orbitals of subsystem I,

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{KS}}[\rho_{\mathrm{I}}](\boldsymbol{r}) + v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}](\boldsymbol{r})\right] \phi^{\mathrm{I}}_i(\boldsymbol{r}) = \varepsilon^{\mathrm{I}}_i \phi^{\mathrm{I}}_i(\boldsymbol{r}), \tag{6.9}$$

where $v_{\mathrm{KS}}[\rho_{\mathrm{I}}](\boldsymbol{r})$ is the KS effective potential of the isolated subsystem I containing the usual terms of the nuclear potential, the Coulomb potential of the electrons, and the exchange–correlation potential,

$$v_{\mathrm{KS}}[\rho_{\mathrm{I}}](\boldsymbol{r}) = \frac{\delta E[\rho_{\mathrm{I}}]}{\delta \rho_{\mathrm{I}}(\boldsymbol{r})} = v^{\mathrm{I}}_{\mathrm{nuc}}(\boldsymbol{r}) + \int \frac{\rho_{\mathrm{I}}(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} \mathrm{d}^3 r' + \frac{\delta E_{\mathrm{xc}}[\rho_{\mathrm{I}}]}{\delta \rho_{\mathrm{I}}(\boldsymbol{r})}, \tag{6.10}$$

and the effective embedding potential $v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}](\boldsymbol{r})$ describes the interaction of subsystem I with the frozen density and nuclei of subsystem II,

$$\begin{aligned} v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}](\boldsymbol{r}) = \frac{\delta E_{\mathrm{int}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}]}{\delta \rho_{\mathrm{I}}(\boldsymbol{r})} &= v^{\mathrm{II}}_{\mathrm{nuc}}(\boldsymbol{r}) + \int \frac{\rho_{\mathrm{II}}(\boldsymbol{r'})}{|\boldsymbol{r} - \boldsymbol{r'}|} \mathrm{d}^3 r' \\ &+ \left.\frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\boldsymbol{r})}\right|_{\rho_{\mathrm{tot}}} - \left.\frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\boldsymbol{r})}\right|_{\rho_{\mathrm{I}}} + \left.\frac{\delta T_s[\rho]}{\delta \rho(\boldsymbol{r})}\right|_{\rho_{\mathrm{tot}}} - \left.\frac{\delta T_s[\rho]}{\delta \rho(\boldsymbol{r})}\right|_{\rho_{\mathrm{I}}}. \end{aligned} \tag{6.11}$$

This embedding potential accounts for the presence of the frozen environment when determining the electron density of the active subsystem with KS-DFT. Note that the embedding potential is a local potential that depends only on the electron densities of the two subsystems.

The first two terms of the embedding potential of Eqn. (6.11) describe the classical electrostatic potential of the nuclei and of the electrons in the frozen environment. In addition, the embedding potential also contains an exchange–correlation component and a kinetic-energy component. These account for the non-classical contributions, such as the Pauli (exchange) repulsion of the electrons in the frozen subsystem and chemical bonding (i.e., orbital interactions) between the subsystems. While the electrostatic part of the embedding potential can be evaluated directly for given subsystem densities, this is not possible for the exchange–correlation and kinetic energy parts, as these require the knowledge of the corresponding exact functionals.

The same embedding potential can also be derived for the case that a wavefunction based description is used for the active subsystem.[271, 272] In this case a wavefunction $\Psi_I(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2, \dots)$ is used to represent the electron density $\rho_{\mathrm{I}}(\boldsymbol{r})$ of subsystem I. By using that

$$E_{\mathrm{I}} = E[\rho_{\mathrm{I}}] = E^{(I)}_{\mathrm{NN}} + \langle \Psi_{\mathrm{I}} | \hat{T} + \hat{V}^{\mathrm{I}}_{\mathrm{nuc}} + \hat{V}_{\mathrm{ee}} | \Psi_{\mathrm{I}} \rangle, \tag{6.12}$$

where $\hat{T}$, $\hat{V}^{\mathrm{I}}_{\mathrm{nuc}}$, and $\hat{V}_{\mathrm{ee}}$ are the operators of the kinetic energy, the electron–nuclear attraction energy, and of the electron–electron interaction, respectively, the total energy of Eqn. (6.3) and (6.6) can be rewritten as

$$E_{\mathrm{tot}} = E[\Psi_{\mathrm{I}}, \rho_{\mathrm{II}}] = E^{(I)}_{\mathrm{NN}} + \langle \Psi_{\mathrm{I}} | \hat{T} + \hat{V}^{\mathrm{I}}_{\mathrm{nuc}} + \hat{V}_{\mathrm{ee}} | \Psi_{\mathrm{I}} \rangle + E_{\mathrm{int}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] + E_{\mathrm{II}}, \tag{6.13}$$

where $E_{\mathrm{int}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}]$ and $E_{\mathrm{II}} = E[\rho_{\mathrm{II}}]$ are the interaction energy and the energy of subsystem II as defined in Eqns. (6.8) and (6.7), respectively.

The wavefunction describing subsystem I in the presence of the frozen density $\rho_{\mathrm{II}}(\boldsymbol{r})$ can then be obtained by minimizing this total energy functional with respect to $\Psi_{\mathrm{I}}$ while keeping the electron density $\rho_{\mathrm{II}}$ of the environment frozen, under the constraint that the number of electron $N_{\mathrm{I}}$ in subsystem I is conserved. This leads to the condition,

$$\begin{aligned} 0 &= \left(\hat{T} + \hat{V}^{\mathrm{I}}_{\mathrm{nuc}} + \hat{V}_{\mathrm{ee}}\right)\Psi_{\mathrm{I}} + \int \frac{\delta E_{\mathrm{int}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}]}{\delta \rho_{\mathrm{I}}(\boldsymbol{r'})} \frac{\delta \rho_{\mathrm{I}}(\boldsymbol{r'})}{\delta \Psi_{\mathrm{I}}} \, \mathrm{d}^3 r' - \lambda \, \Psi_{\mathrm{I}} \\ &= \left[\hat{T} + \hat{V}^{\mathrm{I}}_{\mathrm{nuc}} + \hat{V}_{\mathrm{ee}} + \hat{V}^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}]\right] \Psi_{\mathrm{I}} - \lambda \, \Psi_{\mathrm{I}}, \end{aligned} \tag{6.14}$$

with the embedding operator

$$\hat{V}^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] = \sum_i v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}](\boldsymbol{r_i}), \tag{6.15}$$

that is, the wavefunction of subsystem I in the presence of the frozen density $\rho_{\mathrm{II}}(\boldsymbol{r})$ can be determined by solving an eigenvalue equation, in which the embedding potential of Eqn. (6.11) enters as an additional one-electron operator. However, the eigenvalue $\lambda$ in this embedded Schrödinger equation does not correspond to an energy. Instead, the energy has to be evaluated using Eqn. (6.13) once the embedded wavefunction $\Psi_{\mathrm{I}}$ has been determined.

For solving this embedded Schrödinger equation, the common approximations of wavefunction based quantum chemistry can be applied. Note that the derivation given here differs from the one in ref. [272], where an approximate wavefunction of subsystem I was introduced before performing the energy minimization. In this case, the embedding potential contains an additional term correcting for the difference between the approximate and the exact wavefunctions. However, it can be argued that a correction for deficiencies of an employed wavefunction approximation should not be contained in the embedding potential.[273, 274] Therefore, the derivation given here avoids this correction by introducing an approximate wavefunction only at a later stage.

**Polarization of the environment.**

In an exact embedding calculation using a frozen environment density $\rho_{\mathrm{II}}$ the electron density $\rho_{\mathrm{I}}$ of the active subsystem should be determined such that the total electron density $\rho_{\mathrm{tot}} = \rho_{\mathrm{I}} + \rho_{\mathrm{II}}$ is identical to the one obtained from a calculation of the full system. This can be achieved by minimizing the total energy with respect to the density (or wavefunction) of the active subsystem, and leads to the local embedding potential derived above.

However, such an agreement with the results of a full calculation is only possible if the frozen density fulfills certain conditions.[275, 276] In particular, the frozen density $\rho_{\mathrm{II}}$ has to be smaller than or equal to the correct total density $\rho_{\mathrm{tot}}$ at every point in space, i.e., $\rho_{\mathrm{II}}(\boldsymbol{r}) \leq \rho_{\mathrm{tot}}(\boldsymbol{r})$. Otherwise, the complementary density of the active system would have to be negative, which is not possible. In addition, this complementary density $\rho_{\mathrm{tot}} - \rho_{\mathrm{II}}$ has to be noninteracting $v_s$-representable in the case of a KS-DFT description for the active system or interacting $v$-representable in the case of a wavefunction based treatment.

In particular the first condition is usually not fulfilled for most approximate frozen densities. Usually, these will be too small in some regions and too large in others. Consequently, the application of the embedding potential of Eqn. (6.11) does not lead to the exact total density with such choices for $\rho_{\mathrm{II}}$. This problem can be alleviated by switching from an embedding method to a subsystem approach in which both the densities of subsystem I and II are determined. That is, the densities of both subsystems are determined separately, but in each case the (frozen) density of the other subsystem is taken into account. When using KS-DFT, this can be formulated as a set of coupled equations for the KS orbitals of the two subsystems,

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{KS}}[\rho_{\mathrm{I}}](\boldsymbol{r}) + v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}](\boldsymbol{r})\right] \phi^{\mathrm{I}}_i(\boldsymbol{r}) = \varepsilon^{\mathrm{I}}_i \phi^{\mathrm{I}}_i(\boldsymbol{r}) \tag{6.16}$$

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{KS}}[\rho_{\mathrm{II}}](\boldsymbol{r}) + v^{\mathrm{II}}_{\mathrm{emb}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}](\boldsymbol{r})\right] \phi^{\mathrm{II}}_i(\boldsymbol{r}) = \varepsilon^{\mathrm{II}}_i \phi^{\mathrm{II}}_i(\boldsymbol{r}) \tag{6.17}$$

Note that because of the different roles of the active and frozen densities in Eqn. (6.11), the embedding potentials in these two equations differ.

The simplest strategy for solving these coupled equations for the two subsystems is through so-called freeze-and-thaw iterations [277]. First the density of subsystem I is determined in the presence of an approximate frozen density for subsystem II. Subsequently, the roles of the two subsystems are interchanged and the density calculated for subsystem I in the previous step is now frozen,

whereas an updated density is determined for subsystem II. This is repeated iteratively until convergence is reached. Alternatively, the two sets of equations (Eqns. (6.16) and (6.17)) can be solved simultaneously.[268] Note that the resulting partitioning into subsystems is not unique, because density can be moved between the two subsystems without changing the total electron density. However, a unique partitioning can be obtained when requiring that both subsystems share the same embedding potential, i.e., that $v^{\mathrm{I}}_{\mathrm{emb}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}] = v^{\mathrm{II}}_{\mathrm{emb}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$. [267, 278]

When focussing on one subsystem of interest, such an iterative subsystem scheme can be considered as an embedding scheme that not only accounts for the effect of the environment on the active subsystem but also includes the polarization of the environment caused by the active subsystem. Thus, such a polarizable embedding goes beyond a scheme in which a fixed frozen density is employed for the environment. Consequently, it will always converge to the same density as a full treatment, irrespective of the initial choice of $\rho_{\mathrm{II}}$ if no further approximations are introduced.

**Excitation energies and response properties.**

When treating excited states within FDE theory, one has to distinguish between the two available theoretical approaches: time-independent (state-specific) methods and response theory. The conceptually simpler theory is obtained in the case of time-independent methods in which a wavefunction is calculated explicitly for each excited state of interest. In this case, the embedding theory outlined above can be applied directly, with the only difference that the wavefunction and electron density of the active subsystem are different for each excited state. Thus, the embedding potential is different for each excited state, and for each excited state the environment density has to be determined iteratively (e.g., in freeze-and-thaw iterations). The theoretical justification for such a state-specific treatment of excited states is given in ref. [279].

In a formalism based on response theory [280–283] (see also paper D.3), the (time-dependent) electron densities $\rho_{\mathrm{I}}(\boldsymbol{r},t)$ and $\rho_{\mathrm{II}}(\boldsymbol{r},t)$ of the two subsystems are in a KS-DFT framework represented by two separate Slater determinants $|\tilde{\Phi}_{\mathrm{I}}\rangle$ and $|\tilde{\Phi}_{\mathrm{II}}\rangle$, respectively. Consequently, the total time-averaged quasi-energy can be expressed as

$$\{Q(t)\}_T = \{Q[\rho_{\mathrm{I}}](t)\}_T + \{Q[\rho_{\mathrm{II}}](t)\}_T + \{E_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}](t)\}_T, \tag{6.18}$$

where the first two terms are the time-averaged quasi-energies of the two subsystems according to Eqn. (5.60), and the third term is the time-average of the interaction energy given in Eqn. (6.8). Subsequently, an exponential parametrization with the parameters $\boldsymbol{\kappa}_{\mathrm{I}}$ and $\boldsymbol{\kappa}_{\mathrm{II}}$ can be introduced for both subsystems (cf. Eqn. (5.61)), i.e., the total time-dependent density is expressed as

$$\rho(\boldsymbol{r},t) = \rho_{\mathrm{I}}(\boldsymbol{r},\boldsymbol{\kappa}_{\mathrm{I}}) + \rho_{\mathrm{II}}(\boldsymbol{r},\boldsymbol{\kappa}_{\mathrm{II}}). \tag{6.19}$$

With this parametrization, the matrix $\boldsymbol{F}$ determing the poles of the response function assumes a block structure,

$$\boldsymbol{F} = \frac{\partial^2\{Q^{(2)}\}_T}{\partial\boldsymbol{\kappa}^{(1)}\partial\boldsymbol{\kappa}^{(1)}} = \begin{pmatrix} \boldsymbol{F}_{\mathrm{I,I}} & \boldsymbol{F}_{\mathrm{II,I}} \\ \boldsymbol{F}_{\mathrm{I,II}} & \boldsymbol{F}_{\mathrm{II,II}} \end{pmatrix} \quad \text{with} \quad \boldsymbol{F}_{\mathrm{n,m}} = \frac{\partial^2\{Q^{(2)}\}_T}{\partial\boldsymbol{\kappa}^{(1)}_n\partial\boldsymbol{\kappa}^{(1)}_m}. \tag{6.20}$$

By separating the contributions arising from the different terms in Eqn. (6.18), this matrix can be decomposed into

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{F}_{\mathrm{I,I}} & \boldsymbol{F}_{\mathrm{II,I}} \\ \boldsymbol{F}_{\mathrm{I,II}} & \boldsymbol{F}_{\mathrm{II,II}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_{\mathrm{I}} & 0 \\ 0 & \boldsymbol{F}_{\mathrm{II}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{F}^{\mathrm{I}}_{\mathrm{int}} & \boldsymbol{F}^{\mathrm{II,I}}_{\mathrm{int}} \\ \boldsymbol{F}^{\mathrm{I,II}}_{\mathrm{int}} & \boldsymbol{F}^{\mathrm{II}}_{\mathrm{int}} \end{pmatrix}, \tag{6.21}$$

where $\boldsymbol{F}_{\mathrm{I}}$ and $\boldsymbol{F}_{\mathrm{II}}$ arise from the differentiation of the quasi-energies of the isolated subsystems (as in Eqn. (5.63)) while the second contribution originates from the differentiation of the interaction energy. This interaction contribution contains blocks $\boldsymbol{F}^{\mathrm{I}}_{\mathrm{int}}$ and $\boldsymbol{F}^{\mathrm{II}}_{\mathrm{int}}$, which modify the diagonal of $\boldsymbol{F}$ and therefore be regarded as modifying the isolated subsystem $\boldsymbol{F}_{\mathrm{I}}$ and $\boldsymbol{F}_{\mathrm{II}}$ matrices,

$$\boldsymbol{F}^{(\mathrm{n})}_{\mathrm{int}} = \frac{\partial^2\{E^{(2)}_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]\}_T}{\partial\boldsymbol{\kappa}^{(1)}_n\partial\boldsymbol{\kappa}^{(1)}_n} = \iint w^{nn}_{\mathrm{emb}}(\boldsymbol{r},\boldsymbol{r}')\frac{\partial\rho^{(1)}_n(\boldsymbol{r})}{\partial\boldsymbol{\kappa}^{(1)}_n}\frac{\partial\rho^{(1)}_n(\boldsymbol{r}')}{\partial\boldsymbol{\kappa}^{(1)}_n}\,\mathrm{d}^3r\,\mathrm{d}^3r' + \int v^{n}_{\mathrm{emb}}(\boldsymbol{r})\frac{\partial^2\rho^{(2)}_n(\boldsymbol{r})}{\partial\boldsymbol{\kappa}^{(1)}_n\partial\boldsymbol{\kappa}^{(1)}_n}\,\mathrm{d}^3r \tag{6.22}$$

with the embedding kernel

$$w^{nn}_{\mathrm{emb}}(\boldsymbol{r},\boldsymbol{r}') = \frac{\delta^2 E_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]}{\delta\rho_n(\boldsymbol{r})\delta\rho_n(\boldsymbol{r}')} = \left.\frac{\delta^2 E_{\mathrm{xc}}[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right|_{\rho_{\mathrm{tot}}} - \left.\frac{\delta^2 E_{\mathrm{xc}}[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right|_{\rho_n} + \left.\frac{\delta^2 T_s[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right|_{\rho_{\mathrm{tot}}} - \left.\frac{\delta^2 T_s[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right|_{\rho_n}. \tag{6.23}$$

These lead to additional embedding contributions to the subsystem $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices (cf. Eqns. (5.65)-(5.66))

$$A^{nn}_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon^n_a - \varepsilon^n_i) + 2(ia|bj) + (ia|f_{\mathrm{xc}}|bj) + (ia|w^{nn}_{\mathrm{emb}}|bj), \tag{6.24}$$

$$B^{nn}_{ia,jb} = 2(ia|jb) + (ia|f_{\mathrm{xc}}|jb) + (ia|w^{nn}_{\mathrm{emb}}|jb), \tag{6.25}$$

where all orbital indices refer to the considered subsystem and the contributions arising from the second term in Eqn. (6.22) have been included in the orbital energies (i.e., it is assumed that the orbitals are obtained from Eqn. (6.9)).

Second, the off-diagonal blocks introduce a coupling between the subsystems, which is given by

$$\boldsymbol{F}^{\mathrm{I,II}}_{\mathrm{int}} = \frac{\partial^2\{E^{(2)}_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]\}_T}{\partial\boldsymbol{\kappa}^{(1)}_{\mathrm{I}}\partial\boldsymbol{\kappa}^{(1)}_{\mathrm{II}}} = \iint w^{\mathrm{I,II}}_{\mathrm{emb}}(\boldsymbol{r},\boldsymbol{r}')\frac{\partial\rho^{(1)}_{\mathrm{I}}(\boldsymbol{r})}{\partial\boldsymbol{\kappa}^{(1)}_{\mathrm{I}}}\frac{\partial\rho^{(1)}_{\mathrm{II}}(\boldsymbol{r}')}{\partial\boldsymbol{\kappa}^{(1)}_{\mathrm{II}}}\,\mathrm{d}^3r\,\mathrm{d}^3r' \tag{6.26}$$

with the embedding kernel

$$w^{\mathrm{I,II}}_{\mathrm{emb}}(\boldsymbol{r},\boldsymbol{r}') = \left.\frac{\delta^2 E_{\mathrm{xc}}[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right|_{\rho_{\mathrm{tot}}} + \left.\frac{\delta^2 T_s[\rho]}{\delta\rho(\boldsymbol{r})\delta\rho(\boldsymbol{r}')}\right|_{\rho_{\mathrm{tot}}} + \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{6.27}$$

These give rise to $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices corresponding to coupling between subsystems, with elements

$$A^{\mathrm{I,II}}_{i_{\mathrm{I}}a_{\mathrm{I}},j_{\mathrm{II}}b_{\mathrm{II}}} = (i_{\mathrm{I}}a_{\mathrm{I}}|w^{\mathrm{I,II}}_{\mathrm{emb}}|b_{\mathrm{II}}j_{\mathrm{II}}), \tag{6.28}$$

$$B^{\mathrm{I,II}}_{i_{\mathrm{I}}a_{\mathrm{I}},j_{\mathrm{II}}b_{\mathrm{II}}} = (i_{\mathrm{I}}a_{\mathrm{I}}|w^{\mathrm{I,II}}_{\mathrm{emb}}|j_{\mathrm{II}}b_{\mathrm{II}}), \tag{6.29}$$

where the subscripts indicate that orbital indices refer to the different subsystems.

In should be noted that the subsystem response theory discussed above in the framework of TDDFT can also be generalized to a wavefunction based description of the active subsystem. In this case, the DFT quasi-energy of subsystem I in Eqn. (6.18) is replaced by the quasi-energy Lagrangian of a wavefunction based method, i.e.,

$$\{L(t)\}_T = \{L[\rho_{\mathrm{I}}](t)\}_T + \{Q[\rho_{\mathrm{II}}](t)\}_T + \{E_{\mathrm{int}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}](t)\}_T, \tag{6.30}$$

and an appropriate parametrization of the wavefunction of subsystem I is introduced. This then translates to a parametrization of the total time-dependent density

$$\rho(\boldsymbol{r},t) = \rho_{\mathrm{I}}(\boldsymbol{r},\boldsymbol{\lambda}_{\mathrm{I}},\bar{\boldsymbol{\lambda}}_{\mathrm{I}}) + \rho_{\mathrm{II}}(\boldsymbol{r},\boldsymbol{\kappa}_{\mathrm{II}}) \tag{6.31}$$

in which the density of subsystem I depends on the parameters $\boldsymbol{\lambda}_{\mathrm{I}}$ and possibly the multipliers $\bar{\boldsymbol{\lambda}}_{\mathrm{I}}$. Embedding contributions to the isolated subsystem matrices $\boldsymbol{F}_{\mathrm{I}}$, $\boldsymbol{A}_{\mathrm{I}}$, and $\boldsymbol{J}_{\mathrm{I}}$ as appearing in the response function (Eqn. (5.59)) can then be derived by differentiating the interaction energy with respect to the parameters and multipliers. As in the TDDFT case, these will introduce both embedding contributions entering into the subsystem response and embedding contributions that couple the two subsystems. In general, linear-response function in Eqn. (5.59) will now involve the matrix

$$\begin{pmatrix} \boldsymbol{F}_{\mathrm{I,I}} & \boldsymbol{A}_{\mathrm{I,I}} & \boldsymbol{A}_{\mathrm{II,I}} \\ \boldsymbol{A}^T_{\mathrm{I,I}} & \boldsymbol{J}_{\mathrm{I,I}} & \boldsymbol{F}_{\mathrm{II,I}} \\ \boldsymbol{F}_{\mathrm{I,II}} & \boldsymbol{A}_{\mathrm{I,II}} & \boldsymbol{F}_{\mathrm{II,II}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}_{\mathrm{I}} & \boldsymbol{A}_{\mathrm{I}} & 0 \\ \boldsymbol{A}^T_{\mathrm{I}} & \boldsymbol{J}_{\mathrm{I}} & 0 \\ 0 & 0 & \boldsymbol{F}_{\mathrm{II}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{F}^{\mathrm{I}}_{\mathrm{int}} & (\boldsymbol{A}^{\mathrm{I}}_{\mathrm{int}})^T & \boldsymbol{F}^{\mathrm{II,I}}_{\mathrm{int}} \\ \boldsymbol{A}^{\mathrm{I}}_{\mathrm{int}} & \boldsymbol{J}^{\mathrm{I}}_{\mathrm{int}} & \boldsymbol{A}^{\mathrm{II,I}}_{\mathrm{int}} \\ \boldsymbol{F}^{\mathrm{I,II}}_{\mathrm{int}} & \boldsymbol{A}^{\mathrm{I,II}}_{\mathrm{int}} & \boldsymbol{F}^{\mathrm{II}}_{\mathrm{int}} \end{pmatrix}, \tag{6.32}$$

due to the presence of additional coupling blocks. Note that because of the non-linear dependence of the interaction energy on the multipliers (via the density $\rho_{\mathrm{I}}$) the matrix $\boldsymbol{J}_{\mathrm{I,I}}$ will in general not be zero anymore, even if the isolated subsystem contribution $\boldsymbol{J}_{\mathrm{I}}$ is. For further discussion and explicit equations, we refer to paper D.3. We note that this formalism is general enough to also allow for the definition of WFT-in-WFT approaches [284, 285].

## 6.2 Approximate Embedding Methods

*This section partially reproduces the contents of section 4 of* ***ASP Gomes, CR Jacob, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem., 2012, 108, 222–277***

While the FDE theory presented in the previous section provides an exact theoretical framework for embedding methods in quantum chemistry, it is not directly suitable for numerical applications. In particular, it requires the knowledge of the exact nonadditive kinetic-energy functional, which are not easily available in practice. Therefore, numerous approximate embedding schemes have been developed instead. For discussing these in the following section, it is useful to establish a classification scheme for such methods.

One scheme for classifying approximate embedding methods has been proposed by Bakowies and Thiel (BT)[286] and follows the steps taken above for presenting the FDE theory.

- **Mechanical coupling**: The simplest way to setup an embedding scheme is to treat each subsystem individually and introduce the coupling only through the total energy. To this end, an interaction energy between the subsystems is calculated according to Eqn. (6.8) or some approximation to it. With such a simple scheme only the geometrical structure of the active subsystem is altered. All electronic properties, in particular the electron density, are identical to those obtained when treating the subsystem of interest in isolation.

- **Electronic coupling**: Embedding schemes that include the effect of the environment in the quantum-chemical treatment of the active subsystem in some way form the next category. This can be achieved by including the embedding potential of Eqn. (6.11) both in a KS-DFT or in a wavefunction based treatment of the subsystem of interest. In practice, this embedding potential is usually approximated. With such a coupling through an embedding potential, the electronic properties of the subsystem of interest can be affected by its environment.

- **Polarizable embedding**: In the simplest case of embedding schemes with electronic coupling, the (approximate) embedding potential is determined solely by the geometric structure of the environment. For instance, the frozen electron density calculated for the isolated environment (i.e., in the absence of the subsystem of interest) can be used in Eqn. (6.11). More advanced schemes can be set up by including the polarization of the environment due to the presence of the active subsystem. This leads to schemes where the embedding potential has to be determined iteratively.

- **Embedding including environment response:** When treating electronic excitation energies or other response properties, a fourth category — not contained in the original BT classification — can be introduced. In this case one can distinguish whether or not the response of the environment to the electronic excitation is included. With state-specific methods, this can be achieved by iteratively updating the environment density for each excited state instead of employing one common frozen environment density for all excited states. Within response theory, the response of the environment can be included through the additional subsystem and coupling contributions to the response matrices discussed in the previous section. Several different strategies can be introduced to approximate these contributions. First, the coupling contributions can be neglected and only those modifying the subsystem response are retained. Second, the coupling contributions can be included in an approximate fashion. Finally, it is also possible to fold the contributions of the environment as well as the coupling in an approximate fashion into the response matrices of the active subsystem. While this will not account for coupling between individual excitation energies, it does allow for an efficient inclusion of the polarization of the environment density.

Among those groups, different methods can be classified according to the approximations that are introduced for calculating the interaction energy and the embedding potential. First, in continuum solvation models the discrete molecular structure of the environment is neglected and replaced by a

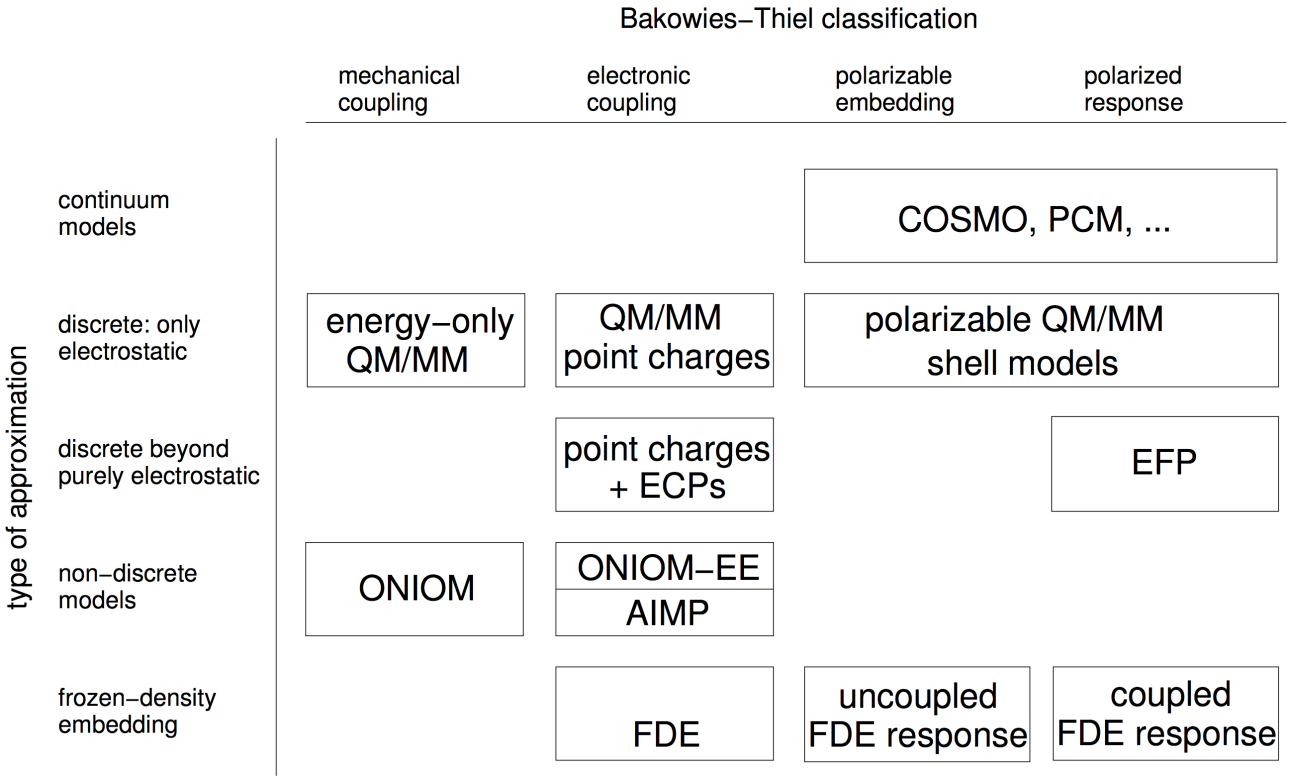



Figure 6.1: Overview of some of the available approximate embedding scheme. On the horizontal axis are the categories of the extended Bakowies–Thiel classification, while the vertical axis sorts different approaches according to the models employed for the environment.

dielectric continuum. Next, a large variety of embedding methods uses a discrete description for the environment and models the electrostatic part of the environment using point charges or localized multipole moments. Such a purely electrostatic description can be augmented with additional terms accounting for non-classical interactions. As a further step, there are embedding methods that retain a full electron density for the environment using the embedding potential of Eqn. (6.11), but employing an approximate electron density for the environment and introduce approximations for the nonadditive kinetic-energy functional. Finally, embedding methods that do not introduce such approximations have also been proposed for embedding accurate wavefunction theory calculations into an environment treated with DFT.

**In the following, we will focus exclusively in the approximations to the FDE exact theory, but refer the reader to [25], where an extensive discussion of the different approximations is presented.**

We note, however, that a classification in terms of sophistication does not imply an equivalence in terms of accuracy. Often, seemingly simple schemes have been parametrized so that they provide rather accurate results. At all levels of approximations, embedding methods can operate in the different categories in the BT classification scheme. This is sketched in the two-dimensional overview in Fig. 6.1. In all cases further distinctions can be made according to the methods used for describing the active subsystems. However, in most cases this does not affect the embedding methodology significantly, so we will only make this distinction in a few cases.

### Frozen-density embedding with approximate kinetic-energy functionals

The frozen-density embedding (FDE) scheme aims at providing a full description of environment effects by approximating the exact embedding potential of Eqn. (6.11). However, for the exchange–correlation functional $E_{\mathrm{xc}}[\rho]$ and for the nonadditive kinetic-energy functional $T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}]$ (Eqn. (6.5)) and its

functional derivative

$$v_T[\rho_{\mathrm{I}},\rho_{\mathrm{II}}](\boldsymbol{r}) = \frac{\delta T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]}{\delta\rho_{\mathrm{I}}(\boldsymbol{r})} = \left.\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})}\right|_{\rho_{\mathrm{tot}}} - \left.\frac{\delta T_s[\rho]}{\delta\rho(\boldsymbol{r})}\right|_{\rho_{\mathrm{I}}} \tag{6.33}$$

one now has to introduce approximations. Thus, in the interaction energy given in Eqn. (6.8) and in the embedding potential the contributions of the kinetic energy and of the exchange–correlation energy are approximated. The idea to use an approximate kinetic-energy functional to evaluate the interaction energy between two fixed electron densities dates back to the work of Kim and Gordon.[287, 288] Later, it was extended to electron densities determined using the above embedding potential, both in subsystem approaches[264–266] and in embedding schemes.[263]

If both the active subsystem and its environment are described with DFT using an approximate exchange–correlation functional that depends only locally on the electron density (i.e., LDA or GGA functionals), the exchange–correlation contribution can be treated consistently. With hybrid functionals or with orbital-dependent exchange–correlation potentials, a local functional has to be used for the nonadditive exchange–correlation contributions, which constitutes an additional approximation.[289, 290]

**Approximation to $T_s^{\mathbf{n}add}[\rho_{\mathbf{I}},\rho_{\mathbf{II}}]$ and $v_T[\rho_{\mathbf{I}},\rho_{\mathbf{II}}]$.**

With a local exchange–correlation functional, differences between a full DFT calculation and an embedding treatment in which the densities of both subsystems are optimized are due to the approximations applied for the kinetic energy, provided that the full supermolecular basis set expansion is used for both subsystems.[277, 291] Thus, comparing the electron densities from such calculations offers a way for assessing the quality of approximations for $v_T[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$, whereas a comparison of the total energies also probes the quality of $T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$. These strategies have been used to develop and tests approximations both for the kinetic-energy component of the embedding potentials and to the nonadditive kinetic energy. Here, we will only give a brief overview of the most widely used approximations and highlight some more recent developments. Dedicated reviews on kinetic-energy functionals in general[292, 293] and in the context of the FDE scheme[275] are available in the literature (for more recent overviews, see, e.g., the introductions of refs. [276, 294–296]).

The simplest class of approximations applies an approximate kinetic energy functional in Eqn. (6.5) as well as for the functional derivative in Eqn. (6.33). These are referred to as *decomposable approximations*. Early studies employed the well-known Thomas–Fermi functional, but following a series of tests for hydrogen-bonded systems,[297, 298] Wesolowski proposed the use of generalized-gradient approximation (GGA) kinetic-energy functionals within the FDE scheme. In particular, he recommended[298] the use of the PW91k functional of Lembarki and Chermette,[299] which has been used almost exclusively in applications of the FDE scheme in the past decades. More recently, new GGA kinetic-energy functionals have also been proposed for the use in decomposable approximations to the kinetic-energy component of the FDE interaction energy and embedding potential.[296]

Several studies have assessed the quality of the available decomposable approximations. Generally, these provide a good accuracy of the electron densities and molecular properties as long as the interaction between the subsystems is dominated by weak, non-covalent interactions such as hydrogen bonding. This has, for instance, been demonstrated by comparing the electron densities obtained in FDE calculations to those of a full treatment.[300, 301] For interaction energies, the PW91k approximation provides a typical accuracy of ca. 1–2 kcal/mol in hydrogen bonded complexes.[295]

Even though successful for weak interactions between the subsystems, several shortcomings of the available decomposable approximations based on GGA functionals have been pointed out. In the limit of infinitely separated subsystems the potential shows a wrong form at the frozen subsystem, which affect the resulting orbital energies and can lead to spuriously low excitation energies.[302] This shortcoming can partly be addressed with so-called *non-decomposable approximations*,[294, 302] in which the non-additive kinetic energy or the potential $v_T[\rho_{\mathrm{I}},\rho_{\mathrm{II}}]$ are approximated directly. Even more severe is the failure of all presently available approximations for subsystems connected by covalent

bonds,[301] even if these covalent bonds are very weak.[303] These problems have so far not been addressed satisfactorily, but recent work provides some possible directions for future improvements.[276] Alternatively, the insufficiencies of the currently available approximations can be circumvented by using a more general partitioning that introduces capping groups.[304]

**DFT-in-DFT embedding.**

Initial applications of the FDE embedding potential in combination with approximate kinetic-energy functionals focussed on ionic crystals.[265, 266] This initial work aimed at a subsystem formulation of DFT, and treats all subsystems—in this case the individual ions—on the same footing, i.e., the density of all subsystems are optimized iteratively in freeze-and-thaw iterations. In the scheme of Cortona, only spherical ions are considered and the embedding potential is spherically averaged.[305–307] Mehl and coworkers extended this scheme to general, non-spherical fragments.[308–310] Wesolowski and Warshel pioneered the use of the approximate FDE embedding potential in applications that focus on a specific subsystem of interest, while its environment is kept frozen.[263] Their initial applications concerned the solvation of lithium ions in water as well as the solvation free energies of water and methane. In these applications, an additional approximations was introduced: Instead of obtaining the electron density $\rho_{\mathrm{II}}$ of the environment from a full DFT calculation (or from a fully self-consitent subsystem DFT calculation), it was approximated as the sum of the densities of isolated solvent molecules. Such approximate ways of constructing the environment density are key to efficient DFT-in-DFT FDE calculations.

The applications discussed so far focussed on ground-state properties. Excited states can be treated in such DFT-in-DFT embedding calculations either with a state-specific approach or within response theory. A state-specific approach is realized if excited states of the active subsystem are described using a ΔDFT or ΔSCF-DFT approach (see section 5.1). This was applied by Wesolowski and coworkers to study crystal field splittings for impurities in ionic crystals[311–313], using the ground-state embedding potential also for the excited states.

As outlined in section 6.1, the FDE theory can also be extended to a description of excited states within response theory, in particular with TDDFT. In applications of such an approach, one usually—in addition to the use of approximations to the kinetic-energy contribution to the embedding kernel—introduces approximations for the treatment of the embedding contributions to the response equations (Eqn. (6.21)). The simplest approximation is to neglect the off-diagonal coupling blocks $\boldsymbol{F}_{\mathrm{int}}^{\mathrm{I,II}}$ and $\boldsymbol{F}_{\mathrm{int}}^{\mathrm{II,I}}$ arising from the embedding contribution.[314] This leads to a decoupling of the response equations of the two subsystems, and the energies of local excitations of the active subsystem can be determined by considering only the matrix $\boldsymbol{F}_{\mathrm{I,I}} = \boldsymbol{F}_{\mathrm{I}} + \boldsymbol{F}_{\mathrm{int}}^{\mathrm{I,I}}$, where the additional embedding contribution is determined by the embedding kernel given in Eqn. (6.23). Such an approximation corresponds to a neglect of the response of the environment.

Such a scheme can be employed for the calculation of solvent effects on local excitation energies by combining it with approximate construction of the solvent electron density.[315] Because the TDDFT response calculation is limited to the active subsystem describing the solute molecule this results in an efficient treatment and allows for the inclusion of large frozen solvent shells as well as an averaging over a sufficient number of solvent structures. The simplest approximation for the solvent density is to use the sum of the densities of isolated water molecules. Such a description can be further refined by updating the density of a few solvent molecules close to the active subsystem in freeze-and-thaw iterations.[269, 316] Similar schemes can be used to treat local excitations in protein environments.[317, 318]

With a fixed frozen density, such FDE calculations correspond to electronic embedding in the BT classification. Accounting for the ground-state polarization of the environment leads to a polarizable embedding scheme, but within the approximation discussed so far the polarized response of the environment density is not included. A discussion of these different contributions and a comparison to a polarizable QM/MM description can be found in ref. [319].

A computational strategy for an efficient treatment of the full embedding contributions to the

response matrices has been devised by Neugebauer.[281, 282] In his subsystem TDDFT scheme, off-diagonal coupling contributions to the response matrices are not neglected. Instead, the excitation energies of the individual subsystems are determined first, and in a second step the coupling contributions are included only for those excitations that are of interest. This allows for an efficient treatment of both the polarization of the environment[282, 320] and of couplings between local excitations.[281] Neugebauer and coworkers have employed their scheme in several studies of photosynthetic systems, in particular light harvesting complexes.[317, 321] At about the same time of paper D.3, Pavanello looked at the question of coupled excitations from the perspective of Dyson equations[283], and in the years that followed Wasserman and coworkers have investigated the time-dependent formulation of partition DFT[322–324]. Pavanello and coworkers have also reported a real-time TDDFT implementation[325], based on a periodic implementation of FDE[326, 327], as well as different FDE-based approaches to treat charge-transfer processes[328, 329]. Recent reviews on the calculation of excitation energies with subsystem TDDFT and on the related applications are available[24, 330–332].

**WFT-in-DFT and WFT-in-WFT embedding.**

The application of the FDE embedding potential in combination with approximations for the kinetic-energy functional for embedding a wavefunction based description of the active subsystem in an environment described by DFT (WFT-in-DFT embedding) was pioneered by Carter and coworkers.[271, 333] Their work focussed on the description of molecules absorbed on metallic surfaces. Their pilot application concerned the description of ground-state properties of CO on a Cu(111) surface and used a scheme in which the density of the active subsystem is obtained with CI or CASSCF and is updated iteratively, while the total density is obtained from a periodic DFT calculation and is kept fixed. Subsequently, this scheme was extended to the treatment of excited states within these state-specific methods for the active subsystem to describe the local excitations of CO on a Pt(111) surface.[334, 335]

The limitations of these scheme where addressed in later work, in which the constraint that the total density is kept fixed was relaxed.[336, 337] Instead, the environment density is chosen as $\rho_{\mathrm{II}} = \rho_{\mathrm{tot}} - \rho_{\mathrm{I}}^{\mathrm{bare}}$, where the total density is obtained from a periodic DFT calculation and $\rho_{\mathrm{I}}^{\mathrm{bare}}$ is the density of the isolated subsystem I. This environment density is then kept frozen, i.e., the polarization of the environment is only included through the periodic DFT description, but not updated according to the wavefunction based calculation. This scheme has been applied in a number of studies of ground state properties[338–341] and of local excited states.[342] For a review of these WFT-in-DFT embedding approaches and their applications, see ref. [21–23].

For the calculation of excitation energies, wavefunction based methods are often required because of the well-known limitations of TDDFT, but nevertheless a DFT calculation provides an adequate ground-state density. Therefore, a simplified WFT-in-DFT embedding scheme, in which the embedding potential is obtained from a DFT-in-DFT embedding calculation (either using a fixed approximate environment density or an environment density polarized in freeze-and-thaw iterations) has been proposed in paper D.1. This simplified scheme has been applied to study electronic excitations of heavy element species in ionic crystals (papers D.1 and D.6) or surrounded by noble gas species (paper D.4) using IHFSCC and DFT methods. The performance of this approach is discussed below.

In applications of WFT-in-DFT embedding, the response of the environment via the intra-subsystem kernel contributions has been accounted for in papers D.3 and D.5, with a first application for WFT-in-WFT embedding for local excitation by Höfener and Visscher[284], who later extended this work to describe coupled excitations[285] in an approximate (RI-CC2) formalism[343].

With the time-independent wavefunction based methods that were mainly employed (i.e., CASSCF, CI, and IHFSCC), the incorporation of the response of the environment is possible by using state-specific embedding potentials.[279, 344] While the state-specific approach can be computationally advantageous as it avoids the calculation of a large number of excited states on both subsystems, it has the disadvantage of making it more difficult to calculate transition moments since the different electronic states of the active subsystem will no longer be orthogonal. In connection to state-specific approaches, the question of whether external orthogonality would be required between the subsystems[345] has

been posed, but recent works show that for linearly independent basis FDE would not require it.[346]

### Frozen-density embedding with optimized effective potentials

For calculating local excitations and other local molecular properties, FDE calculations employing approximate kinetic-energy functional can provide an accurate description of environment effects in certain cases. In particular, the available approximations are applicable if the interaction between the subsystem of interest and its environment is weak or dominated by electrostatic interactions. However, even in these cases the available approximations to the kinetic-energy component $v_T[\rho_\mathrm{I}, \rho_\mathrm{II}]$ of the embedding potential have deficiencies. These inevitably introduce small, but sometimes not negligible errors into the calculated excitation energies. These can be reduced by increasing the size of the active subsystem, but especially in WFT-in-DFT embedding calculations this is not desirable and often not feasible. On the other hand, a description of covalent bonds between subsystems is not possible with the currently available approximations.

Therefore, variants of the FDE scheme that avoid such approximations for the kinetic-energy functional or its functional derivative have been developed in recent years by several groups. The evaluation of the kinetic-energy component $v_T[\rho_\mathrm{I}, \rho_\mathrm{II}]$ of the embedding potential requires the evaluation of the functional derivative of the noninteracting kinetic-energy functional $T_s[\rho]$ for two different densities, the total density $\rho_\mathrm{tot} = \rho_\mathrm{I} + \rho_\mathrm{II}$ and the density of the active subsystem $\rho_\mathrm{I}$. By using the Euler-Lagrange equation for the KS system of noninteracting electron with a given, fixed density $\rho(\boldsymbol{r})$ this functional derivative can be related to the local potential $v_s[\rho](\boldsymbol{r})$ that has this density $\rho(\boldsymbol{r})$ as its ground state,[347]

$$\frac{\delta T_s[\rho]}{\delta \rho(\boldsymbol{r})} = -v_s[\rho](\boldsymbol{r}) + \mu, \tag{6.34}$$

where $\mu$ is a constant shift that is related to the chemical potential. Thus, the kinetic-energy component of the embedding potential can be evaluated from[302]

$$v_T[\rho_\mathrm{I}, \rho_\mathrm{II}](\boldsymbol{r}) = v_s[\rho_\mathrm{I}](\boldsymbol{r}) - v_s[\rho_\mathrm{tot}](\boldsymbol{r}) + \Delta\mu \tag{6.35}$$

as the difference between the local potentials yielding the density of the active system and the total density, respectively. These local potentials yielding a certain density can be evaluated numerically. Different algorithms for such a potential reconstruction (often also referred to as optimization of effective potentials or OEP methods) have been developed. In the context of embedding calculations, the algorithms of van Leeuwen and Baerends [139] and of Zhao, Morrison, and Parr (ZMP) [348] as well as schemes based on the direct optimization algorithm of Wu and Yang (WY) [349] have been employed. While the van Leeuwen–Baerends and the ZMP schemes employ a numerical representation of the potential on a grid, the algorithm of Wu and Yang expands the potential in a suitable basis set. Even though the details differ, all embedding schemes avoiding approximations for $v_T[\rho_\mathrm{I}, \rho_\mathrm{II}]$ are based on such an optimization of an effective potential.

Within DFT-in-DFT embedding schemes, approaches calculating an approximate embedding potential are usually computationally not advantageous, especially if one is only interested at ground-state properties. For determining the embedding potential, one or more calculations on the full system are required, which embedding schemes usually aim to avoid. Nevertheless, such calculations can be employed to demonstrate that schemes based on the embedding potential of Eqn. (6.11) do indeed reproduce the electron density of a full calculation. Moreover, the reconstruction of accurate embedding potentials can further guide the development of new approximations to the kinetic-energy component of the embedding potential $v_T[\rho_\mathrm{I}, \rho_\mathrm{II}]$. The latter was the aim of the study of Fux *et al.*[276], and a similar study performed by Goodpaster *et al.*[350] However, such schemes are not suitable for practical calculations. Therefore, Goodpaster *et al.* extended their scheme[351] by introducing a pairwise approximation that relies on a further partitioning of the frozen environment density (see Eqn. (6.2). Instead of calculating the kinetic-energy component of the embedding potential as in Eqn. (6.35), it is approximated as

$$\tilde{v}_T^{(\mathrm{pair})}[\rho_\mathrm{I}, \rho_\mathrm{II}^{(n)}] \approx \sum_n v_s[\rho_\mathrm{I}] - v_s[\rho_\mathrm{I} + \rho_\mathrm{II}^{(n)}] \tag{6.36}$$

This approximation turns out to be very accurate for small water clusters, and might provide a way to the efficient simulation of condensed-phase systems. More recently, Artiukhin *et al.* [352] have performed calculations of local excitations with reconstructed potentials. While their approach required KS-DFT calculations on the total system and the use of a localization procedure in order to divide the system into subsystems, the TD-DFT calculations – which can amount to a substantial part of the calculation – were performed only for the active subsystems. Their results underscore the important gain in accuracy afforded by replacing the kinetic energy density functionals.

As a DFT calculation on the full system might be less expensive than the correlated WFT calculation on the small subsystem, the use of accurate embedding potentials becomes a much more feasible proposition for WFT-in-DFT embedding schemes. This was realized by Roncero *et al.*, who were the first to propose the use of OEP methods in the context of WFT-in-DFT embedding.[353] Their scheme starts with a DFT or HF calculation on the full system, from which the total electron density $\rho_{\text{tot}}(\boldsymbol{r})$ is obtained. This density is then partitioned into an active subsystem and its environment, and an accurate embedding potential for the active subsystem is determined with the ZMP algorithm by requiring that the chosen density of the active subsystem is reproduced in a DFT or HF calculation. This potential is then included in the wavefunction based treatment of the active subsystem. Note that, even though no approximate kinetic-energy functional is used, such a scheme still introduces several approximations. First, the total electron density is calculated with approximate DFT or HF. Second, neither the total density nor the embedding potential are refined to account for the difference between the DFT or HF electron density and the one obtained from a correlated wavefunction based treatment (i.e, an accurate DFT-in-DFT embedding potential is used as approximation to the WFT-in-DFT embedding potential). Finally, when constructing a suitable partitioning of the total density, it is difficult to ensure that the subsystem densities are $v_s$-representable. In particular when localized orbitals are employed, these densities usually contain nodes, which makes them difficult to reproduce with a local potential expanded in a finite basis set (see also the discussion in refs. [276, 354, 355]). To address the latter problem, Roncero *et al.* extended their scheme to allow for an iterative refinement of the density partitioning.[356]

However, as for the DFT-in-DFT studies in refs. [276] and [350] discussed above, the resulting density partitioning — and thus also the embedding potential — are not unique. This shortcoming was addressed by Carter and co-workers, who defined a unique partitioning by using the idea of partition density-functional theory (P-DFT) of Wasserman and co-workers[267, 357] to require that the active subsystem and its environment share a common embedding potential.[278] Subsequently, they presented a reformulation of the embedding theory in terms of an optimization of the embedding potential, that allows for a conceptually simple implementation of WFT-in-DFT embedding schemes that do not rely on approximate kinetic-energy functionals.[358, 359] Even though results were only presented for DFT-in-DFT embedding calculations, this scheme can be easily extended to obtain accurate WFT-in-DFT embedding potentials, provided a correlated WFT method that allows for an efficient calculation of the electron density is used.

In such a complete WFT-in-DFT scheme (i.e., one in which the density calculated with WFT is used to construct an accurate embedding potential), the only remaining approximations are those inherent to the (approximate) WFT treatment of the subsystem of interest and the (approximate) DFT treatment of the environment as well as the use of an approximate functional for the exchange–correlation component of the embedding potential. However, all these approximations are justified and controllable. The largest remaining obstacle for such complete WFT-in-DFT schemes is the need for OEP methods. In combination with a finite orbital basis set the reconstruction of the local potential corresponding to a given density is an ill-posed problem.[360, 361] Therefore, the embedding potentials obtained with finite-basis set OEP methods are in general not unique. This will affect the energy and density from a correlated WFT calculation on the active subsystem as well as molecular properties. Thus, numerically stable OEP methods that provide unambiguous embedding potentials are required [362–364] and a new approach addressing these issues has been developed recently.[365]

The existing complete WFT-in-DFT methods using accurate embedding potentials[278, 358] can also be applied directly to a state-specific WFT approaches for calculating excited states. This

can either be done in an approximate fashion using a common frozen environment density or in a full treatment that determines a state-specific embedding potential. However, the extension to response theory is still an open issue, because this will also require the calculation of an accurate kinetic-energy contribution to the embedding kernel, unless a hybrid approach – where embedding potentials are obtained with OEP and the kernel and any other higher-order contributions with approximate kinetic energy density functionals – is put in place.

## 6.3 Actinyls in $Cs_2UO_2Cl_4$

Having discussed in section 5.3 the electronic spectra calculations for actinyl species in the gas-phase, I now turn my attention to the electronic spectra in the $Cs_2UO_2Cl_4$ crystal for uranyl (pure crystal) and neptunyl (which is taken as a impurity) using FDE. The full studies can be found in appendices D.1 for neptunyl and D.6 for uranyl, respectively. For neptunyl we have relied exclusively on IHFSCCSD calculations with a two-component Hamiltonian and double-zeta basis sets, whereas for uranyl we were able to perform four-component calculations with triple-zeta basis sets and both IHFSCCSD and CAM-B3LYP. In all cases, embedding potentials were constructed with the ADF code with triple-zeta basis sets, the spinfree ZORA Hamiltonian and the PW91k kinetic energy density functional, and exported to be used in Dirac with the help of the PyADF scripting framework (see paper D.2 for a detailed description of PyADF).

The $Cs_2UO_2Cl_4$ crystal [214] is composed by a network of actinyl tetrachloride dianions surrounded by cesium cations. In it, the actinyl tetrachloride species shows a $C_{2h}$ site symmetry, slightly deviating from the $D_{4h}$ structure due to the plane containing the chloride being slightly tilted. This species is particularly well-suited as a benchmark system for theoretical approaches as its electronic spectra as been experimentally investigated in great detail (one-photon [366] and two-photon [367] absorption measurements for the valence excited states, for excited states [368] and for core excitations [369]).

Before discussing any results, it is important to address the question of how to partition the system into active and environment subsystems. Here the goal is to try using the smallest possible active subsystem that can still capture the essential physical processes behind the property we want to model. Since we are interested in the electronic spectra of actinyl chorides ($AnO_2Cl_4^{2-}$, An = U, Np), we have the choice of considering an actinyl chloride species as the active subsystem and the rest of the crystal as the environment, or attempt a partition where the chlorides ligands are part of the environment as well. This second approach, if suitable, is interesting from the perspective of computations (it reduces the size of the system on which we will compute excitation energies) as well as of analyses (it makes it easier to relate the results to the bare uranyl results).

Once the partition into subsystems is chosen, a second point to be addressed is how the different components of the total system are to be treated. In figure 6.2 one finds the models for the active subsystems and environments explored for uranyl chloride: first, there are four models for which interactions with the host crystal are not take into account: apart from the bare uranyl (model a), which has been discussed in section 5.3, there is a rather crude model (model b) which consists of considering the equatorial ligands as point charges, and another (model c) in which the chlorides are represented by an embedding potential from FDE; and finally (model d) we consider the complete uranyl chloride. In model (c) one would have the choice of performing subsystem DFT calculations (e.g. optimizing both the uranyl and the chlorides' densities via freeze-thaw cycles) or via FDE (where the chlorides' densities would be determined from a calculation on the isolated subsystem and used to construct the embedding potential). In all of these we have maintained the idealized $D_{4h}$ structure (though calculations are performed in $D_{2h}$ symmetry).

In models (e) and (f) the host crystal is represented by a two-layer model: the first layer consists of an ensemble of 20 uranyl chloride units and about 90 cesium ions surrounding the active subsystem, treated via FDE, whereas the second layer (not shown) is composed of a large array of point charges obtained in such a way as to represente the long-range electrostatic potential in the solid. As can be seen in figure 6.2, models (e) and (f) differ in the subsystems which are allowed to have their

densities related: in (e) only the equatorial chlorides are relaxed, and in (f) the freeze-thaw procedure is extended to the cesium ions nearest to the active uranyl unit. Since in these models we employ the X-ray structure, we retain in the calculations the $C_{2h}$ symmetry. In the case of four-component calculations, this has as consequence the use of complex algebra and a memory usage that is twice as large as the calculations in $D_{2h}$ symmetry.

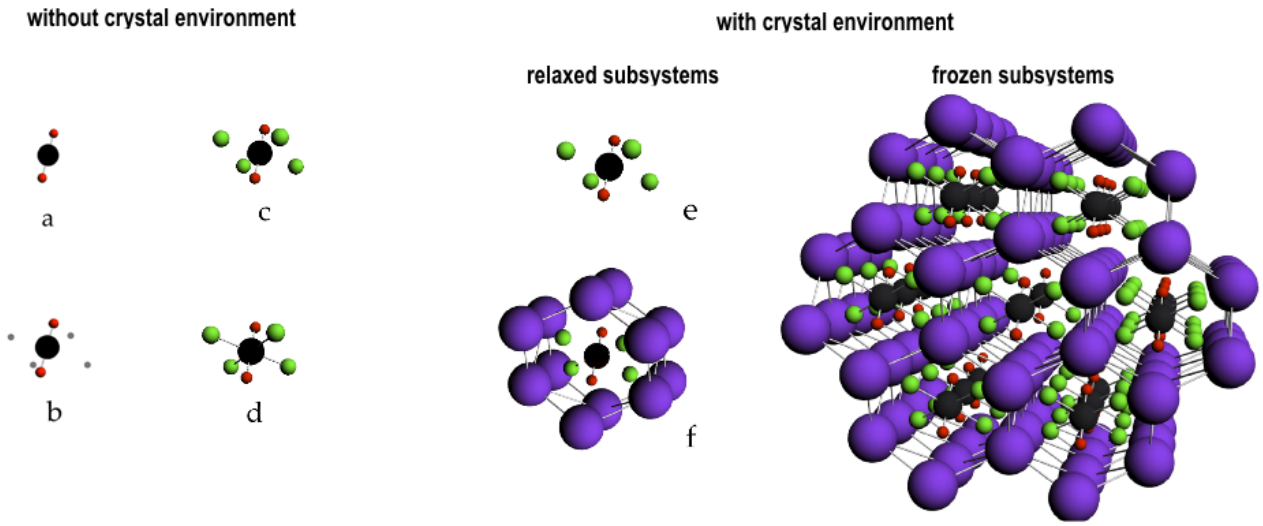



Figure 6.2: Models for the description of environment effect on actinyl chloride species. Models without the crystal environment: (a) the bare actinyl; (b) actinyl with point-charge embedding; (c) actinyl with FDE embedding; (d) actinyl chloride. Models including the crystal environment (shown to the right), including the relaxation of the environment densities of: (e) only the neareast chlorides; (f) the nearest chlorides and 12 cesium ions. (uranium: black; oxygen: red; cesium: purple; chlorine: green)

In table 6.1 selected results from papers D.6 and D.1 are shown. It should be noted that in the latter models (e) and (f) were not explored, and the two-layered model was used together with the neptunyl chloride as the active subsystem. Starting with the neptunyl results, we see that the simple point-charge embedding (model b) is far from being a quantitatively correct model but it does get the correct ordering of states with respect to neptunyl chloride (model d) and experiment and is off by about 1000 $cm^{-1}$ for the three lowest excited states but shows significantly larger discrepancies for the $\Pi$ states. We see that model (c) represents a major improvement over model (b), yielding results that are quite close to model (d) and to experiment, even though in both cases the energy of state V is still strongly overestimated with respect to experiment. Here it is important to note that if the equatorial ligands are not relaxed, as is the case in model (c)†, results are no better than those for model (b). Because of this finding, in paper D.6 we have only performed calculations where the chlorides' density was relaxed.

The addition of the crystal environment (model (e)‡) hardly affects states II and III but does shift states I, IV upward by about 200 and 800 $cm^{-1}$ with respect to model (d), and state V by about 4000 $cm^{-1}$. We did not attempt to explore the reasons for such discrepancies, and it would be interesting to verify whether this is due to the deficiencies in the model space, the orbital basis, or some other factor. In comparison to the prior work of Matsika and Pitzer [370], we see that our WFT-in-DFT calculations are closer to experiment for the first three states, but show poorer agreement for the $\Pi$ states.

Considering now the case of uranyl chloride, we again see that model (c) is a rather good approximation to model (d), yielding excitation energies that are roughly 1000 $cm^{-1}$ lower than those from model (d) for all states considered. While we believe this is a rather good accuracy for such a simple model it may be that the agreement to model (d) may still be improved if the kernel contributions to the excitation energies are included. These were not taken into account in paper D.6 in order to better compare to the IHFSCC results, for which kernel contributions are not yet defined.

Table 6.1: $f$-$f$ excited state energies (in $cm^{-1}$) for $AnO_2^{2+}$ and $AnO_2Cl_4^{2-}$ (An = U, Np), obtained with the inclusion of environment effects via point-charge, DFT-in-DFT and WFT-in-DFT Embedding. All calculations were performed using the X-ray structures [214] ($r_{NpO}$ = 1.775 Å, $r_{NpCl}$ = 2.653 Å, $r_{UO}$ = 1.774 Å, $r_{NpCl}$ = 2.675 Å). Results taken from papers D.1 and D.6.

| | | Excited Electronic States | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Ac | Method/Model | I | II | III | IV | V | VI | VII | VIII |
| | | $\Delta_{3/2u} + \Phi_{5/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ | | | |
| Np | IHFSCC/Q (b) | 2093 | 8032 | 7828 | 23326 | 26321 | | | |
| | (c)† | 2243 | 8150 | 7677 | 23323 | 26433 | | | |
| | (c) | 690 | 7108 | 8176 | 18870 | 21776 | | | |
| | (d) | 886 | 7679 | 9262 | 20018 | 22445 | | | |
| | (e)‡ | 1156 | 7738 | 9137 | 20857 | 26305 | | | |
| | SO-CI (AIMP) [370] | 1663 | 5775 | 8463 | 18367 | 20575 | | | |
| | Exp. [215] | 900-1050 | 6880 | 7890 | 17241 | 20081 | | | |
| | | $B_{2g}$ | $B_{3g}$ | $B_{1g}$ | $A_g$ | $B_{2g}$ | $B_{3g}$ | $A_g$ | $B_{1g}$ |
| U | CAMB3LYP (b) | 17265 | 17265 | 16341 | 16239 | 17681 | 17681 | 19394 | 19660 |
| | (c) | 18115 | 18115 | 18321 | 17981 | 19565 | 19565 | 20539 | 20829 |
| | (d) | 19018 | 19018 | 19934 | 19288 | 20970 | 20970 | 21745 | 21592 |
| | (c') | 18114 | 18112 | 17975 | 18317 | 19568 | 19552 | 20536 | 20825 |
| | (e) | 18134 | 18136 | 17938 | 18263 | 19520 | 19501 | 20514 | 20826 |
| | (f) | 18119 | 18120 | 17913 | 18236 | 19494 | 19475 | 20494 | 20085 |
| | IHFSCC/Q112(c) | 17998 | 17998 | 18705 | 19409 | 20689 | 20689 | 21797 | 21855 |
| | (c') | 18151 | 18154 | 18874 | 19522 | 20860 | 20843 | 21944 | 22025 |
| | (f) | 18128 | 18124 | 18816 | 19492 | 20760 | 20768 | 21848 | 21905 |
| | Exp. [371] | 20095.7 | 20097.3 | 20406.5 | 21316 | 22026.1 | 22076 | 22406 | 22750 |

† FDE calculation without freeze-thaw procedure.
‡ Active subsystem is $NpO_2Cl_4^{2-}$, instead of $NpO_2^{2+}$ as in figure 6.2

Model (d) shows very good agreement overal with experiment, usually underestimating the transition energies with respect to the latter by about 1000 $cm^{-1}$. One slightly unfortunate feature of CAM-B3LYP in this case is that the ordering of the lowest states does not conform to the experimental assignment (the first $A_g$ should not be lower than the first $B_{1g}$, for instance). This comparison so far can only be made on the basis of CAM-B3LYP calculations, as we were unable to perform IHFSCCSD calculations on uranyl chloride due to convergence difficulties. That said, the IHFSCCSD/Q112 results we have for models (c) show no such change in the ordering of states. Unlike the case of neptunyl, adding the crystal environment (models e and f) has a rather limited effect on the excitation energies, both for DFT-in-DFT and WFT-in-DFT results, with shifts of the order of 100 $cm^{-1}$. A closer inspection shows that this is due to the fact that the crystal environment lowers the energies of the occupied and virtual spinors by roughly the same amount so that the effect largely cancels out for excitation energies.

In addition to the electronic spectrum, we have investigated the effect of the environment on the ionization potential for uranyl chloride in the gas-phase and in the $Cs_2UO_2Cl_4$ crystal. From our results for the first ionization potential, summarized in Table 6.2, we see that WFT-in-DFT calculations yield results which are in very good agreement with experiment for the gas-phase (model c) and far superior than the results from the simple point-charge embedding (model b). As we used the same static embedding potential obtained from DFT-in-DFT calculations in both DFT-in-DFT and WFT-in-DFT calculations, the underestimation seen for CAM-B3LYP arises form the use of spinor energies as an approximation to the ionization potential. The difference of about 1 eV with respect to the IHFSCCSD results is furthermore in line with what has been observed for uranyl, NUN and

$NUO^+$.

For the IHFSCCSD calculations, we see that the effect of employing a reduced virtual space (IHFSCCSD/$Q_1$) brings about a slight decrease (0.3 eV) of the ionization energies and by chance brings it closer to experimental the experimental result of Dau and coworkers [372]. One should recall that in all of our models we use the X-ray structures and therefore have not investigated whether using the complex's geometry in the gas-phase would lower the calculated ionization energies and bring them closer to the experimental values. Other source of errors in our calculations that could bring the ionization energies down are the lack of higher-order correlation corrections such as triples that could help to capture further relaxation effects (these are not available for IHFSCC), as well as the use of an embedding potential obtained for the non-ionized species – after ionization, we expect a non-negligible change in the environment due to a polarization induced by the now even more positively charge uranyl subsystem, that could help lower the energy required to ionize the species.

For the calculations in the $Cs_2UO_2Cl_4$ crystal, we have the same difference of about 1 eV between WFT and DFT results. However, a result that is surprising at first glance is the better agreement of CAM-B3LYP with experiment (about 0.1 eV difference) than IHFSCCSD (1 eV difference). While we have not yet performed and further explorations on this matter, we note that in the ionization from a solid the final state will be stabilized due to the polarization of the species' closest environment, which is something our calculations have not taken into account, and in other halide crystals such a relaxation has been estimated to be of about 1eV [371]. Assuming this to hold in our case, our values would have to be decreased by the same amount and we would be in a similar situation as for the gas-phase, with IHFSCC being quite close to the experimental results and CAM-B3LYP underestimating the ionization energies.

Table 6.2: CAM-B3LYP and IHFSCCSD absolute ionization energies (in eV), corresponding to ionization from the highest $e_{1u}$ occupied spinor, for different models for embedded uranyl.

| phase | Method/Model | IP |
|---|---|---|
| gas-phase | CAM-B3LYP(b) | 6.16 |
| | IHFSCCSD/$Q_2$ (b) | 7.78 |
| | CAM-B3LYP(c) | 4.24 |
| | IHFSCCSD/$Q_2$ (c) | 5.81 |
| | IHFSCCSD/$Q_1$ (c) | 5.51 |
| | Exp. [372] | $\simeq$5 |
| $Cs_2UO_2Cl_4$ | CAM-B3LYP(f) | 9.32 |
| | IHFSCCSD/$Q_1$(f) | 10.61 |
| | Exp. [371] | $\simeq$9.4 |

From these results we see DFT-in-DFT and WFT-in-DFT embedding allow us to define relatively compact models for the actinyl species in environment such as ionic crystals, where only the actinyl species is taken as the active subsystem and equatorial ligands and other species in the vicinity are treated as an effective potential. This model has introduced errors of the order of hundreds of wavenumbers for neptunyl, and of about a thousand wavenumbers for uranyl, but we cannot tell at this point whether this will remain the case for other species such as plutonyl, or for different ligands. Furthermore, we have been able to show that the crystal environment does not greatly affect excitation energies – thus justifying the approach taken in most investigations to ignore long-range interactions – but its proper description is absolutely vital for other properties such as ionization potentials.

## 6.4 The shortcomings of the kinetic energy density functionals: CUO-Ng complexes

The promising results obtained for actinyl species discussed above motivated us to, in paper D.4, use WFT-in-DFT and DFT-in-DFT to try to shed some light on the question of what is the ground-state for the CUO molecule in neon and argon noble gas matrices. Our expectation was that since the interaction between CUO and the noble gases is weak, approximate density functionals such as PW91k would not have difficulty in describing the non-additive kinetic energy contributions to the embedding potential. Here one should recall that the non-additive kinetic energy contributions to the embedding potential act to offset the strong attraction between the electron density of the active subsystem and the nuclear framework in the frozen subsystems, and if this balance between repulsion and attraction is not achieved, the electron density will tend to spuriously localize on, or around, the frozen subsystem [302, 373, 374].

Much to our surprise we observed that, while this held true to a good extent for neon, as the noble gas became heavier there was a marked deterioration of results. This can be seen, first, by comparing the energies of the valence orbitals for conventional DFT calculations on the CUO-$Ng_4$ complexes and those for the CUO in subsystem calculations using the PW91k functional, shown in figure 6.3: there, we see that the unoccupied $f_\sigma$ orbital that in conventional DFT calculations should be destabilized along the series until it becomes close in energy to the highest $\delta$ component, whereas the $\phi$ and the other low-lying $\delta$ orbitals should be stabilized and become very close in energy for Ng = Xe. For the embedding calculations, however, the trends are nearly opposite to that: while there is a destabilization of the $f_\sigma$ orbital for Ng = Ne, this orbital gets strongly stabilized along the series, and both the $\phi$ and the other low-lying $\delta$ orbitals become destabilized.

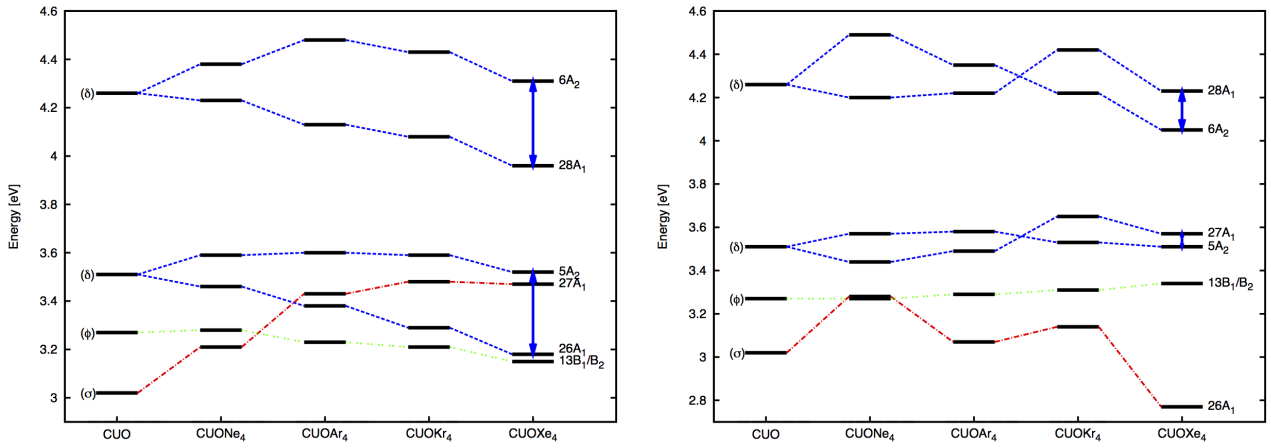



Figure 6.3: Lowest-lying valence orbital energies of the CUO ($^1\Sigma^+$) and the CUO$Ng_4$($^1A_1$) complexes from the supermolecular and DFT-in-DFT/PBE0 calculations using the PW91k kinetic energy functional. Orbital energies are given relative to the energy of HOMO for all the compounds.

Other measures of the difficulties of the approximate kinetic energy density functionals are the integrated errors in the electron density [375], as well as the errors in the dipole moments, between the supermolecular and subsystem calculations shown in table 6.3. From the table we see that if the $\Delta^{rms}$ and, particularly, $|\Delta_\mu|$ are reasonable for FDE calculations for Ng = Ne, already for Ng = Ar they increase significantly and remain large for Ng = Kr, Xe. This indicates spurious localization of charge in these species.

In order to better understand what is causing such behavior, it is instructive to look at figures 6.4 to 6.7, where plots of the embedding potentials and the different contributions to it are plotted along the Ng-U-Ng bond for an idealized structure in which the noble gas species are on the same plane as the uranium atom. Even though in paper D.4 we investigated the Thomas-Fermi kinetic energy functional as well as a scheme to enforce exact conditions a posteriori [302], in these figures only results for the PW91k (the one yielding the smallest absolute errors in the electron density for most systems)

and NDSD [294] (constructed with the goal of correcting for spurious charge localization by enforcing exact conditions at the nuclei of the frozen subsystems) functionals are shown.

Table 6.3: Integrated errors in the electron density: $\Delta^{abs}$ (absolute) and $\Delta^{rms}$(root mean square), magnitude of the error in the dipole moment $|\Delta_\mu|$ for the sum of fragments and the DFT-in-DFT/PBE0 calculations with a different approximate non-additive kinetic energy functionals. The most accurate values are marked in boldface.

| System | | $\Delta^{abs} \times 10^{-3}$ | $\Delta^{rms} \times 10^{-3}$ | $\|\Delta_\mu\|$(D) |
|---|---|---|---|---|
| $^1$CUONe$_4$ | Sum of fragments | 1.35 | 0.99 | 0.183 |
| | FDE (NDSD) | 4.84 | 0.03 | 0.033 |
| | FDE (PW91K) | **2.32** | **0.02** | **0.010** |
| $^1$CUOAr$_4$ | Sum of fragments | 2.76 | 1.39 | 0.354 |
| | FDE (NDSD) | 1.40 | **0.06** | **0.139** |
| | FDE (PW91K) | **1.39** | 0.07 | 0.261 |
| $^1$CUOKr$_4$ | Sum of fragments | 2.71 | 1.23 | 0.225 |
| | FDE (NDSD) | 1.10 | 0.06 | 0.115 |
| | FDE (PW91K) | **1.03** | 0.06 | 0.188 |
| $^1$CUOXe$_4$ | Sum of fragments | 2.27 | 0.11 | 0.205 |
| | FDE (NDSD) | 1.41 | 0.07 | 0.222 |
| | FDE (PW91K) | 1.40 | 0.08 | 0.302 |

The figures show an important difference between the Ng = Ne and the other cases, in that the embedding potentials (red curves) pass from having a small repulsive wall (at about 3 Å from the uranium atom) to showing a well at roughly the same position, that can contribute to artificially localizing charge in the region between the uranium and noble gas atoms. In the idealized structures this would have no effect on the dipole moment, but in the actual CUO-noble gas complexes where the uranium is slighlty above the plane defined by the ligands, such a charge accumulation will lead to a modified dipole moment. The change from a repulsive to an attractive interaction in the region of the noble gases also goes towards explaining, for instance, why the $f_\sigma$ orbital which is mostly located at the region between the noble gas ligands and the CUO molecule as shown in figure 2 of paper D.4, is not shifted upward along the series for the embedding calculations and is instead stabilized. Our calculations also underscore the difficulty of improving functionals such as Thomas-Fermi or PW91k by enforcing when the nuclear charge is very large as done for NDSD – though it must be said that these aproaches have mostly been developed and validated with relatively light elements in mind. We observe that while qualitatively both PW91k and NDSD yield the same results (though PW91k does perform slightly better overall), PW91k potentials are in general much smoother than NDSD ones in the vicinity of the atoms in the environment.

Table 6.4: Low-lying SOC vertical excitation energies of the CUO and CUONg$_4$ complexes from IHFSCCSD-in-DFT. via sector $(0h, 2p)$, embedding potential from DFT-in-DFT calculations using the PW91k kinetic energy functional.

| State | CUO | CUONe$_4$ | CUOAr$_4$ | CUOKr$_4$ | CUOXe$_4$ |
|---|---|---|---|---|---|
| $^3\Sigma^+_{0,1}$† | 0.81 | 1.02 | 0.85 | 0.76 | 0.65 |
| $^3\Phi_2$ | 0.94 | 0.84 | 0.89 | 0.93 | 0.97 |
| $^3\Phi_3$ | 1.01 | 0.94 | 0.98 | 1.02 | 1.06 |

† Difference between $^3\Sigma^+_0$ and $^3\Sigma^+_1$ is less than 0.01 eV

In spite of these shortcomings, it is nevertheless instructive to see how these perform in WFT-in-DFT calculation, and to what extend we can obtain qualitative trends that would point to a possible change in ground state. IHFSCCSD-in-DFT results with embedding potentials obtained with the PW91k functional are presented in table 6.4 for some of the low-lying states. We have not explored the use of NDSD potentials for the calculation of excited states, but given their less smooth behavior

one may wonder how suitable they would really be for such calculations. From these calculations we observe that environment effects are probably not sufficient to change the nature of the ground-state, but do seem enough to place the $^3\Phi_2$ state below – or at least very close to – the $^3\Sigma^+_{0,1}$ states and therefore may well change the ordering of the excited states. But without performing calculations with more accurate potentials (possibly such as those obtained with OEP approaches) one cannot rule out such changes.

Tecmer and coworkers have recently tried to address the question with DMRG calculations on the neon and argon complexes [376], and have concluded on the basis of spin-free calculations that such a change is possible. However, considering the importance of dynamical correlation effects for the CUO system discussed in section 5.4 these DMRG results (which are not guaranteed to account for most of the dynamical correlation effects) must be viewed with caution, and I consider this question is still not fully resolved.

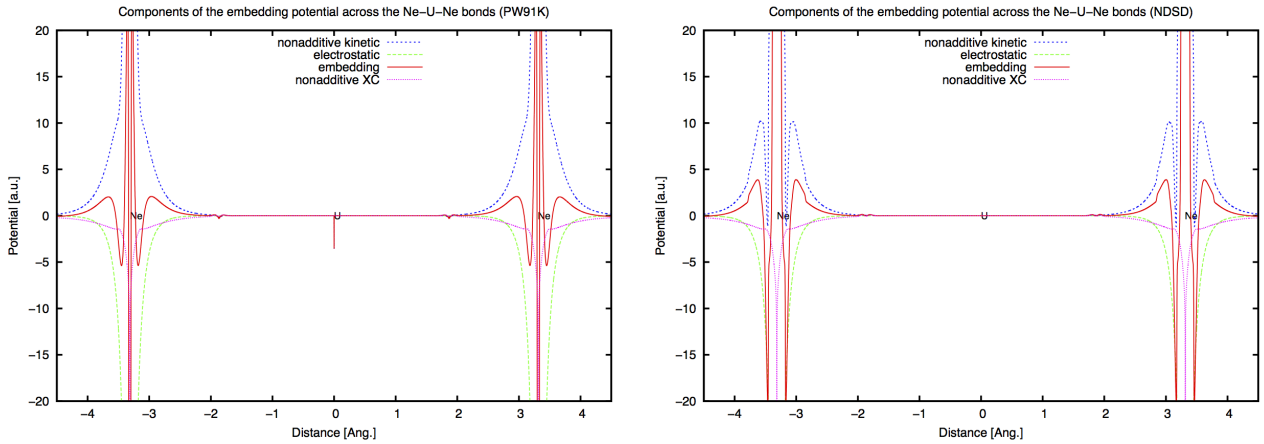



Figure 6.4: Embedding potentials for the CUO-$Ne_4$ complex employing the PW91k and NDSD kinetic energy functionals

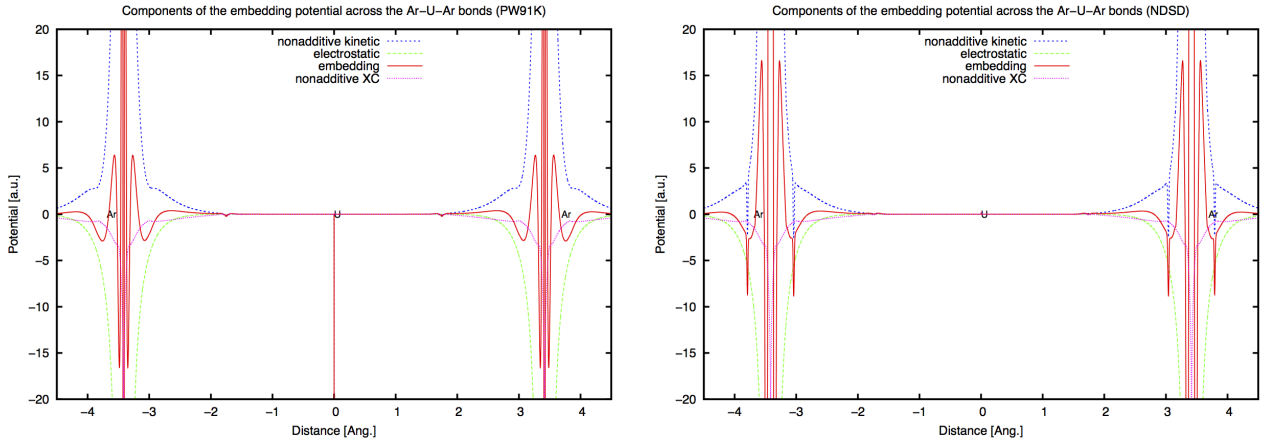



Figure 6.5: Embedding potentials for the CUO-$Ar_4$ complex employing the PW91k and NDSD kinetic energy functionals

## 6.5 FDE Second-order magnetic properties: formalism and first applications

In what follows I outline the development of FDE for the calculation of second-order magnetic properties for a four-component relativistic Hamiltonian, discussed in detail in paper D.7 along with the study of $H_2X$–$H_2O$ (X = Se, Te, Po) model systems of selected properties. It should be noted, in contrast to the great activity in determining excited state properties with FDE, there has been next to

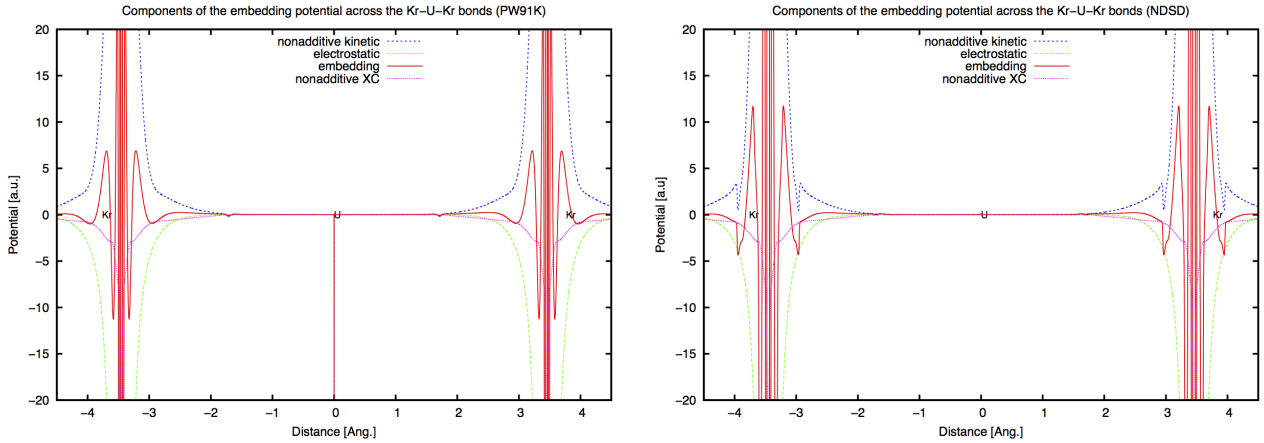



Figure 6.6: Embedding potentials for the CUO-$Kr_4$ complex employing the PW91k and NDSD kinetic energy functionals

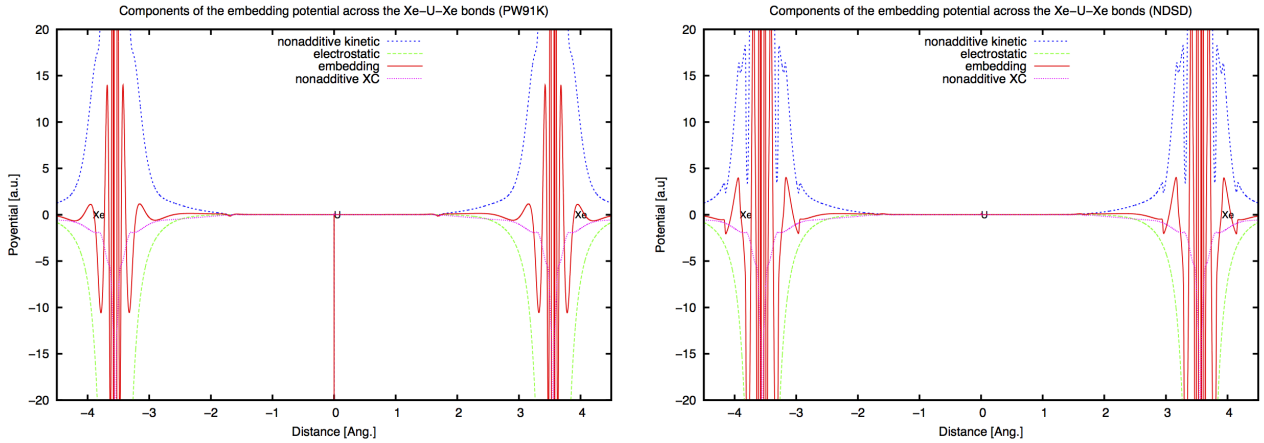



Figure 6.7: Embedding potentials for the CUO-$Xe_4$ complex employing the PW91k and NDSD kinetic energy functionals

no activity on the calculation of magnetic properties, even though Jacob and Visscher [377] have outlined the theory in 2006 (and shown how promising the method is in a first application [378]) bar the implementation of indirect spin-spin constants [379]. These works mostly focused on non-relativistic Hamiltonians, though they can be used with the approximate ZORA Hamiltonian as well. This makes ours the first effort in a four-component framework.

Before outlining the theory for FDE, we recall that for closed-shell systems, a static magnetic field induces only even-order changes in the total energy: [380]

$$E(\varepsilon) = E_0 + \frac{1}{2}\frac{d^2E}{d\varepsilon_1 d\varepsilon_2}\varepsilon_1\varepsilon_2 + \frac{1}{4!}\frac{d^4E}{d\varepsilon_1 \dots d\varepsilon_4}\varepsilon_1 \dots \varepsilon_4 + \dots \tag{6.37}$$

where $E_0$ denotes the energy at zero field and $\{\varepsilon_n\}$ the field strengths of applied perturbations collected in vector $\varepsilon$. The coefficients of this expansion, taken in the zero-field limit, define molecular properties in the Born–Oppenheimer approximation. The focus here is on three second-order magnetic properties arising from a perturbation of an external field $\vec{B}$ or the field of nuclear magnetic dipole moments, $\{\vec{m}_A\}$: the NMR shielding tensor of a nucleus K,

$$\sigma^K_{\alpha\beta} = \left.\frac{d^2E}{dB_\alpha dm_{K;\beta}}\right|_{\vec{B},\{\vec{m}_A\}=0}, \tag{6.38}$$

the reduced spin-spin coupling tensor of nuclei K and L,

$$K^{KL}_{\alpha\beta} = \left.\frac{d^2E}{dm_{K;\alpha} dm_{L;\beta}}\right|_{\{\vec{m}_A\}=0}, \tag{6.39}$$

related to the indirect spin-spin coupling constants observed in NMR experiment,

$$J^{KL} = (\hbar/2\pi)\gamma_{\mathrm{K}}\gamma_{\mathrm{L}}K^{KL}_{\alpha\beta}, \tag{6.40}$$

with $\gamma_{\mathrm{M}}$ denoting the gyromagnetic ratio of a given isotope of $M$, and the molecular magnetizability tensor,

$$\xi_{\alpha\beta} = -\frac{d^2E}{dB_\alpha dB_\beta}\bigg|_{\vec{B}=0}. \tag{6.41}$$

### Subsystem time-independent linear response (LR) theory

Considering now the case of time-independent perturbations with strengths $\varepsilon_1$ and $\varepsilon_2$, the second-order molecular property can be written as:

$$\frac{d^2E}{d\varepsilon_1 d\varepsilon_2}\bigg|_{\varepsilon=0} = \frac{\partial^2 E}{\partial\kappa_{pq}\partial\varepsilon_2}\frac{\partial\kappa_{pq}}{\partial\varepsilon_1}\bigg|_{\varepsilon=0} + \frac{\partial^2 E}{\partial\varepsilon_1\partial\varepsilon_2}\bigg|_{\varepsilon=0}, \tag{6.42}$$

assuming that the energy is optimized with respect to variational parameters at all field strengths, $\partial E/\partial\kappa_{pq} = 0$. The first contribution is determined perturbatively, with the first-order orbital rotation amplitudes, $\partial\kappa_{pq}/\partial\varepsilon_1$, obtained from the LR equations:

$$0 = \frac{d}{d\varepsilon_1}\left(\frac{\partial E}{\partial\kappa_{pq}}\right)\bigg|_{\varepsilon=0} = \left(\frac{\partial^2 E}{\partial\kappa_{pq}\partial\varepsilon_1} + \frac{\partial^2 E}{\partial\kappa_{pq}\partial\kappa_{rs}}\frac{\partial\kappa_{rs}}{\partial\varepsilon_1}\right)\bigg|_{\varepsilon=0}, \tag{6.43}$$

which can be recast in a matrix form as: [176]

$$0 = \mathbf{E}^{[1]}_{\varepsilon_1} + \mathbf{E}^{[2]}_0\mathbf{X}_{\varepsilon_1}. \tag{6.44}$$

Here, $\mathbf{E}^{[2]}_0$ is the electronic Hessian, $\mathbf{E}^{[1]}_{\varepsilon_1}$ the property gradient and $\mathbf{X}_{\varepsilon_1}$ the solution vector yielding $\{\kappa^{\varepsilon_1}_{rs}\}$. While the Hessian is independent on a perturbation, the property gradient is calculated as the first-order derivative of the KS matrix with respect to field strength of applied perturbation,

$$\mathbf{E}^{[1]}_{\varepsilon_1} = \begin{bmatrix} g^{\varepsilon_1} \\ g^{*\varepsilon_1} \end{bmatrix}, \qquad g^{\varepsilon_1}_{ai} = \frac{\partial E_{\varepsilon_1}}{\partial\kappa^*_{ai}}\bigg|_0 = \langle 0|[-\hat{a}^\dagger_i\hat{a}_a, \hat{h}_{\varepsilon_1}]|0\rangle = -\tilde{F}^{\varepsilon_1}_{ai}. \tag{6.45}$$

In particular, if $\varepsilon_1 = \vec{B}$, the property gradient is calculated in OMO basis and requires additional contributions involving derivatives of LAOs and of matrices T [192, 194]. Once $\mathbf{X}_{\varepsilon_1}$ has been determined, one can construct the static linear response function:

$$\langle\langle\varepsilon_1;\varepsilon_2\rangle\rangle_0 = E^{[1]\dagger}_{\varepsilon_1}X_{\varepsilon_2} = -E^{[1]\dagger}_{\varepsilon_1}\left(E^{[2]}_0\right)^{-1}E^{[1]}_{\varepsilon_2} \tag{6.46}$$

which constitutes the response contribution to the molecular property expressed by the first term of Eq. 6.42. The second term of Eq. 6.42 can be thought of as an expectation value, which due to the linearity of the DC Hamiltonian in applied perturbations (see Eq. 5.106) is non-zero only in perturbation-dependent basis sets. This brings about to the final form of the properties mentioned above:

$$K^{KL}_{\alpha\beta} = \langle\langle m_{K;\alpha}; m_{L;\beta}\rangle\rangle_0 \tag{6.47}$$

$$\sigma^{K}_{\alpha\beta} = \langle\langle m_{K;\alpha}; B_\beta\rangle\rangle_0 \tag{6.48}$$

$$\xi_{\alpha\beta} = -\left(\langle\langle B_\alpha; B_\beta\rangle\rangle_0 + \frac{\partial^2 E}{\partial B_\alpha\partial B_\beta}\bigg|_0\right) \tag{6.49}$$

with the LAO basis used for the last two.

The properties in question being static (=time/frequency-independent), we can consider the zero-frequency limit of the embedding response formalism discussed above and arrive, as shown in detail in Paper D.7, at equations analogous to Eqs. 6.42 and 6.43, where the electronic Hessian and

property gradient are subdivided into isolated subsystem and interaction contributions, leading to a (coupled) system of LR equations

$$(\mathbf{E}_0^{[2];M,M} + \mathbf{E}_{0;int}^{[2];M,M})\mathbf{X}_{\varepsilon_1}^M + \mathbf{E}_{0;int}^{[2];M,N}\mathbf{X}_{\varepsilon_1}^N = -(\mathbf{E}_{\varepsilon_1}^{[1];M} + \mathbf{E}_{\varepsilon_1;int}^{[1];M}), \tag{6.50}$$

$$\mathbf{E}_{0;int}^{[2];N,M}\mathbf{X}_{\varepsilon_1}^M + (\mathbf{E}_0^{[2];N,N} + \mathbf{E}_{0;int}^{[2];N,N})\mathbf{X}_{\varepsilon_1}^N = -(\mathbf{E}_{\varepsilon_1}^{[1];N} + \mathbf{E}_{\varepsilon_1;int}^{[1];N}), \tag{6.51}$$

where the response vector has also been split into blocks pertaining to each subsystem, $\mathbf{X}_{\varepsilon_1} = [\mathbf{X}_{\varepsilon_1}^M; \mathbf{X}_{\varepsilon_1}^N]^\dagger$. The matrix elements of each subblock have the form

$$\mathbf{E}_0^{[2];M,M} = \frac{\partial^2 E_M}{\partial\kappa_{pq}^M \partial\kappa_{rs}^M}; \qquad \mathbf{E}_{0;int}^{[2];M,N} = \frac{\partial^2 E_{int}}{\partial\kappa_{pq}^M \partial\kappa_{rs}^N} \tag{6.52}$$

for the Hessian and

$$\mathbf{E}_{\varepsilon_1}^{[1];M} = \frac{\partial^2 E_M}{\partial\kappa_{ai}^M \partial\varepsilon_1}; \qquad \mathbf{E}_{\varepsilon_1;int}^{[1];M} = \frac{\partial^2 E_{int}}{\partial\kappa_{ai}^M \partial\varepsilon_1} \tag{6.53}$$

for the property gradient. The Hessian terms will have the same general form as in the case of electric perturbations, but will contain additional terms due to the spin-density contributions in the (S)DFT framework. The embedding contributions to the property gradient and expectation values, which have respectively the general form

$$\left.\frac{\partial^2 E_{int}}{\partial\kappa_{pq}^M \partial\varepsilon_1}\right|_0 = \int \left.\frac{\delta E_{int}}{\delta\rho_k^M}\frac{\partial^2\rho_k^M}{\partial\kappa_{pq}^M\partial\varepsilon_1}\right|_0 + \iint \left.\frac{\delta^2 E_{int}}{\delta\rho_k^M\delta\rho_{k'}^N}\frac{\partial\rho_k^M}{\partial\kappa_{pq}^M}\right|_0 \left.\frac{\partial\rho_{k'}^N}{\partial\varepsilon_1}\right|_0 \tag{6.54}$$

$$\left.\frac{\partial^2 E_{int}}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 = \sum_{M=I,II}\int \left.\frac{\delta E_{int}}{\delta\rho_k^M}\frac{\partial^2\rho_k^M}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 + \sum_{M,N=I,II}\iint \left.\frac{\delta^2 E_{int}}{\delta\rho_k^M\delta\rho_{k'}^N}\frac{\partial\rho_k^M}{\partial\varepsilon_1}\right|_0 \left.\frac{\partial\rho_{k'}^N}{\partial\varepsilon_2}\right|_0. \tag{6.55}$$

are identically zero unless the (generalized) density components carry a dependence on the perturbation. This is the case for electric properties (and because of that we have not explicitly carried over these terms) and for the indirect spin-spin couplings, but in the case of NMR shieldings and magnetizabilities the use of London atomic orbitals introduces such a dependency and therefore these terms should be considered.

By evaluating the partial derivatives we have the expressions for the property gradient for the determination of NMR shieldings and magnetizabilities in ($M \neq N$),

$$\left.\frac{\partial}{\partial B_\alpha}\frac{\partial E_{\mathrm{int}}}{\partial\kappa_{ai}^M}\right|_0 = -\int v_{emb;k}^M(\vec{r})\breve{\Omega}_{ia;k}^{B_\alpha;M}\mathrm{d}\vec{r} \tag{6.56}$$

$$- \iint w_{emb;k,k'}^{M,M}(\vec{r}_1,\vec{r}_2)\Omega_{ia;k}^M(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\alpha;M}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 \tag{6.57}$$

$$- \iint w_{emb;k,k'}^{M,N}(\vec{r}_1,\vec{r}_2)\Omega_{ia;k}^M(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\alpha;N}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2. \tag{6.58}$$

containing embedding potential and embedding kernel contributions, the latter being further subdivided into intra-subsystem and coupling contributions as was the case for the Hessian. The expectation value embedding contributions to the magnetizabilities can also be subdivided into intra-subsystem

$$\left.\frac{\partial^2 E_{int}^M}{\partial B_\alpha\partial B_\beta}\right|_0 = \int v_{emb;k}^M(\vec{r})\breve{\Omega}_{ii;k}^{B_\alpha B_\beta;M}\mathrm{d}\vec{r} \tag{6.59}$$

$$+ \iint w_{emb;k,k'}^{M,M}(\vec{r}_1,\vec{r}_2)\breve{\Omega}_{ii;k}^{B_\alpha;M}(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\beta;M}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2, \tag{6.60}$$

and coupling contributions

$$\left.\frac{\partial^2 E_{int}^{MN}}{\partial B_\alpha\partial B_\beta}\right|_0 = \iint w_{emb;k,k'}^{M,N}(\vec{r}_1,\vec{r}_2)\breve{\Omega}_{ii;k}^{B_\alpha;M}(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\beta;N}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2, \tag{6.61}$$

both containing embedding kernel terms.

### Tensor Expressions for the molecular properties and their representation in terms of magnetically induced currents

The theory discussed above is sufficient to determine the properties of interest in the subsystem approach. However, these properties can also be presented in a different mathematical form using the linearity of the 4c DC Hamiltonian in applied perturbations, complemented by the formulation involving magnetically induced current densities, which more directly conveys the physical characteristics of each property. In what follows, we use the definitions of Mason [381] for the isotropic and the anisotropic parts of a tensor.

#### NMR shielding and indirect spin-spin coupling tensors

The NMR shielding or the NMR indirect spin-spin coupling tensors in Eqs. 6.47 and 6.48 can be recast in a computationally advantageous form [192] in terms of expectation values involving the hyperfine operator for a nuclei $L$, the unperturbed spinors $|\psi_i\rangle$ and the first-order perturbed spinors [192, 194] $|\tilde{\psi}_i^{\epsilon_i}\rangle$, yielding the general expression

$$M_{\alpha\beta}^{\epsilon;L} = \sum_i \left\{ \langle \tilde{\psi}_i^{\epsilon_\alpha} | \hat{h}_{m_{L;\beta}} | \psi_i \rangle + \langle \psi_i | \hat{h}_{m_{L;\beta}} | \tilde{\psi}_i^{\epsilon_\alpha} \rangle \right\}. \tag{6.62}$$

The expression for the shielding tensor $\sigma_{\alpha\beta}^L$ is therefore obtained from Eq. 6.62 employing the spinors perturbed by the external magnetic field ($\epsilon = \mathbf{B}$), $|\tilde{\psi}_i^{B_\alpha}\rangle$, and by the same token the spin-spin coupling tensor $K_{\alpha\beta}^{KL}$ is obtained by employing the spinors perturbed by the nuclear magnetic dipole ($\epsilon = \vec{m}_K$), $|\psi_i^{m_{K;\alpha}}\rangle$.

In the FDE case, as each subsystem is described by its own set of externally-orthogonal orbitals, we can rewrite the expression in Eq. 6.62 as

$$M_{\alpha\beta}^{\epsilon;L} = \sum_{i\in I} \left\{ \langle \tilde{\psi}_i^{\epsilon_\alpha} | \hat{h}_{m_{L;\beta}} | \psi_i \rangle + \langle \psi_i | \hat{h}_{m_{L;\beta}} | \tilde{\psi}_i^{\epsilon_\alpha} \rangle \right\} + \sum_{j\in II} \left\{ \langle \tilde{\psi}_j^{\epsilon_\alpha} | \hat{h}_{m_{L;\beta}} | \psi_j \rangle + \langle \psi_j | \hat{h}_{m_{L;\beta}} | \tilde{\psi}_j^{\epsilon_\alpha} \rangle \right\} \tag{6.63}$$

The FDE expression for $\sigma_{\alpha\beta}^L$ or $K_{\alpha\beta}^{KL}$ can be further approximated by neglecting the terms arising from subsystem $II$. In the case of NMR shieldings, assuming nucleus $L$ belongs to subsystem $I$, this approximation should be good, especially if the overlap between two subsystems is small, but whatever the case we can estimate this missing contribution by the magnetically-induced current density formulation outlined in section 6.5. For the spin-spin tensor this approximation should also be good due to the local nature of the hyperfine operator, if both $K$ and $L$ belong to subsystem $I$ (a restriction in our current implementation).

#### Magnetizability tensor

Contrary to NMR properties, the magnetizability tensor is not a local property as the Zeeman operator (Eq. 5.106) affects the whole system. It can be expressed in terms of the sum of (interacting) intra-subsystem and inter-subsystem contributions

$$\xi_{\alpha\beta} = \xi_{\alpha\beta}^{I,(II)} + \xi_{\alpha\beta}^{II,(I)} \tag{6.64}$$

where

$$\xi_{\alpha\beta}^{I,(II)} = \left[ \frac{\partial^2 E_I}{\partial B_\alpha \partial B_\beta}\bigg|_0 + \frac{\partial^2 E_{int}^{M=I}}{\partial B_\alpha \partial B_\beta}\bigg|_0 \right] + \left[ \frac{\partial^2 E_I}{\partial \kappa_{pq}^I \partial B_\beta} + \frac{\partial^2 E_{int}}{\partial \kappa_{pq}^I \partial B_\beta} \right] \frac{\partial \kappa_{pq}^I}{\partial B_\alpha}. \tag{6.65}$$

#### Tensors in terms of induced currents

The first-order derivatives of the relativistic current density vector with respect to perturbations [382],

$$\vec{j}^{\varepsilon_1}(\vec{r}) = -e \sum_i \left\{ \tilde{\psi}_i^{\varepsilon_1 \dagger} c\vec{\alpha} \psi_i + \psi_i^\dagger c\vec{\alpha} \tilde{\psi}_i^{\varepsilon_1} \right\} \tag{6.66}$$

allow to construct property densities [383], which may be visualized on a grid and integrated – giving the value of the corresponding property. Thus, the properties discussed here can be written as:

$$\sigma^K_{\alpha\beta} = -\frac{1}{c^2}\int \frac{1}{r_K^3}\left(\vec{r}_K \times \vec{j}^{B_\alpha}\right)_\beta d\vec{r}, \tag{6.67}$$

$$K^{KL}_{\alpha\beta} = -\frac{1}{c^2}\int \frac{1}{r_K^3}\left(\vec{r}_K \times \vec{j}^{m_{L;\alpha}}\right)_\beta d\vec{r}. \tag{6.68}$$

$$\xi_{\alpha\beta} = -\frac{1}{2}\int \left(\vec{r}_G \times \vec{j}^{B_\alpha}\right)_\beta d\vec{r}. \tag{6.69}$$

where the first-order current density perturbed by an external magnetic field is calculated with LAOs [384].

### NMR shieldings

The DC calculations of isotropic and anisotropic parts of NMR shielding tensor are summarized in Table 6.5. We present only the results for nuclei of the active subsystem ($H_2X$): the heavy center (X) and the hydrogen involved in the hydrogen bond ($H_b$). In the table we shall present the results in terms of the difference ($\Delta$) between the refence calculations on the dimers and the subsystem treatment, be it with FDE our without the inclusion of electronic contributions to embedding – we note that, since the structures have been optimized for the dimer (see paper D.7 for further details), the structures resemble in fact what would be closer to the outcome of a mechanical embedding calculation than of calculations on truly isolated subsystems. Furthermore, in the case of FDE calculations, we introduce additional notation in order to discern FDE-LAO contributions to the property gradient (NMR shielding and magnetizability) and to the expectation value (magnetizability), which can be either neglected (FDE[0]), limited to the embedding potential (FDE[$v$]) or to the embedding potential and the uncoupled kernel (FDE[$v+w_u$]) or also incorporating the coupling kernel ($w_c$) terms (FDE[$v+w_{\mathrm{all}}$]).

We see from our results for isotropic shieldings that FDE calculations generally outperform the calculations on the isolated subsystems. A striking trend for the shielding on the heavy elements is that the discrepancy between the dimer and FDE results decreases as the element becomes heavier, whereas the difference between the reference calculations and those for the isolated subsystems grows. Interestingly, the same is not seen for $H_b$, for which the agreement of both FDE and isolated calculations with the reference ones improves with increased atomic number of the heavy center.

For the anisotropy, on the other hand, if FDE still outperforms the calculations on isolated subsystems, its agreement with the reference is not very good and actually deteriorates as the systems become heavier. From the more detailed analysis of our results in paper D.7, we believe the key to further improve these results is in ameliorating the leading FDE-LAO contribution which arises from the inclusion of the embedding potential to the property gradient ($v$), and a possible way to do so is via the use of accurate approximations to the non-additive kinetic energy contributions, as recently done by Artiukhin for TDDFT [352].

These tendencies can be better seen in the plots of the differential isotropic shielding density in figures 6.8 for the heavy centers and 6.9 $H_b$, respectively. These figures exhibit positive (pink) and negative (blue) isosurfaces, which, respectively, depict more shielded and more deshielded areas in a dimer than in the embedded subsystems, and which upon integration give the corresponding values of $\Delta\sigma^{isol}_{iso}$.

We observe from figure 6.8 that the plots are rather similar for Se and Te nuclei, exhibiting small negative isosurfaces centered on a heavy nucleus, surrounded by much larger positive isosurface elongated on X-$H_b$ bond. In case of Po the differential shielding density is represented by much more complex isosurfaces around heavy center, as the negative isosurface centered on Po is larger than observed for Se and Te and surrounded by many well-separated positive lobes. This indicates that even though the property shift $\Delta^{isol}_{iso}$(Po) turns out to be relatively modest compared to the

Table 6.5: Absolute Dirac-Coulomb isotropic and anisotropic shielding values ($\sigma_{iso}^{super}$ and $\sigma_{aniso}^{super}$, in ppm) of nuclei in $H_2X$ (X = Se, Te, Po) subsystems in $H_2X$-$H_2O$, and absolute shifts ($\Delta\sigma$, in ppm) for the isolated ("isol") and embedded ("FDE") $H_2X$ molecules in the presence of $H_2O$. For FDE the values for different approximations in the FDE-LAO treatment ($c : [v + w_u]$, $d : [v + w_{\mathrm{all}}]$) are shown.

| Atom | $\sigma_{iso}^{super}$ | $\Delta\sigma_{iso[c]}^{FDE}$ | $\Delta\sigma_{iso[d]}^{FDE}$ | $\Delta\sigma_{iso}^{isol}$ | $\sigma_{aniso}^{super}$ | $\Delta\sigma_{aniso[c]}^{FDE}$ | $\Delta\sigma_{aniso[d]}^{FDE}$ | $\Delta\sigma_{aniso}^{isol}$ |
|---|---|---|---|---|---|---|---|---|
| Se | 2378.03 | -12.55 | -12.33 | 38.25 | 609.27 | -0.43 | -0.19 | 0.17 |
| $H_b$ | 30.88 | -0.67 | -0.62 | -2.09 | 22.19 | 5.76 | 5.70 | 7.29 |
| Te | 4667.85 | -9.16 | -8.88 | 67.48 | 1189.67 | -1.22 | -1.21 | 3.60 |
| $H_b$ | 35.62 | -0.44 | -0.42 | -1.70 | 14.59 | 0.24 | 0.28 | -1.31 |
| Po | 13985.80 | -3.52 | -3.13 | 137.84 | 5556.67 | -18.36 | -18.61 | -463.51 |
| $H_b$ | 40.80 | -0.09 | -0.09 | -0.57 | 105.25 | -1.67 | -1.67 | -6.01 |

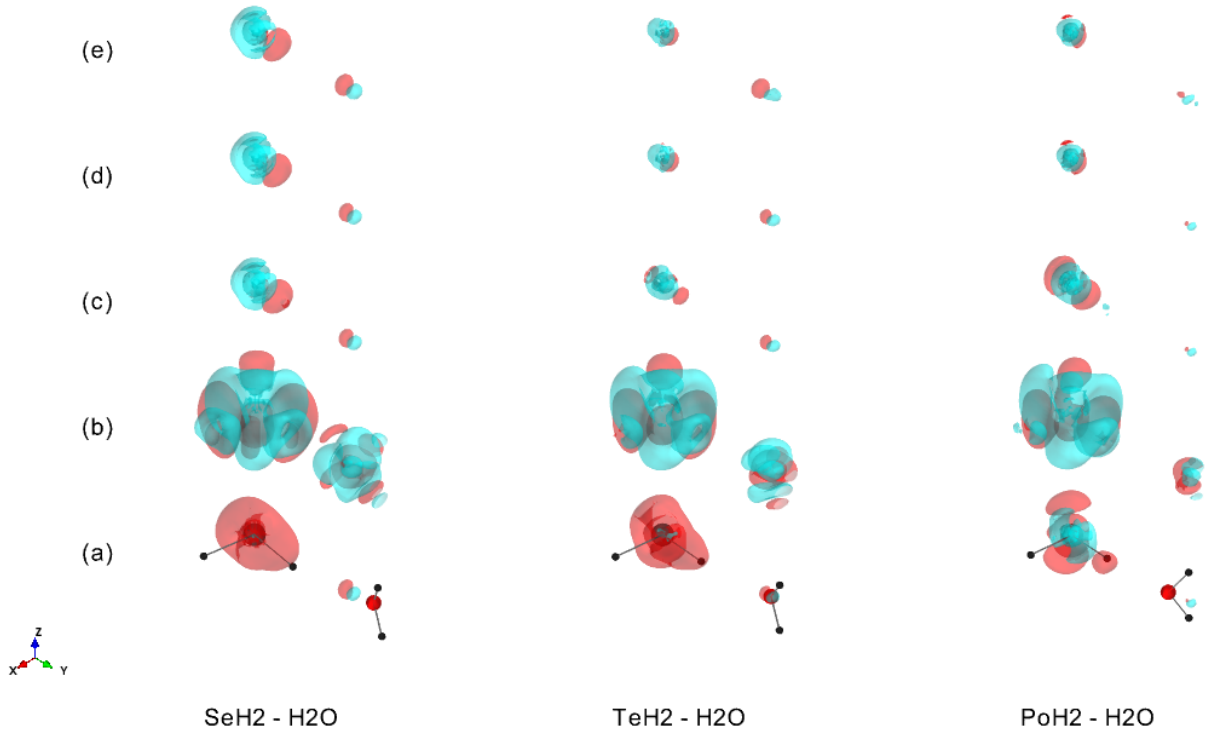



Figure 6.8: Differential isotropic shielding density isosurfaces (isovalues at +0.53 ppm (red) and -0.53 ppm (blue))for $XH_2$ - $H_2O$ systems for X = Se, Te, Po, calculated as a difference between dimer shielding densities and the sum of subsystem shielding densities approximated as: (a) isolated fragments (b) FDE[0](c) FDE[v] (d) FDE[v+$w_u$] (e) FDE[v+$w_{\mathrm{all}}$]. Color of atoms: X (blue), O(red), H(grey).

value of the absolute shielding, it is a result of shielded and deshielded areas cancelling out upon integration, reflecting the intricate interplay between environmental and relativistic effects, which are quite different for Te and Se.

Towards the heavier neighbouring center, hydrogen nuclei experience larger HALA effects [385], as reflected by increasing absolute values of $\sigma_{iso}^{super}(H_b)$ contributing to the shielding of $H_b$ nuclei in figure 6.9 and competing with the deshielding effect caused by the interaction with water molecule. We also observe the difference in the non hydrogen bonded hydrogen shielding between $PoH_2$ and the other species, which could also be a consequence of the HALA effect.

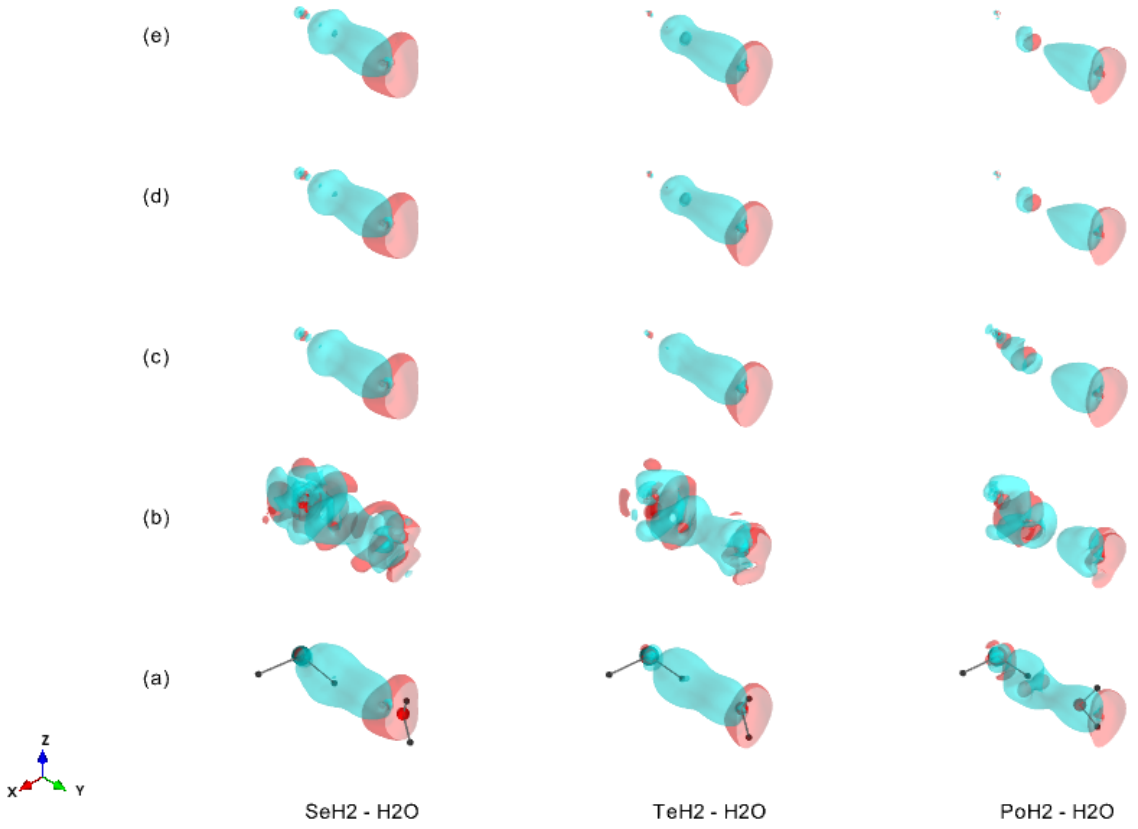



Figure 6.9: Differential isotropic shielding density isosurfaces (isovalues at +0.53 ppm (red) and -0.53 ppm (blue))for $H_b$ in the $XH_2$ - $H_2O$ systems, calculated as a difference between dimer shielding densities and the sum of subsystem shielding densities approximated as: (a) isolated fragments (b) FDE[0](c) FDE[v] (d) FDE[v+$w_u$] (e) FDE[v+$w_{all}$]. Color of atoms: X (blue), O(red), H(grey).

### Magnetizabilities

The magnetizability tensors calculated with the DC Hamiltonian are summarized in Table 6.6, where we present the isotropic ($\xi_{iso}$) and first anisotropy ($\xi_{aniso1}$) values. Due to it being an extensive property, in the case of subsystem calculations the total tensors are obtained as the sum of tensors for each subsystem.

An interesting aspect of magnetizabilities is that its value in a molecule can be very well approximated by a sum of contributing atomic susceptibilities [386–388], with only few exceptions [386, 389] The approximate additivity of magnetizability tensor, known as the Pascal's rule [390, 391], has been attributed to the local diamagnetic nature of atoms in molecules [392, 393] and the breakdown of this rule to the *"long-range circulation of electrons"* not accounted for in the atomic picture [392]. Recent studies have shown that Pascal's rule is particularly useful when analysed in terms of the magnetically induced current density, as the interaction of induced currents in neighbouring molecules and the increase of paramagnetic component of magnetizability tensor can be connected to the breakdown of the additivity rule.

As can be seen in Table 6.6, the $\Delta\xi_{iso}^{isol}$ values are small along the series, so that one could consider that Pascal's rule holds rather well here, whereas for anisotropies we see an increase in $\Delta\xi_{aniso1}^{isol}$ with increased atomic number of the heavy center. Moreover, the latter change sign along the series, indicating a complex interplay between relativistic and environmental effects. FDE-LAO contributions are also present in the property gradient for magnetizabilities, and for them we have found that the inclusion of FDE-LAO kernel contributions are even more important than for NMR shieldings in bringing about a good agreement with the reference calculations. We observe that for Se FDE is outperformed by the isolated calculations, but for Te and in particular for Po, FDE is much

Table 6.6: Dirac-Coulomb isotropic and first anisotropic magnetizabilities ($\xi_{iso}$ and $\xi_{aniso1}$, in SI units) and absolute shifts ($\Delta\sigma$, in SI) for the $H_2X$-$H_2O$ (X = Se, Te, Po) systems as well as for the $H_2X$ and $H_2O$ subsystems, the latter as isolated species ("isol") or via FDE calculations. In the case of subsystem calculations the total $\xi_{iso}$ and $\xi_{aniso1}$ were obtained as the sum of the subsystem values. For FDE the values for the different approximations in the FDE-LAO treatment ($c : [v + w_u]$, $d : [v + w_{\mathrm{all}}]$) are shown.

| System | $\xi_{iso}^{super}$ | $\Delta\xi_{iso[c]}^{FDE}$ | $\Delta\xi_{iso[d]}^{FDE}$ | $\Delta\xi_{iso}^{isol}$ | $\xi_{aniso1}^{super}$ | $\Delta\xi_{aniso1[c]}^{FDE}$ | $\Delta\xi_{aniso1[d]}^{FDE}$ | $\Delta\xi_{aniso1}^{isol}$ |
|---|---|---|---|---|---|---|---|---|
| $SeH_2$-$H_2O$ | -836.26 | 4.06 | 4.04 | 0.05 | -57.94 | -6.31 | -6.28 | -5.47 |
| $TeH_2$-$H_2O$ | -1080.67 | 1.74 | 1.66 | -4.28 | -81.63 | 4.19 | 4.39 | 4.73 |
| $PoH_2$-$H_2O$ | -1184.04 | -0.13 | -1.15 | -9.92 | -89.81 | 8.22 | 9.21 | 8.71 |

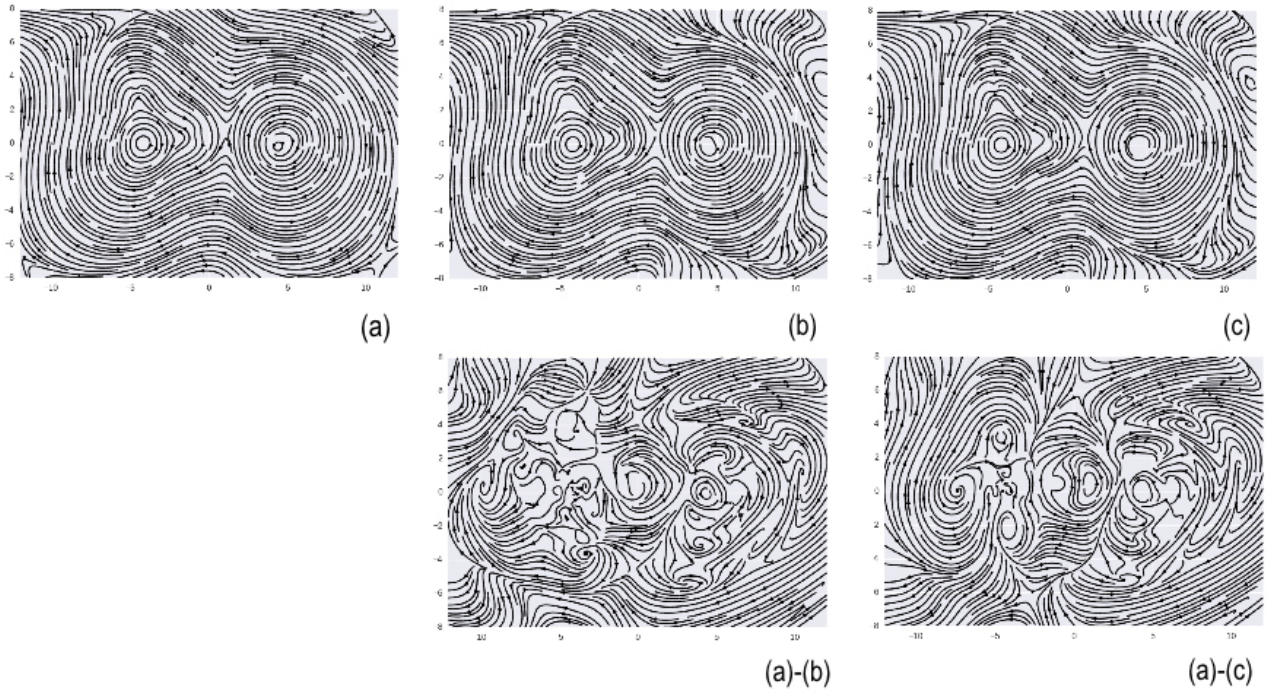



Figure 6.10: Streamline plots (xy plane, z = 0) of magnetically induced current densities for (a) reference calculations on the dimer; (b) subsystem calculations with FDE; and (c) calculations on isolated subsystems. The coordinates of Po and O are approximately (-5,0) and (5,0), respectively.

closer to the reference, even though the inclusion coupling kernel term for Po makes the agreement worse.

The good agreement between reference and subsystem magnetizabilities can be perhaps better understood by comparing the streamline plots for the total magnetically induced currents (shown in figure 6.10 for the Po system, on the plane containing the H-bond) since, from Eq. 6.69, the magnetizability is proportional to the induced current. There, we have a hard time to visually distinguish the difference in current between the reference calculations, the FDE-LAO and the isolated subsystem ones. If, on the other hand, we plot the difference in current between the reference and each of the subsystems, we see that in the vicinity of the molecules, the plot corresponding to "reference - FDE" shows a larger area where there are no currents (and therefore the FDE and reference one are nearly the same), whereas the plot of "reference - isolated" such a cancellation is less marked. That said, there is still a significant portion of the plots where there appears to be significant differences between the reference and subsystem calculations.

This preliminary analysis makes it clear, in my view, that in order to gain a better understanding of the behavior of subsystem calculations of magnetic properties, it is crucial to analyze the induced

currents and property densities as done here. In the case of the former, we should follow in the footsteps of Kiewisch [300] and perform a topological analysis of the induced current, as preconized by Bader [388] and Lazzeretti [394, 395].

# Chapter 7

# Perspectives

The previous chapters have underscored the successes of relativistic multireference and single-reference correlation approaches to determine electronic spectra and properties for heavy element systems in the gas-phase and, with the aid of FDE, in condensed phases as well. In the discussion some of the challenges faced by these approaches have also become clear. It is my intention to outline the actions I intend to take, over the course of the next years, to address these challenges.

## 7.1 The treatment of electron correlation for open-shell ground and excited states

The (IH)FSCCSD approach has been shown to yield very accurate results for vertical and adiabatic excitation energies, electron affinities, ionization potentials and equilibrium bond lengths – though vibrational frequencies tend to be overestimated – provided suitable model spaces can be selected.

A first important shortcoming of this approach has to do with convergence difficulties arising from the presence of intruder states that often manifest themselves despite the use of the intermediate Hamiltonian formulation. This makes (IH)FSCC a cumbersome method to use in exploring potential energy curves/surfaces (as attested in the case of the $I_3^-$, and of actinyl species) or for investigating states arising from excitations of subvalence or innermost electrons (as discussed for TsO or the species isoelectronic to the bare uranyl).

A second shortcoming has to do with the current (IH)FSCC implementation in Dirac, which can treat states that differ from the closed-shell reference by the addition/removal of up to two quasiparticles. This has two undesirable consequences: it renders a great deal of lanthanide, actinide and transition metal species exhibiting three or more unpaired electrons out of reach ($PuO_3$ being the latest example I encountered), but also does not allow for relaxation of configurations that belong to different Fock-space sectors as it would be necessary to describe $f^6$ or $f^8$ species, which can be considered relativistic closed shells (such as the Eu(III) ion or the Pu atom).

Third, employing a true multireference approach such as (IH)FSCCSD to investigate molecules whose ground-state is well-represented by a closed-shell (such as uranyl (VI) species) has turned out in practice to be rather costly in terms of computer resources (memory and disk usage) and human intervention (setting up model spaces, controlling convergence). In these circumstances, approaches such as EOM-CC offer the possibility of doing away with these difficulties with only a modest (if any) decrease in accuracy. The same can be said of some radical species (such as the XO and $I_3$ species), whose ground and excited states can be reached by electron attachment or electron detachment calculations from a closed-shell reference.

In view of these, I believe it is time to pursue a more flexible strategy in Dirac to address coupled

cluster calculations for excited states of both open and closed-shell species. In my future research I therefore plan to follow two directions: first, to pursue the development of EOM-CC approaches covering the sectors of Fock space currently covered by the (IH)FSCC implementation. Second, to focus on expanding the current (IH)FSCC implementation so that higher and mixed sectors [237, 396] of Fock space (for example, $(2,1),(2,2)$ or $(3,1)$) can be reached.

My main focus in the short term will be to continue working on EOM-CC approaches, to capitalize on our recent implementation of EOMCC-EE/EA/IP excitation energies and first-order properties. The discussion in chapter 5 shows that we can treat ionized states with the same accuracy as (IH)FSCC, and the results of our study on $UO_2^{2+}$ indicate EOM-CC should provide a very good approximation to the $(1,1)$ sector. One interesting extension to the EOM code would be the implementation of EA/IP approaches to treat ionized states corresponding to the $(0,2)$ and $(2,0)$ (or higher) sectors. These approaches have been implemented for non-relativistic Hamiltonians and appear to be sufficiently accurate[81–83, 397] and it would be quite interesting to see how these compare to (IH)FSCC. Even if they are only reliable enough to obtain semi-quantitative results, I'd still consider their availability to be of great value by eliminating the need of setting up model spaces and investigating convergence issues due to intruder states. Another development that may be potentially useful to improve the accuracy for singly excited energies is the implementation of the similarity-transformed EOM-CC (STEOM-CC) approach, which aims to bring EOM-CC closer to FSCC by carrying out a second similarity transformation on the Hamiltonian. I believe these efforts on the EOM-CC implementation will be useful in setting up the infrastructure for attacking the implementation of higher and mixed sectors of Fock space.

Prior to that, however, I am interested in addressing a very important aspect currently lacking for excited states in DIRAC: the inclusion of triples (and higher) corrections. For higher sectors of Fock Space previous studies on atomic systems have shown these are essential to retain the accuracy [398] found for lower sectors with the CCSD approach. Furthermore, their absence for the lower sectors prevent us from estimating the relative importance of higher-order correlation effects on our results (for the energies), and may well be the missing ingredient to obtaining vibrational frequencies in good agreement with experiment. The inclusion of triple excitations is a rather challenging aspect, due to their steep computational cost, scaling as $N^7$ for non-iterative approaches and $N^8$ for iterative ones.

For EOM-CC I plan to start exploring noniterative schemes such as the Λ-(EOM-)CCSD(T) approaches introduced by Bartlett and coworkers [77] which is closely related to the usual CCSD(T) method and is applicable both to the single-reference CC and EOM-CC cases (and seems to correct for the wrong dissociation behavior of CCSD(T) for a closed-shell reference), and other approaches (see [399] and references therein) for correcting EOM-IP states. In the long run, however, it will be important to have an implementation of iterative triples, since these would be necessary for the Fock-space implementation of higher sectors and, as we shall discuss below, for the determination of molecular properties.

Finally, while I believe coupled cluster approaches will remain the methods of choice for accurately determining the electronic structure of species containing one or two heavy centers with partially filled $f, d$ shells, this will likely not be the case for systems containing several such centers, as it is the case polyoxo and other species currently receiving a fair amount of interest in the research on magnetic materials or otherwise. Thus, it will be important in the coming years to closely follow the development of approaches such as DMRG in the relativistic domain [400] which can also include dynamical electron correlation in an efficient manner. I intend to closely follow these developments in the near future through the thesis of Sophie Kervazo, which I co-supervise with Valérie Vallet and Florent Réal.

## 7.2 The calculation of molecular properties with coupled cluster approaches

A second major shortcoming in coupled cluster implementations in the DIRAC code that requires attention is the currently limited capability of determining molecular properties other than energies. On a practical level that has prevented us, for instance, from calculating oscillator strengths for the studies carried out with (IH)FSCCSD. But that also means that one misses a unique opportunity to investigate properties for complex multireference ground states.

With the current EOM-CC implementation we are taking the first steps towards bridging this gap since, as discussed before, once we obtain the left eigenvectors of the similarity transformed Hamiltonian we have, in effect, the zeroth-order Lagrange multipliers required to define first-order properties for the different excited states. While this is an interesting by itself, it also open interesting perspectives in the context of FDE as it will be discussed below.

Our next goal within the EOM-CC framework (EOM-IP/EA/EE) is to lay out the framework to calculate second-order properties arising from electric perturbations, such as electric polarizabilities, which require at the very least the determination of first-order perturbed cluster amplitudes via the solution of linear response equations. If one does not take into account orbital relaxation, via the solution of CPHF equations, such an implementation requires relatively small additions to the infrastructure put in place to obtain excitation energies, and I expect to be able to achieve it in a relatively short time. Such a development will also contribute to other activities within the group, as accurate polarizability values are of great importance to the development of polarizable force-fields for heavy elements [36, 401].

For electric properties, the neglect of orbital relaxation has shown to be of relatively minor importance, since the $T_1$ amplitudes introduce a good deal of orbital relaxation. However, for properties arising from magnetic perturbation this is not a good approximation and the solution of CPHF equations must be considered, especially if London orbitals are used. In the case of second-order magnetic properties, which are currently of great interest to me in the context of FDE developments and applications (as it will be discussed below), an additional complication in relativistic calculations is that the diamagnetic contributions arise from rotations between electronic and positronic solutions, and therefore such a relaxation must also be taken into account.

The proper solution to the question of orbital relaxation in general and specifically for the case of magnetic properties will require significant efforts in terms of developing the formalism and, more importantly, putting in place an efficient implementation. Such an effort is being discussed in a collective manner within the community of DIRAC developers to the coupled cluster model. I intend to contribute to these efforts.

Apart from the activities mentioned above, which by and large focus on the EOM approach, I believe it is important to construct a framework with which we can also implement properties for (IH)FSCC wavefunctions, in the mid- to long-term, but keeping in mind the strategy outlined above and focus on using Fock-space whenever one stands to have a significant advantage over the EOM implementation(s).

## 7.3 Generalizing molecular properties for DFT-in-DFT and CC-in-DFT embedding

FDE is an essential ingredient in my efforts to provide tools for the accurate modeling of molecular properties of heavy element species in the condensed phase FDE, given its fully quantum-mechanical nature and the seamless way with which it lets us combine DFT and coupled cluster approaches.

With the recent implementation of second-order magnetic properties at DFT-in-DFT level in paper D.7, one clear short-term goal to me is to apply this implementation to systems of experimental interest, even if intermediate benchmark studies will still be important due to the relative little ex-

perience the community has in calculating magnetic properties with FDE. A particularly important aspect in this exploratory phase will be to further develop the tools to visualize and analyze magnetically induced currents, characterize how they differ between FDE and conventional DFT calculations, and how such information could be used to devise better subspace partitioning schemes or, eventually improve kinetic energy density functionals.

Looking beyond that, one can see that so far developments of FDE within the framework of response have mostly focused on the linear regime, be it in the time-dependent or time-independent cases, given the importance of second-order properties. This focus, however, has left something of a blind spot in most if not all implementations, which are not sufficiently flexible to explore other, equally interesting properties that can formally be associated with lighter orders in the response treatment – a notable example is the two-photon absorption (TPA) cross-sections, which can be extracted from quadratic response calculations – or even with second-order properties arising from mixing electric and magnetic perturbations.

Dirac offers us a unique opportunity in that the response code is sufficiently general, and should currently allow us mix magnetic and electric perturbations for second-order properties (though so far no such calculations have been attempted) as well as to perform up to quadratic response calculations. As such, I am very interested in attempting to formulate and implement FDE quadratic response in order to obtain TPA cross-sections, as well as excited state properties such as dipole moments, first at a purely DFT level (DFT-in-DFT) so that we can gain experience both in implementation and in applications.

In order to tackle a quadratic response DFT-in-DFT implementation in Dirac, or even venture in calculations involving mixed perturbations one must, however, first remedy the current limitation in the FDE response implementation, as it does not allow us to calculate the coupling of the subsystems' response to the perturbation (the off-diagonal blocks of the electronic Hessian in Eq. 6.26). Removing this limitation will already be quite interesting, as we shall be able to calculate electric polarizabilities or other second-order extensive properties in a subsystem manner, and will eliminate any handicaps in future developments.

If these DFT-in-DFT developments are quite important in themselves, one should not loose from sight that our ultimate goal is to develop CC-in-DFT approaches. As has been the case for DFT, I see tremendous potential for applications of CC-in-DFT for the calculation of magnetic properties such as NMR shieldings or indirect spin-spin couplings, provided the FDE description is sufficiently accurate to justify the additional cost of performing coupled cluster calculations, as opposed to simply performing standard DFT calculations on an extended system.

The implementation of CC-in-DFT in Dirac is in a much less advanced stage than for DFT-in-DFT, so here I shall focus first on lifting the constraint of only performing calculations with the so-called "static" potential approach outline in section 6.2 and paper D.1 by implementing the second-order (kernel) contributions arising from the response of the embedding potential, since now with the EOM-CC implementation we have the CC Jacobian at our disposal. The ability of describing the coupling of the subsystems' response to the perturbation will be of great importance for any further extensions, so I shall focus on that after the intra-subsystem kernel contributions are implemented.

The extension of the FDE CC-in-DFT response implementation to magnetic properties, though not strictly requiring the description of the coupling contributions, hinges nevertheless on the progress of the standard implementation of coupled cluster NMR properties, and cannot be envisaged for the immediate future.

## 7.4 Improving the accuracy of embedding approaches for strongly interacting systems

Most FDE calculations are performed with approximate kinetic energy functionals since these are very efficient from a computational standpoint and are often sufficiently accurate to describe relatively weak

interactions such as hydrogen bonds – though as the example of our $CUO(Ng)_4$ study attests D.4, it can fail even in such weak interaction cases.

However, if FDE is to become a generally applicable approach, one must be able to have more freedom in fragmenting the total system, including by cutting through covalent bonds, without a significant worsening of accuracy. While it would be ideal, from the computational point of view, to continue to rely on density functionals for obtaining the non-additive kinetic energy contributions to the interaction energy, their relatively slow pace of development means that the alternative of approximating the derivative of the non-additive kinetic energy contributions directly with OEP schemes is likely the most efficient route to improve accuracy.

I am currently working with Christoph Jacob and Claudia Filippi on a practical realization of such a potential reconstruction approach, and its application to WFT-in-DFT calculations. While our results are promising, we have so far only investigated light systems and much can be improved in terms of efficiency and modularity for its implementation. One aspect we wish to explore in the near future is the construction of unique embedding potentials, in the spirit of [267, 402] since these would greatly facilitate calculations with accurate potentials for several subsystems since they'd all share the same reconstructed potential A second aspect is the scheme's performance for heavy systems, and I believe it would be important to revisit the $CUO(Ng)_4$ case to see whether such approach would also perform well for systems containing heavy elements.

On a practical level, one thing I intend to do is to work towards having a stand-alone version of the code currently in use (which is strongly connected to the ADF suite) so that it can serve as a platform for future developments as well as to be used by anyone that does not have access to ADF. From there, a next step would be to verify the performance of such potentials in the calculation of molecular properties, both at WFT-in-DFT and DFT-in-DFT. There, I would be particularly interested to see whether such approaches can improve in any way our results for second-order magnetic properties for lighter atoms, where we observed for instance a relatively poor performance for FDE in comparison to heavy elements for NMR shieldings, as was the case for TDDFT excitation energies [352]. If results are again promising, a next step would be to also reconstruct the derivative of the reconstructed potential in order to see whether second-order (kernel) contributions to molecular properties, and consequently the subsystem response coupling, would also be improved.

## 7.5 Combining a molecular description of properties with a periodic description of the environment

All of the approaches I've developed and so far discussed in this manuscript have in common the fact that they have been formulated and implemented for discrete molecular species, and when applied to systems in the condensed phase, such as the $Cs_2UO_2Cl_4$ crystal D.6 or for species in solution, we have introduced approximations to the description of the environment beyond a certain distance from the subsystem of interest.

These approximations include adding point charges surrounding the QM region treated with FDE in order to properly represent the long-range electrostatic interactions with the bulk, or even disregarding any long-range effects in property calculations and considering only structures determined in the presence of such interactions as it is often the case of FDE calculations of solvent using snapshots from molecular dynamics results.

While such embedded cluster approaches yield good results for localized electronic excitations, one may wonder whether they would be equally effective in the case of molecules interacting with extended, non-ionic materials – not to mention the fact that, from a practical perspective, the determination of point charges for embedded clusters is somewhat cumbersome if one wishes to explore different geometries for the molecule and interacting surfaces, or study the electronic relaxation of the surroundings upon processes such as ionization.

I believe a more satisfying approach would be to make use of FDE to connect solid-state[326, 327]

and molecular codes, as I'm currently pursuing in collaboration with the group of Prof. Pavanello at Rutgers University Newark. We so far have only carried out preliminary assessments on light systems, but are currently revisiting the ionization potential of uranyl chloride in the $Cs_2UO_2Cl_4$ system, in order to compare to our embedded cluster results.

My goal with this approach is to attempt to devise a computational protocol based on the combination of periodic FDE calculations to obtain structures and molecular FDE calculations to obtain molecular properties, and to apply it to the investigation of NMR spectra of heavy centers for molecules adsorbed on surfaces. In this project corresponds as well to the PhD thesis of Yassine Bouchafra, which started in October 2016.

Once such a protocol is validated, I intend to explore its use in other challenging cases, such as in determining the magnetic properties and electronic spectra of polymetallic $Mo_6X_4^{2-}$ (X = halides) species in solution [403]. These species, important as building blocks of different materials, still pose a challenge to theory due to the size of the systems in the molecular dynamics and property calculation steps, and I believe these difficulties can be surmounted with subsystem approaches.

## Bibliography

[1] I. McGill, *Rare Earth Elements* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[2] G. Winstel, *Electroluminescent Materials and Devices* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[3] T. Jüstel, S. Möller, H. Winkler and W. Adam, *Luminescent Materials* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[4] Radiometals for Combined Imaging and Therapy. C. S. Cutler, H. M. Hennkens, N. Sisay, S. Huclier-Markai and S. S. Jurisson. *Chem. Rev.* **113**, 858–883 (2013).

[5] Enigmatic astatine. D. S. Wilbur. *Nature Chem.* **5**, 246–246 (2013).

[6] 211At-labeled agents for alpha-immunotherapy: On the in vivo stability of astatine-agent bonds. T. Ayed, J. Pilmé, D. Tézé, F. Bassal, J. Barbet, M. Chérel, J. Champion, R. Maurice, G. Montavon and N. Galland. *European Journal of Medicinal Chemistry* pp. 1–38 (2016).

[7] O. Deutschmann, H. Knözinger, K. Kochloefl and T. Turek, *Heterogeneous Catalysis and Solid Catalysts* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[8] Recent Developments in the Synthesis of Supported Catalysts. P. Munnik, P. E. de Jongh and K. P. de Jong. *Chem. Rev.* **115**, 6687–6718 (2015).

[9] W. Stoll, *Thorium and Thorium Compounds* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[10] M. Peehs, T. Walter, S. Walter and M. Zemek, *Uranium, Uranium Alloys, and Uranium Compounds* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[11] J.-P. Glatz and R. J. M. Konings, *Plutonium and Other Transuranium Elements* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2000).

[12] Complexation and Extraction of Trivalent Actinides and Lanthanides by Triazinylpyridine N-Donor Ligands. P. J. Panak and A. Geist. *Chem. Rev.* **113**, 1199–1236 (2013).

[13] Solution and Solid-State Structural Chemistry of Actinide Hydrates and Their Hydrolysis and Condensation Products. K. E. Knope and L. Soderholm. *Chem. Rev.* **113**, 944–994 (2013).

[14] Recent Developments in Synthesis and Structural Chemistry of Nonaqueous Actinide Complexes. M. B. Jones and A. J. Gaunt. *Chem. Rev.* **113**, 1137–1198 (2013).

[15] Die Renaissance der nichtwässrigen Uranchemie. S. T. Liddle. *Angew. Chem.* **127**, 8726–8764 (2015).

[16] Yucca Mountain Looking ten thousand years into the future. R. C. Eckhardt. *Los Alamos Science* **26**, 464–489 (2000).

[17] Recent Advances in Aqueous Actinide Chemistry and Thermodynamics. M. Altmaier, X. Gaona and T. Fanghänel. *Chem. Rev.* **113**, 901–943 (2013).

[18] Mineral–Water Interface Reactions of Actinides. H. Geckeis, J. Lützenkirchen, R. Polly, T. Rabung and M. Schmidt. *Chem. Rev.* **113**, 1016–1062 (2013).

[19] Actinide Colloids and Particles of Environmental Concern. C. Walther and M. A. Denecke. *Chem. Rev.* **113**, 995–1015 (2013).

[20] ASTEC (Accident Source Term Evaluation Code) Software System, developed by the French Institute for Radiological Protection and Nuclear Safety (IRSN).

[21] Advances in Correlated Electronic Structure Methods for Solids, Surfaces, and Nanostructures. P. Huang and E. A. Carter. *Annu. Rev. Phys. Chem.* **59**, 261–290 (2008).

[22] Embedded Correlated Wavefunction Schemes: Theory and Applications. F. Libisch, C. Huang and E. A. Carter. *Acc. Chem. Res.* **47**, 2768–2775 (2014).

[23] Frozen–Density Embedding Strategy for Multilevel Simulations of Electronic Structure. T. A. Wesolowski, S. Shedge and X. Zhou. *Chem. Rev.* **115**, 5891–5928 (2015).

[24] Subsystem density-functional theory. C. R. Jacob and J. Neugebauer. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **4**, 325–362 (2014).

[25] Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems. A. S. P. Gomes and C. R. Jacob. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **108**, 222–277 (2012).

[26] Electronic Structure and Bonding in Actinyl Ions and their Analogs. R. G. Denning. *J. Phys. Chem. A* **111**, 4125 (2007).

[27] The Chemical Complexities of Plutonium. D. L. Clark. *Los Alamos Science* **26**, 364–381 (2000).

[28] The Chemical Interactions of Actinides in the Environment. W. Runde. *Los Alamos Science* **26**, 392–415 (2000).

[29] Uranyl Bearing Hybrid Materials: Synthesis, Speciation, and Solid-State Structures. M. B. Andrews and C. L. Cahill. *Chem. Rev.* **113**, 1121–1136 (2013).

[30] Clusters of Actinides with Oxide, Peroxide, or Hydroxide Bridges. J. Qiu and P. C. Burns. *Chem. Rev.* **113**, 1097–1120 (2013).

[31] Relativistic Effects in Chemistry: More Common Than You Thought. P. Pyykkö. *Ann. Rev. Phys. Chem* **63**, 45–64 (2012).

[32] Relativistic Hamiltonians for chemistry: a primer. T. Saue. *ChemPhysChem* **12**, 3077–3094 (2011).

[33] K. G. Dyall and K. Fægri, *Introduction to relativistic quantum chemistry* (Oxford University Press, New York, 2007).

[34] M. Reiher and A. Wolf, *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science* (WILEY-VCH Verlag, 2009).

[35] The Physics behind Chemistry and the Periodic Table. P. Pyykkö. *Chem. Rev.* **112**, 371–384 (2012).

[36] Structural, dynamical, and transport properties of the hydrated halides: How do $At^-$ bulk properties compare with those of the other halides, from $F^-$ to $I^-$ ? F. Réal, A. Severo Pereira Gomes, Y. O. Guerrero Martínez, T. Ayed, N. Galland, M. Masella and V. Vallet. *J. Chem. Phys.* **144**, 124513–13 (2016).

[37] Relativistic effects in homogeneous gold catalysis. D. J. Gorin and F. D. Toste. *Nature* **446**, 395–403 (2007).

[38] Mercury Methylation by Cobalt Corrinoids: Relativistic Effects Dictate the Reaction Mechanism. T. B. Demissie, B. D. Garabato, K. Ruud and P. M. Kozlowski. *Angew. Chem.* **128**, 11675–11678 (2016).

[39] Relativity and the Lead-Acid Battery. R. Ahuja, A. Blomqvist, P. Larsson, P. Pyykkö and P. Zaleski-Ejgierd. *Phys. Rev. Lett.* **106**, 018301–4 (2011).

[40] J. Olsen and P. Jorgensen, Time-Dependent Response Theory with Applications to Self-Consistent Field and Multiconfiguration Self-Consistent Field Wave Functions, in *Modern Electronic Structure Theory - Part 2*, ed.D. R. Yarkony, pp. 857–990 (World Scientific, Singapore, 1995).

[41] Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. O. Christiansen, P. Jørgensen and C. Hättig. *Int. J. Quantum Chem.* **68**, 1–52 (1998).

[42] D. J. Tannor, *Introduction to Quantum Mechanics: A Time-Dependent Perspective* (University Science Books, 2006).

[43] Recent advances in wave function-based methods of molecular-property calculations. T. Helgaker, S. Coriani, P. Jorgensen, K. Kristensen, J. Olsen and K. Ruud. *Chem. Rev.* **112**, 543–631 (2012).

[44] Inhomogeneous Electron Gas. P. Hohenberg and W. Kohn. *Phys. Rev.* **136**, B864 (1964).

[45] Self-Consistent Equations Including Exchange and Correlation Effects. W. Kohn and L. J. Sham. *Phys. Rev.* **140**, A1133 (1965).

[46] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed. (Wiley-VCH, Weinheim, 2001).

[47] On the accuracy of density functional theory in transition metal chemistry. J. N. Harvey. *Annu. Rep. Prog. Chem., Sect. C* **102**, 203–226 (2006).

[48] D. Rappoport, N. R. M. Crawford, F. Furche and K. Burke, Approximate Density Functionals: Which Should I Choose?, in *Computational Inorganic and Bioinorganic Chemistry*, eds.E. I. Solomon, R. A. Scott and R. B. King, pp. 159–172 (Wiley, 2009).

[49] Insights into Current Limitations of Density Functional Theory. A. J. Cohen, P. Mori-Sanchez and W. Yang. *Science* **321**, 792–794 (2008).

[50] Challenges for Density Functional Theory. A. J. Cohen, P. Mori-Sánchez and W. Yang. *Chem. Rev.* **112**, 289–320 (2012).

[51] Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. O. Gunnarsson and B. I. Lundqvist. *Phys. Rev. B* **13**, 4274 (1976).

[52] Local-density theory of multiplet structure. U. von Barth. *Phys. Rev. A* **20**, 1693 (1979).

[53] Some Fundamental Issues in Ground-State Density Functional Theory: A Guide for the Perplexed. J. P. Perdew, A. Ruzsinszky, L. A. Constantin, J. Sun and G. I. Csonka. *J. Chem. Theory Comput.* **5**, 902–908 (2009).

[54] Lack of Hohenberg-Kohn Theorem for Excited States. R. Gaudoin and K. Burke. *Phys. Rev. Lett.* **93**, 173001 (2004).

[55] On the calculation of multiplet energies by the hartree-fock-slater method. T. Ziegler, A. Rauk and E. J. Baerends. *Theor. Chim. Acta* **43**, 261–271 (1977).

[56] Density functional theory applied to the excited states of coordination compounds. C. Daul. *Int. J. Quantum Chem.* **52**, 867–877 (1994).

[57] Variational Density-Functional Theory for an Individual Excited State. M. Levy and A. Nagy. *Phys. Rev. Lett.* **83**, 4361–4364 (1999).

[58] Self-Consistent Field Calculations of Excited States Using the Maximum Overlap Method (MOM). A. T. B. Gilbert, N. A. Besley and P. M. W. Gill. *J. Phys. Chem. A* **112**, 13164–13171 (2008).

[59] Rydberg energies using excited state density functional theory. C.-L. Cheng, Q. Wu and T. Van Voorhis. *J. Chem. Phys.* **129**, 124112 (2008).

[60] Assessment of the ?SCF density functional theory approach for electronic excitations in organic dyes. T. Kowalczyk, S. R. Yost and T. V. Voorhis. *J. Chem. Phys.* **134**, 054128 (2011).

[61] On the relation between time-dependent and variational density functional theory approaches for the determination of excitation energies and transition moments. T. Ziegler, M. Seth, M. Krykunov, J. Autschbach and F. Wang. *J. Chem. Phys.* **130**, 154102 (2009).

[62] The formulation of a self-consistent constricted variational density functional theory for the description of excited states. J. Cullen, M. Krykunov and T. Ziegler. *Chem. Phys.* **391**, 11–18 (2011).

[63] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, Chichester, 2000).

[64] Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications. R. G. Szalay, T. M. an G. Gidofalvi, H. Lischka and R. Shepard. *Chem. Rev.* **112**, 108–181 (2012).

[65] Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules. A. Dreuw and M. Head-Gordon. *Chem. Rev.* **105**, 4009–4037 (2005).

[66] Spectroscopy of low-coordinated surface sites: Theoretical study of MgO. A. L. Shluger, P. V. Sushko and L. N. Kantorovich. *Phys. Rev. B* **59**, 2417–2430 (1999).

[67] Electronic structure of excited states at low-coordinated surface sites of MgO. P. V. Sushko and A. L. Shluger. *Surf. Sci.* **421**, L157–L165 (1999).

[68] Assessment of TD-DFT methods and of various spin scaled CIS(D) and CC2 versions for the treatment of low-lying valence excitations of large organic dyes. L. Goerigk and S. Grimme. *J. Chem. Phys.* **132**, 184103 (2010).

[69] Benchmarking Second Order Methods for the Calculation of Vertical Electronic Excitation Energies: Valence and Rydberg States in Polycyclic Aromatic Hydrocarbons. H. H. Falden, K. R. Falster-Hansen, K. L. Bak, S. Rettrup and S. A. Sauer. *J. Phys. Chem. A* **113**, 11995 (2009).

[70] A doubles correction to electronic excited states from configuration interaction in the space of single substitutions. M. Head-Gordon, R. J. Rico, M. Oumi and T. J. Lee. *Chem. Phys. Lett.* **219**, 21–29 (1994).

[71] A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. B. O. Roos, P. R. Taylor and P. E. M. Siegbahn. *Chem. Phys.* **48**, 157 (1980).

[72] Assessing Metal–Metal Multiple Bonds in Cr-Cr, Mo-Mo, and W-W Compounds and a Hypothetical U-U Compound: A Quantum Chemical Study Comparing DFT and Multireference Methods. G. Li Manni, A. L. Dzubak, A. Mulla, D. W. Brogden, J. F. Berry and L. Gagliardi. *Chem.–Eur. J.* **18**, 1737–1749 (2012).

[73] The Density Matrix Renormalization Group in Quantum Chemistry. G. K.-L. Chan and S. Sharma. *Annu. Rev. Phys. Chem.* **62**, 465–481 (2011).

[74] New electron correlation theories for transition metal chemistry. K. H. Marti and M. Reiher. *Phys. Chem. Chem. Phys.* **13**, 6750–6759 (2011).

[75] Size extensive modification of local multireference configuration interaction. A. Venkatnathan, A. B. Szilva, D. Walter, R. J. Gdanitz and E. A. Carter. *J. Chem. Phys.* **120**, 1693 (2004).

[76] Coupled-cluster theory in quantum chemistry. R. J. Bartlett and M. Musial. *Rev. Mod. Phys.* **79**, 291–351 (2007).

[77] Coupled-cluster theory and its equation-of-motion extensions. R. J. Bartlett. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2**, 126–138 (2011).

[78] Excitation energies from the coupled cluster singles and doubles linear response function (CCSDLR). Applications to Be, $CH^+$, CO, and $H_2O$. H. Koch, H. J. A. Jensen, P. Jorgensen and T. Helgaker. *J. Chem. Phys.* **93**, 3345–3350 (1990).

[79] Calculation of size-intensive transition moments from the coupled cluster singles and doubles linear response function. H. Koch, R. Kobayashi, A. S. de Meras and P. Jorgensen. *J. Chem. Phys.* **100**, 4393–4400 (1994).

[80] Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space. A. I. Krylov. *Ann. Rev. Phys. Chem* **59**, 433–462 (2008).

[81] Doubly electron-attached and doubly ionized equation-of-motion coupled-cluster methods with 4-particle–2-hole and 4-hole–2-particle excitations and their active-space extensions. J. Shen and P. Piecuch. *J. Chem. Phys.* **138**, 194102–17 (2013).

[82] Doubly electron-attached and doubly ionised equation-of-motion coupled-cluster methods with full and active-space treatments of 4-particle–2-hole and 4-hole–2-particle excitations: the role of orbital choices. J. Shen and P. Piecuch. *Mol. Phys.* **112**, 868–885 (2014).

[83] The equation-of-motion coupled cluster method for triple electron attached states. M. Musiał, M. Olszówka, D. I. Lyakh and R. J. Bartlett. *J. Chem. Phys.* **137**, 174102–10 (2012).

[84] A new method for excited states: Similarity transformed equation-of-motion coupled-cluster theory. M. Nooijen and R. J. Bartlett. *J. Chem. Phys.* **106**, 6441–9 (1997).

[85] Similarity transformed equation-of-motion coupled-cluster study of ionized, electron attached, and excited states of free base porphin. M. Nooijen and R. J. Bartlett. *J. Chem. Phys.* **106**, 6449–8 (1997).

[86] Similarity transformed equation of motion coupled cluster theory revisited: a benchmark study of valence excited states. J. Sous, P. Goel and M. Nooijen. *Mol. Phys.* **0**, 1–23 (2013).

[87] M. Wladyslawski and M. Nooijen, Analytical Energy Gradients for Excited-State Coupled-Cluster Methods: Automated Algebraic Derivation of First Derivatives for Equation-of-Motion Coupled-Cluster and Similarity Transformed Equation-of-Motion Coupled-Cluster Theories, in *Advances in Quantum Chemistry Volume 49*, pp. 1–101 (Elsevier, 2005).

[88] Similarity transformed coupled cluster response (ST-CCR) theory - A time-dependent similarity transformed equation-of-motion coupled cluster (STEOM-CC) approach. A. Landau. *J. Chem. Phys.* **139**, 014110–19 (2013).

[89] Multireference Nature of Chemistry: The Coupled-Cluster View. D. I. Lyakh, M. Musial, V. F. Lotrich and R. J. Bartlett. *Chem. Rev.* **112**, 182–243 (2012).

[90] Multireference state-specific coupled-cluster methods. State-of-the-art and perspectives. V. V. Ivanov, D. I. Lyakh and L. Adamowicz. *Phys. Chem. Chem. Phys.* **11**, 2355–2370 (2009).

[91] A Fock space coupled cluster study on the electronic structure of the $UO_2$, $UO_2^+$, $U^{4+}$ and $U^{5+}$ species. I. Infante, E. Eliav, M. J. Vilkas, Y. Ishikawa, U. Kaldor and L. Visscher. *J. Chem. Phys.* **127**, 124308 (2007).

[92] Effects of the first hydration sphere and the bulk solvent on the spectra of the f2 isoelectronic actinide compounds: $U^{4+}$, $NpO_2^+$, and $PuO_2^{2+}$. C. Danilo, V. Vallet, J.-P. Flament and U. Wahlgren. *Phys. Chem. Chem. Phys.* **12**, 1116–1130 (2009).

[93] *Ab initio* theoretical study of the electronic structure of $UO_2{}^+$ and $[UO_2(CO_3)_3]^{5-}$. F. Ruipérez, C. Danilo, F. Réal, J.-P. Flament, V. Vallet and U. Wahlgren. *J. Phys. Chem. A* **113**, 1420–1428 (2009).

[94] Zero Field Splitting the chalcogen diatomics using relativistic correlated wave-function methods. J.-B. Rota, S. Knecht, T. Fleig, D. Ganyushin, T. Saue, F. Neese and H. Bolvin. *J. Chem. Phys.* **135**, 114106 (2011).

[95] Electronic Spectra and Ionization Potentials of Halogen Oxides Using the Fock Space Coupled-Cluster Method. N. Vaval, P. Manohar and S. Pal. *Collect. Czech. Chem. Commun.* **70**, 851–863 (2005).

[96] On the connectivity criteria in the open-shell coupled-cluster theory for general model spaces. I. Lindgren and D. Mukherjee. *Phys. Rev.* **151**, 93–127 (1987).

[97] Aspects of separability in the coupled cluster based direct methods for energy differences. D. Mukhopadhyay, S. Mukhopadhyay, R. Chaudhuri and D. Mukherjee. *Theor. Chim. Acta* **80**, 441–467 (1991).

[98] The Fock space coupled cluster method: theory and applications. U. Kaldor. *Theor. Chem. Acc.* **80**, 427–439 (1991).

[99] Fock-space coupled-cluster method in the intermediate Hamiltonian formulation: Model with singles and doubles. L. Meissner. *J. Chem. Phys.* **108**, 9227–9235 (1998).

[100] L. Meissner and M. Musiał, Intermediate Hamiltonian Formulations of the Fock-space coupled-cluster method: details, comparisons, examples, in *Recent Progress in Coupled Cluster Methods*, ed.P. Čársky, p. 395 (Springer, 2010).

[101] Intermediate Hamiltonian Fock-space coupled-cluster method: Excitation energies of barium and radium. A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor. *J. Chem. Phys.* **113**, 9905–9910 (2000).

[102] Intermediate Hamiltonian Fock-space coupled cluster method in the one-hole one-particle sector: Excitation energies of xenon and radon. A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor. *J. Chem. Phys.* **115**, 6862–6865 (2001).

[103] Intermediate Hamltonian Hilbert space coupled cluster method: Theory and pilot applications. E. Eliav, A. Borschevsky, K. R. Shamasundar, S. Pal and U. Kaldor. *Int. J. Quantum Chem.* **109**, 2909–2915 (2009).

[104] Second-order perturbation theory with a CASSCF reference function. K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski. *J. Phys. Chem.* **94**, 5483 (1990).

[105] Second-order perturbation theory with a complete active space self-consistent field reference function. K. Andersson, P.-A. Malmqvist and B. O. Roos. *J. Chem. Phys.* **96**, 1218 (1992).

[106] The multi-state CASPT2 method. J. Finley, P.-A. Malmqvist, B. O. Roos and L. Serrano-Andrés. *Chem. Phys. Lett.* **288**, 299–306 (1998).

[107] Similarities between single reference pertubation theory based on a CASSCF wavefunction and multireference perturbation theory based on a reference space spanned by a CAS. J. P. Finley. *Chem. Phys. Lett.* **283**, 277–282 (1998).

[108] Multiconfiguration perturbation theory with level shift — the $Cr_2$ potential revisited. B. O. Roos and K. Andersson. *Chem. Phys. Lett.* **245**, 215 (1995).

[109] Multiconfigurational perturbation theory with imaginary level shift. N. Forsberg and P.-Aa. Malmqvist. *Chem. Phys. Lett.* **274**, 196 (1997).

[110] A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). G. Ghigo, B. O. Roos and P.-A. Malmqvist. *Chem. Phys. Lett.* **396**, 142 (2004).

[111] Intruder State Avoidance Multireference Moller-Plesset Perturbation Theory. H. A. Witek, Y.-K. Choe, J. P. Finley and K. Hirao. *J. Comput. Chem.* **108**, 1081–1088 (1998).

[112] Introduction of n-electron valence states for multireference perturbation theory. C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger and J. P. Malrieu. *J. Chem. Phys.* **114**, 10252 (2001).

[113] Multiconfigurational Perturbation Theory: An Efficient Tool to Predict Magnetic Coupling Parameters in Biradicals, Molecular Complexes, and Ionic Insulators. C. de Graaf, C. Sousa, I. de P. R. Moreira and F. Illas. *J. Phys. Chem. A* **105**, 11371–11378 (2001).

[114] High-level ab initio calculations on the energetics of low-lying spin states of biologically relevant transition metal complexes: a first progress report. A. Ghosh and P. R. Taylor. *Curr. Opin. Chem. Biol.* **7**, 113–124 (2003).

[115] Ab initio study of the antiferromagnetic coupling in the wheel-shaped $[Cu_{20}Cl(OH)_{24}(H_2O)_{12}(P_8W_{48}O_{184})]^{25-}$ anion. C. de Graaf, X. López, J. L. Ramos and J. M. Poblet. *Phys. Chem. Chem. Phys.* **12**, 2716–2721 (2010).

[116] The CASPT2 method in inorganic electronic spectroscopy: from ionic transition metal to covalent actinide complexes. K. Pierloot. *Mol. Phys.* **101**, 2083 (2003).

[117] A coupled cluster study of the electronic spectroscopy and photochemistry of $Cr(CO)_6$. S. Villaume, A. Strich, C. Daniel, S. A. Perera and R. J. Bartlett. *Phys. Chem. Chem. Phys.* **9**, 6115 (2007).

[118] Performance of CASPT2 and DFT for Relative Spin-State Energetics of Heme Models. S. Vancoillie, H. Zhao, M. Radon and K. Pierloot. *J. Chem. Theory Comput.* **6**, 576 (2010).

[119] Electronic structure and spectrum of $UO_2{}^{2+}$ and $UO_2Cl_4{}^{2-}$. K. Pierloot and E. van Besien. *J. Chem. Phys.* **123**, 204309 (2005).

[120] Water Exchange Mechanism in the First Excited State of Hydrated Uranyl(VI). P. Wahlin, V. Vallet, U. Wahlgren and I. Grenthe. *Inorg. Chem.* **48**, 11310 (2009).

[121] Integral-direct coupled cluster calculations of frequency-dependent polarizabilities, transition probabilities and excited-state properties. O. Christiansen, A. Halkier, H. Koch, P. Jorgensen and T. Helgaker. *J. Chem. Phys.* **108**, 2801–2816 (1998).

[122] Time-Dependent Density-Functional Theory. E. K. U. Gross and W. Kohn. *Adv. Quantum Chem.* **21**, 255 (1990).

[123] M. E. Casida, Time-Dependent Density Functional Response Theory for Molecules, in *Recent Advances in Density-Functional Methods*, ed.D. P. Chong, pp. 155–192 (World Scientific, Singapore, 1995).

[124] Density-functional theory of linear and nonlinear time-dependent molecular properties. P. Sałek, O. Vahtras, T. Helgaker and H. Ågren. *J. Chem. Phys.* **117**, 9630–9645 (2002).

[125] Linear response at the 4-component relativistic density-functional level: application to the frequency-dependent dipole polarizability of Hg, AuH and PtH2. P. Sałek, T. Helgaker and T. Saue. *Chem. Phys.* **311**, 187–201 (2005).

[126] State-selective optimization of local excited electronic states in extended systems. A. Kovyrshin and J. Neugebauer. *J. Chem. Phys.* **133**, 174114 (2010).

[127] Double Pertubation Theory: a powerful tool in computational coordination chemistry. J. Autschbach and T. Ziegler. *Coord. Chem. Rev.* **238/239**, 83–126 (2003).

[128] F. Furche and D. Rappoprt, Density functional methods for excited states: equilibrium structure and electronic spectra, in *Computational Photochemistry*, ed.M. Olivucci, vol. 16 of *Computational and Theoretical Chemistry* (Elsevier, Amsterdam, 2005).

[129] Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling. F. Neese. *Coord. Chem. Rev.* **253**, 526–563 (2009).

[130] Assessment of Functionals for TD-DFT Calculations of Singlet-Triplet Transitions. D. Jacquemin, E. A. Perpéte, I. Ciofini and C. Adamo. *J. Chem. Theory Comput.* **6**, 1532 (2010).

[131] TD-DFT Investigation of the UV Spectra of Pyranone Derivatives. J. Preat, D. Jacquemin, J.-M. A. V. Wathelet and E. A. Perpéte. *J. Phys. Chem. A* **110**, 8144 (2006).

[132] Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules. D. Jacquemin, V. Wathelet, E. A. Perpéte, I. Ciofini and C. Adamo. *J. Chem. Theory Comput.* **5**, 2420 (2009).

[133] The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Y. Zhao and D. Truhlar. *Theor. Chem. Acc.* **120**, 215–241 (2008).

[134] A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. M. A. Rohrdanz, K. M. Martins and J. M. Herbert. *J. Chem. Phys.* **130**, 054112 (2009).

[135] Benchmarks for electronically excited states: Time-dependent density functional theory and density functional theory based multireference configuration interaction. M. Silva-Junior, M. Schreiber, S. P. A. Sauer and W. Thiel. *J. Phys. Chem.* **129**, 104103 (2008).

[136] Benchmarks of electronically excited states: Basis set effects on CASPT2 results. M. Silva-Junior, M. Schreiber, S. Sauer and W. Thiel. *J. Phys. Chem.* **133**, 174318 (2010).

[137] Electronic spectroscopy of $UO_2^{2+}$, $NUO^+$ and NUN: an evaluation of time-dependent density functional theory for actinides. P. Tecmer, A. S. P. Gomes, U. Ekström and L. Visscher. *Phys. Chem. Chem. Phys.* **13**, 6249–6259 (2011).

[138] On the determination of excitation energies using density functional theory. D. J. Tozer and N. C. Handy. *Phys. Chem. Chem. Phys.* **2**, 2117–2121 (2000).

[139] Exchange-correlation potential with correct asymptotic behavior. R. van Leeuwen and E. J. Baerends. *Phys. Rev. A* **49**, 2421 (1994).

[140] Molecular calculations of excitation energies and (hyper)polarizabilities with a statistical average of orbital model exchange-correlation potentials. P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends. *J. Chem. Phys.* **112**, 1344–1352 (2000).

[141] Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. A. Dreuw, J. L. Weisman and M. Head-Gordon. *J. Chem. Phys.* **119**, 2943–2946 (2003).

[142] Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules. A. Dreuw and M. Head-Gordon. *Chem. Rev.* **105**, 4009 (2005).

[143] Charge-Transfer Excitations and Time-Dependent Density Functional Theory: Problems and Some Proposed Solutions. J. Autschbach. *ChemPhysChem* **10**, 1757–1760 (2009).

[144] Propagator corrections to adiabatic time-dependent density-functional theory linear response theory. M. E. Casida. *J. Chem. Phys.* **122**, 054111–054111–9 (2005).

[145] Double excitations within time-dependent density functional theory linear response. N. T. Maitra, F. Zhang, R. J. Cave and K. Burke. *J. Chem. Phys.* **120**, 5932–5937 (2004).

[146] Integral-direct coupled cluster calculations of frequency-dependent polarizabilities, transition probabilities and excited-state properties. O. Christiansen, A. Halkier, H. Koch, P. Jorgensen and T. Helgaker. *J. Chem. Phys.* **108**, 2801 (1998).

[147] The 2nd-order Approximate Coupled-cluster Singles and Doubles Model CC2. O. Christiansen, H. Koch and P. Jørgensen. *Chem. Phys. Lett.* **243**, 409–418 (1995).

[148] CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. C. Hättig and F. Weigend. *J. Chem. Phys.* **113**, 5154 (2000).

[149] Local CC2 electronic excitation energies for large molecules with density fitting. D. Kats, T. Korona and M. Schütz. *J. Chem. Phys.* **125**, 104106 (2006).

[150] T. Saue, Spin-Interactions and the Non-relativistic Limit of Electrodynamics, in *Adv. Quantum Chem.*, pp. 383–405 (Elsevier, 2005).

[151] Principles of direct 4-component relativistic SCF: application to caesium auride. T. SAUE, K. Faegri, T. Helgaker and O. Gropen. *Mol. Phys.* **91**, 937–950 (1997).

[152] Quaternion symmetry in relativistic molecular calculations: The Dirac–Hartree–Fock method. T. Saue and H. J. A. Jensen. *J. Chem. Phys.* **111**, 6211–6222 (1999).

[153] Computational strategies for a four-component Dirac–Kohn–Sham program: Implementation and first applications. L. Belpassi, F. Tarantelli, A. Sgamellotti and H. M. Quiney. *J. Chem. Phys.* **122**, 184109–13 (2005).

[154] Recent advances and perspectives in four-component Dirac–Kohn–Sham calculations. L. Belpassi, L. Storchi, H. M. Quiney and F. Tarantelli. *Phys. Chem. Chem. Phys.* **13**, 12368–27 (2011).

[155] Kinetic balance: A partial solution to the problem of variational safety in Dirac calculations. R. E. Stanton and S. Havriliak. *J. Chem. Phys.* **81**, 1910–1918 (1984).

[156] Kinetic balance in contracted basis sets for relativistic calculations. L. Visscher, P. J. C. Aerts, O. Visser and W. C. Nieuwpoort. *Int. J. Quantum Chem.* **40**, 131–139 (2004).

[157] Approximate molecular relativistic Dirac-Coulomb calculations using a simple Coulombic correction. L. Visscher. *Theor. Chem. Acc.* **98**, 68–70 (1997).

[158] Invited review: Relativistic wave-function based electron correlation methods. T. Fleig. *Chem. Phys.* **395**, 2–15 (2012).

[159] M. Dolg, *Computational Methods in Lanthanide and Actinide Chemistry*, 1st ed. (John Wiley & Sons, 2015).

[160] Formulation and implementation of a relativistic unrestricted coupled-cluster method including noniterative connected triples. L. Visscher, T. J. Lee and K. G. Dyall. *J. Chem. Phys.* **105**, 8769–8776 (1996).

[161] Formulation and implementation of the relativistic Fock-space coupled cluster method for molecules. L. Visscher, E. Eliav and U. Kaldor. *J. Chem. Phys.* **115**, 9720–9726 (2001).

[162] The generalized active space concept for the relativistic treatment of electron correlation. II. Large-scale configuration interaction implementation based on relativistic 2- and 4-spinors and its application. T. Fleig, J. Olsen and L. Visscher. *J. Chem. Phys.* **119**, 2963 (2003).

[163] The generalized active space concept for the relativistic treatment of electron correlation. III. Large-scale configuration interaction and multiconfiguration self-consistent-field four-component methods with application to $UO_2$. T. Fleig, H. J. A. Jensen, J. Olsen and L. Visscher. *J. Chem. Phys.* **124**, 104106 (2006).

[164] Four-Component Relativistic Coupled Cluster and Configuration Interaction Calculations on the Ground and Excited States of the RbYb Molecule. L. K. Sørensen, S. Knecht, T. Fleig and C. M. Marian. *J. Phys. Chem. A* **113**, 12607–12614 (2009).

[165] Accurate calculations of the ground state and low-lying excited states of the $(Rb\text{-}Ba)^+$ molecular ion, a proposed system for ultracold reactive collisions. S. Knecht, L. K. Sørensen, H. J. A. Jensen, T. Fleig and C. M. Marian. *J. Phys. B: At. Mol. Opt. Phys.* **43**, 055101 (2010).

[166] Large-scale parallel configuration interaction. II. Two- and four-component double-group general active space implementation with application to BiH. S. Knecht, H. J. A. Jensen and T. Fleig. *J. Chem. Phys.* **132**, 014108 (2010).

[167] Large-scale parallel configuration interaction. I. Nonrelativistic and scalar-relativistic general active space implementation with application to $(Rb\text{-}Ba)^+$. S. Knecht, H. J. Aa. Jensen and T. Fleig. *J. Chem. Phys.* **128**, 014108 (2008).

[168] The generalized active space concept for the relativistic treatment of electron correlation. II: Large-scale configuration interaction implementation based on relativistic 2- and 4-spinors and its application. T. Fleig, J. Olsen and L. Visscher. *J. Chem. Phys.* **119**, 2963 (2003).

[169] The generalized active space concept for the relativistic treatment of electron correlation. III: Large-scale configuration interaction and multi-configuration self-consistent-field four-component methods with application to $UO_2$. T. Fleig, H. J. Aa. Jensen, J. Olsen and L. Visscher. *J. Chem. Phys.* **124**, 104106 (2006).

[170] The one-particle Green's function method in the Dirac–Hartree–Fock framework. II. Third-order valence ionization energies of the noble gases, CO and ICN. M. Pernpointner. *J. Chem. Phys.* **121**, 8782–11 (2004).

[171] The one-particle Green's function method in the Dirac–Hartree–Fock framework. I. Second-order valence ionization energies of Ne through Xe. M. Pernpointner and A. B. Trofimov. *J. Chem. Phys.* **120**, 4098–10 (2004).

[172] The relativistic polarization propagator for the calculation of electronic excitations in heavy systems. M. Pernpointner. *J. Chem. Phys.* **140**, 084108–11 (2014).

[173] A. Shee, *Relativistic Coupled Cluster Theory – in Molecular Properties and in Electronic Structure*, Phd thesis, Universite Paul Sabatier, Toulouse (2016).

[174] A direct relativistic four-component multiconfiguration self-consistent-field method for molecules. J. Thyssen, T. Fleig and H. J. A. Jensen. *J. Chem. Phys.* **129**, 034109–15 (2008).

[175] First-order MP2 molecular properties in a relativistic framework. J. N. P. van Stralen, L. Visscher, C. V. Larsen and H. J. A. Jensen. *Chem. Phys.* **311**, 81–95 (2005).

[176] Linear response at the 4-component relativistic level: Application to the frequency-dependent dipole polarizabilities of the coinage metal dimers. T. Saue and H. J. A. Jensen. *J. Chem. Phys.* **118**, 522 (2003).

[177] T. Saue, Chapter 7. Post Dirac-Hartree-Fock methods - properties, in *Relativistic Electronic Structure Theory*, ed.P. Schwerdtfeger, vol. 11 of *Theoretical and Computational Chemistry*, pp. 332–400 (Elsevier, 2002).

[178] Relativistic adiabatic time-dependent density functional theory using hybrid functionals and noncollinear spin magnetization. R. Bast, H. J. A. Jensen and T. Saue. *Int. J. Quantum Chem.* **109**, 2091–2112 (2009).

[179] Quadratic response functions in the time-dependent four-component Hartree-Fock approximation. P. Norman and H. J. A. Jensen. *J. Chem. Phys.* **121**, 6145–11 (2004).

[180] On the evaluation of quadratic response functions at the four-component Hartree-Fock level: Nonlinear polarization and two-photon absorption in bromo- and iodobenzene. J. Henriksson, U. Ekström and P. Norman. *J. Chem. Phys.* **124**, 214311–9 (2006).

[181] Quadratic response functions in the relativistic four-component Kohn-Sham approximation. J. Henriksson, T. Saue and P. Norman. *J. Chem. Phys.* **128**, 024105–10 (2008).

[182] The Dirac equation in quantum chemistry: Strategies to overcome the current computational problems. L. Visscher. *J. Comput. Chem.* **23**, 759–766 (2002).

[183] E. Engel, Chapter 10. Relativistic density functional theory: foundation and basic formalism., in *Relativistic Electronic Structure Theory*, ed.P. Schwerdtfeger, vol. 11 of *Theoretical and Computational Chemistry*, pp. 523–621 (Elsevier, 2002).

[184] Four-component relativistic Kohn–Sham theory. T. Saue and T. Helgaker. *J. Comput. Chem.* **23**, 814–823 (2002).

[185] eds.M. Kaupp, M. Bühl and V. G. Malkin, *Calculation of NMR and EPR Parameters. Theory and Applications* (Wiley-VCH, Weinheim, 2004).

[186] On the origin and contribution of the diamagnetic term in four-component relativistic calculations of magnetic properties. G. A. Aucar, T. Saue, L. Visscher and H. J. A. Jensen. *J. Chem. Phys.* **110**, 6208–6218 (1999).

[187] Relativistic corrections to magnetic properties. W. Kutzelnigg. *J. Comput. Chem.* **20**, 1199–1219 (1999).

[188] Diamagnetism in relativistic theory. W. Kutzelnigg. *Phys. Rev. A* **67**, 032109–032120 (2003).

[189] L. Visscher, Magnetic Balance and Explicit Diamagnetic Expressions for Nuclear Magnetic Resonance Shielding Tensors, in *Advances in Quantum Chemistry*, pp. 369–381 (Elsevier, 2005).

[190] M. Repisky, S. Komorovsky, R. Bast and K. Ruud, Chapter 8 Relativistic Calculations of Nuclear Magnetic Resonance Parameters, in *Gas Phase NMR*, pp. 267–303 (The Royal Society of Chemistry, 2016).

[191] 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. S. Komorovský, M. Repiský, O. L. Malkina, V. G. Malkin, I. M. Ondík and M. Kaupp. *J. Chem. Phys.* **128**, 104101–104115 (2008).

[192] Gauge origin independent calculations of nuclear magnetic shieldings in relativistic four-component theory. M. Iliaš, T. Saue, T. Enevoldsen and H. J. A. Jensen. *J. Chem. Phys.* **131**, 124119 (2009).

[193] Gauge origin independent calculations of molecular magnetisabilities in relativistic four-component theory. M. Iliaš, H. J. A. Jensen, R. Bast and T. Saue. *Mol. Phys.* **111**, 1373–1381 (2013).

[194] A simple scheme for magnetic balance in four-component relativistic Kohn–Sham calculations of nuclear magnetic resonance shielding constants in a Gaussian basis. M. Olejniczak, R. Bast, T. Saue and M. Pecul. *J. Chem. Phys.* **136**, 014108 (2012).

[195] T. Saue and L. Visscher, Relativistic all-electron approaches to the study of f element chemistry, in *Computational methods in lanthanide and actinide chemistry*, ed.M. Dolg (Wiley, 2014).

[196] Relativistic regular two-component Hamiltonians. E. v. Lenthe, E. J. Baerends and J. G. Snijders. *J. Chem. Phys.* **99**, 4597–4610 (1993).

[197] Relativistic total energy using regular approximations. E. van Lenthe, E. J. Baerends and J. G. Snijders. *J. Chem. Phys.* **101**, 9783–9792 (1994).

[198] An exact separation of the spin-free and spin-dependent terms of the Dirac–Coulomb–Breit Hamiltonian. K. G. Dyall. *J. Chem. Phys.* **100**, 2118–2127 (1994).

[199] Approximate relativistic electronic structure methods based on the quaternion modified Dirac equation. L. Visscher and T. Saue. *J. Chem. Phys.* **113**, 3996–4002 (2000).

[200] An infinite-order two-component relativistic hamiltonian by a simple one-step transformation. M. Iliaš and H. J. A. Jensen. *private communication* (2009).

[201] An infinite-order two-component relativistic hamiltonian by a simple one-step transformation. M. Iliaš and T. Saue. *J. Chem. Phys.* **126**, 064102 (2007).

[202] The molecular mean-field approach for correlated relativistic calculations. J. Sikkema, L. Visscher, T. Saue and M. Iliaš. *J. Chem. Phys.* **131**, 124116 (2009).

[203] The Douglas–Kroll–Hess Approach. T. Nakajima and K. Hirao. *Chem. Rev.* **112**, 385–402 (2012).

[204] B. Schimmelpfennig, AMFI, an Atomic Mean-Field Integral program (1996).

[205] Two-Component Relativistic Calculations of Electric-Field Gradients Using Exact Decoupling Methods: Spin–orbit and Picture-Change Effects. J. Autschbach, D. Peng and M. Reiher. *J. Chem. Theory Comput.* **8**, 4239–4248 (2012).

[206] Analytic energy derivatives in relativistic quantum chemistry. L. Cheng, S. Stopkowicz and J. Gauss. *International Journal of Quantum Chemistry* **114**, 1108–1127 (2014).

[207] An infinite-order two-component relativistic Hamiltonian by a simple one-step transformation. M. Iliaš and T. Saue. *J. Chem. Phys.* **126**, 064102 (2007).

[208] A mean-field spin-orbit method applicable to correlated wavefunctions. B. A. Hess, C. M. Marian, U. Wahlgren, C. Teichteil, H. Fagerli and O. Gropen. *Chem. Phys. Lett.* **251**, 365 (1996).

[209] On the combination of ECP-based CI calculations with all-electron spin-orbit mean-field integrals. B. Schimmelpfennig, L. Maron, U. Wahlgren, C. Teichteil, H. Fagerli and O. Gropen. *Chem. Phys. Lett.* **286**, 267 (1998).

[210] The restricted active space (RAS) state interaction approach with spin-orbit coupling. P.-A. Malmqvist, B. O. Roos and B. Schimmelpfennig. *Chem. Phys. Lett.* **357**, 230 (2002).

[211] MOLCAS: a program package for computational chemistry. G. Karlström, R. Lindh, P.-A. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady and L. Seijo. *Comput. Mater. Sci.* **28**, 222 (2003).

[212] A two-step uncontracted determinantal effective Hamiltonian-based SO-CI method. V. Vallet, L. Maron, C. Teichteil and J.-P. Flament. *J. Chem. Phys.* **113**, 1391–1402 (2000).

[213] On the performance of the Intermediate Hamiltonian Fock-space coupled-cluster method on linear triatomic molecules: The electronic spectra of $NpO_2^+$, $NpO_2^{2+}$ and $PuO_2^{2+}$. I. Infante, A. S. P. Gomes and L. Visscher. *J. Chem. Phys.* **125**, 074301 (2006).

[214] Structure of Dicaesium Tetrachlorodioxouranium. D. J. Watkin, R. G. Denning and K. Prout. *Acta Cryst.* **C47**, 2517–2519 (1991).

[215] The electronic structure of actinyl ions V. $f-f$ transitions in $[NpO_2Cl_4{}^{-2}]$ and $[NpO_2(NO_3)_3{}^-]$. R. G. Denning, J. O. W. Norris and D. Brown. *Mol. Phys.* **46**, 287–323 (1982).

[216] Electronic Spectra and Excited States of Neptunyl and Its $[NpO_2Cl_4]^{2-}$ Complex. J. Su, W. H. E. Schwarz and J. Li. *Inorg. Chem.* **51**, 3231–3238 (2012).

[217] Actinyl ions in $Cs_2UO_2Cl_4$. S. Matsika and R. M. Pitzer. *J. Phys. Chem. A* **105**, 637–645 (2001).

[218] Interpretation of the solution absorption spectra of the $(PuO_2)^{2+}$ and $(NpO_2)^+$ ions. Eisenstein, J C and Pryce, M H L. *journal of research of the national bureau of standards - A. physical chemistry* **70A**, 165–173 (1966).

[219] Investigation of the low-lying excited states of $PuO_2{}^{2+}$. L. Maron, B. Schimmelpfennig, J.-L. Heully, V. Vallet, U. Wahlgren and O. Gropen. *Chemical Physics* **244**, 1–7 (1999).

[220] Can density functional methods be used for open-shell actinide molecules? Comparison with multiconfigurational spin-orbit studies. C. Clavaguéra-Sarrio, V. Vallet, D. Maynau and C. J. Marsden. *J. Chem. Phys.* **121**, 5312–11 (2004).

[221] Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the lanthanides La–Lu. A. S. P. Gomes, K. G. Dyall and L. Visscher. *Theor. Chem. Acc.* **127**, 369–381 (2010).

[222] A new flowing afterglow-guided ion beam tandem mass spectrometer. Applications to the thermochemistry of polyiodide ions. K. Do, T. P. Klein, C. A. Pommerening and L. S. Sunderlin. *J. Am. Soc. Mass Spectrom.* **8**, 688 (1997).

[223] Excited states and photodissociation dynamics of the triiodine radical ($I_3$). H. Choi, T. R. Taylor, R. R. Bise, A. A. Hoops and D. M. Neumark. *J. Chem. Phys.* **113**, 8608–8614 (2000).

[224] Determination of Stability Constants between Complexing Agents and At(I) and At(III) Species Present at Ultra- Trace Concentrations. J. Champion, C. Alliot, S. Huclier, D. Deniaud, W. Asfari and G. Montavon. *Inorg. Chim. Acta* **362**, 2654–2661 (2009).

[225] Astatine Standard Redox Potentials and Speciation in Acidic Medium. J. Champion, C. Alliot, E. Renault, B. M. Mokili, M. Chérel, N. Galland and G. Montavon. *J. Phys. Chem. A* **114**, 576–582 (2010).

[226] Investigation of Astatine(III) Hydrolyzed Species: Experiments and Relativistic Calculations. J. Champion, A. Sabatié-Gogova, F. Bassal, T. Ayed, C. Alliot, N. Galland and G. Montavon. *J. Phys. Chem. A* **117**, 1983–1990 (2013).

[227] Rationalization of the Solvation Effects on the $AtO^+$ Ground-State Change. T. Ayed, F. Réal, G. Montavon and N. Galland. *J. Phys. Chem. B* **10589–10595**, 117 (2013).

[228] On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. K. A. Peterson, B. C. Shepler, D. Figgen and H. Stoll. *J. Phys. Chem. A* **110**, 13877–13883 (2006).

[229] Photoelectron spectroscopy of $IO^-$. M. K. Gilles, M. L. Polak and W. C. Lineberger. *J. Chem. Phys.* **95**, 4723–4724 (1991).

[230] Rotational spectroscopy of IO $X^2\Pi_i$. C. E. Miller and E. A. Cohen. *J. Chem. Phys.* **115**, 6459–6470 (2001).

[231] Photoelectron spectroscopy of the halogen oxide anions $FO^-$, $ClO^-$, $BrO^-$, $IO^-$, $OClO^-$, and $OIO^-$. M. K. Gilles, M. L. Polak and W. C. Lineberger. *J. Chem. Phys.* **96**, 8012–8020 (1992).

[232] Discharge Flow-Photoionization Mass Spectrometric Study of HOI: Photoionization Efficiency Spectrum and Ionization Energy. P. S. Monks, L. J. Stief, D. C. Tardy, J. F. Liebman, Z. Zhang, S.-C. Kuo and R. B. Klemm. *J. Chem. Phys.* **99**, 16566–16570 (1995).

[233] Experimental Determination of the Ionization Energy of $IO(X^2\Pi_{3/2})$ and Estimations of $\Delta_f H°(IO^+)$ and PA(IO). Z. Zhang, P. S. Monks, L. J. Stief, J. F. Liebman, R. E. Huie, S.-C. Kuo and R. B. Klemm. *J. Phys. Chem.* **100**, 63–68 (1996).

[234] Spin-orbit couplings within the equation-of-motion coupled-cluster framework: Theory, implementation, and benchmark calculations. E. Epifanovsky, K. Klein, S. Stopkowicz, J. Gauss and A. I. Krylov. *J. Chem. Phys.* **143**, 064102–17 (2015).

[235] Improved measurement of the shape of the electron. J. J. Hudson, D. M. Kara, I. J. Smallman, B. E. Sauer, M. R. Tarbutt and E. A. Hinds. *Nature* **473**, 493–496 (2011).

[236] High-resolution spectroscopy on trapped molecular ions in rotating electric fields: A new approach for measuring the electron electric dipole moment. A. E. Leanhardt, J. L. Bohn, H. Loh, P. Maletinsky, E. R. Meyer, L. C. Sinclair, R. P. Stutz and E. A. Cornell. *Journal of Molecular Spectroscopy* **270**, 1–25 (2011).

[237] Mixed-sector intermediate Hamiltonian Fock-space coupled cluster approach. A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor. *J. Chem. Phys.* **121**, 6634 (2004).

[238] Spectroscopic investigations of ThF and $ThF^+$. B. J. Barker, I. O. Antonov, M. C. Heaven and K. A. Peterson. *J. Chem. Phys.* **136**, 104305–10 (2012).

[239] Spectroscopy and Structure of the Simplest Actinide Bonds. M. C. Heaven, B. J. Barker and I. O. Antonov. *J. Phys. Chem. A* **118**, 10867–10881 (2014).

[240] Broadband velocity modulation spectroscopy of $ThF^+$ for use in a measurement of the electron electric dipole moment. D. N. Gresh, K. C. Cossel, Y. Zhou, J. Ye and E. A. Cornell. *Journal of Molecular Spectroscopy* **319**, 1–9 (2016).

[241] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, 2014).

[242] Theoretical investigation of the energies and geometries of photo-excited uranyl(VI) ion: a comparison between wave-function theory and density functional theory. F. Réal, V. Vallet, C. Marian and U. Wahlgren. *J. Chem. Phys.* **127**, 214302 (2007).

[243] Electronic structure and spectrum of $UO_2{}^{2+}$ and $UO_2Cl_4{}^{2-}$. K. Pierloot and E. van Besien. *J. Chem. Phys.* **123**, 204309 (2005).

[244] Transformation of the Hamiltonian in excitation energy calculations: Comparison between Fock-space multireference coupled-cluster and equation-of-motion coupled-cluster methods. L. Meissner and R. J. Bartlett. *J. Chem. Phys.* **94**, 6670–6676 (1991).

[245] A dressing for the matrix elements of the singles and doubles equation-of-motion coupled-cluster method that recovers additive separability of excitation energies. L. Meissner and R. J. Bartlett. *J. Chem. Phys.* **102**, 7490–7498 (1995).

[246] Intermediate Hamiltonian Fock-space multireference coupled-cluster method with full triples for calculation of excitation energies. M. Musial and R. J. Bartlett. *J. Chem. Phys.* **129**, 044101 (2008).

[247] Multireference Fock-space coupled-cluster and equation-of-motion coupled-cluster theories: The detailed interconnections. M. Musial and R. J. Bartlett. *J. Chem. Phys.* **129**, 134105 (2008).

[248] Benchmark calculations of the Fock-space coupled cluster single, double, triple excitation method in the intermediate Hamiltonian formulation for electronic excitation energies. M. Musial and R. J. Bartlett. *Chem. Phys. Lett.* **457**, 267–270 (2008).

[249] A theoretical study of the ground state and lowest excited states of $PuO^{0/+1/+2}$ and $PuO_2^{0/+1/+2}$. G. La Macchia, I. Infante, R. Juraj, J. K. Gibson and L. Gagliardi. *Phys. Chem. Chem. Phys.* **10**, 7278–7283 (2008).

[250] Ionization Energies for the Actinide Mono- and Dioxides Series, from Th to Cm: Theory versus Experiment. I. Infante, A. Kovacs, G. L. Macchia, A. R. M. Shahi, J. K. Gibson and L. Gagliardi. *J. Phys. Chem. A* **114**, 6997–6015 (2010).

[251] The Kohn–Sham gap, the fundamental gap and the optical gap: the physical meaning of occupied and virtual Kohn–Sham orbital energies. E. J. Baerends, O. V. Gritsenko and R. van Meer. *Phys. Chem. Chem. Phys.* **15**, 16408–18 (2013).

[252] Proof that $\frac{\partial E}{\partial n_i} = E$ in density-functional theory. J. F. Janak. *Phys. Rev. B* **18**, 7165 (1978).

[253] Interpretation of the Kohn–Sham orbital energies as approximate vertical ionization potentials. D. P. Chong, O. V. Gritsenko and E. J. Baerends. *J. Chem. Phys.* **117**, 1760 (2002).

[254] Molecular calculations of excitation energies and (hyper)polarizabilities with a statistical average of orbital model exchange-correlation potentials. P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends. *J. Chem. Phys.* **112**, 1344 (2000).

[255] A bird's-eye view of density-functional theory. K. Capelle. *J. Braz. Phys. Soc.* **36**, 1318 (2006).

[256] Physical Meaning of Virtual Kohn–Sham Orbitals and Orbital Energies: An Ideal Basis for the Description of Molecular Excitations. R. van Meer, O. V. Gritsenko and E. J. Baerends. *J. Chem. Theory Comput.* **10**, 4432–4441 (2014).

[257] Noble gas-actinide compounds: Complexation of the CUO molecule by Ar, Kr, and Xe atoms in noble gas matrices. J. Li, B. E. Bursten, B. Liang and L. Andrews. *Science* **295**, 2242 (2002).

[258] The ground state and electronic spectrum of CUO: a mystery. B. O. Roos, P.-O.Widmark and L. Gagliardi. *Faraday Discuss.* **124**, 57 (2003).

[259] The importance of spin-orbit coupling and electron correlation in the rationalization of the ground state of the CUO molecule. I. Infante and L. Visscher. *J. Chem. Phys.* **121**, 5783 (2004).

[260] The quantum chemistry of d- and f- elements complexes: from an approximate existence to functional hapiness. B. E. Bursten, M. L. Drummond and J. Li. *Faraday Discuss.* **124**, 1 (2003).

[261] Theoretical Chemistry, Vrije Universiteit Amsterdam, ADF, Amsterdam density functional program. URL: http://www.scm.com.

[262] Theoretical Study of the Luminescent States and Electronic Spectra of $UO_2Cl_2$ in an Argon Matrix. J. Su, Y.-L. Wang, F. Wei, W. H. E. Schwarz and J. Li. *J. Chem. Theory Comput.* **7**, 3293–3303 (2011).

[263] Frozen Density Functional Approach for ab Initio Calculations of Solvated Molecules. T. A. Wesolowski and A. Warshel. *J. Phys. Chem.* **97**, 8050–8053 (1993).

[264] Density dependence of the dielectric constant of rare-gas crystals. G. Senatore and K. R. Subbaswamy. *Phys. Rev. B* **34**, 5754 (1986).

[265] Hyperpolarizabilities of alkali halide crystals using the local-density approximation. M. D. Johnson, K. R. Subbaswamy and G. Senatore. *Phys. Rev. B* **36**, 9202 (1987).

[266] Self-consistently determined properties of solids without band-structure calculations. P. Cortona. *Phys. Rev. B* **44**, 8454–8458 (1991).

[267] Partition density-functional theory. P. Elliott, K. Burke, M. H. Cohen and A. Wasserman. *Phys. Rev. A* **82**, 024501 (2010).

[268] Density functional embedding for molecular systems. M. Iannuzzi, B. Kirchner and J. Hutter. *Chem. Phys. Lett.* **421**, 16–20 (2006).

[269] A flexible implementation of frozen-density embedding for use in multilevel simulations. Ch. R. Jacob, J. Neugebauer and L. Visscher. *J. Comput. Chem.* **29**, 1011–1018 (2008).

[270] Fragmentation Methods: A Route to Accurate Calculations on Large Systems. M. S. Gordon, D. G. Fedorov, S. R. Pruitt and L. V. Slipchenko. *Chem. Rev.* **112**, 632–672 (2012).

[271] Electronic-structure calculations by first-principles density-based embedding of explicitly correlated systems. N. Govind, Y. A. Wang and E. A. Carter. *J. Chem. Phys.* **110**, 7677–7688 (1999).

[272] Embedding a multideterminantal wave function in an orbital-free environment. T. A. Wesolowski. *Phys. Rev. A* **77**, 012504 (2008).

[273] Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory. A. S. P. Gomes, Ch. R. Jacob and L. Visscher. *Phys. Chem. Chem. Phys.* **10**, 5353–5362 (2008).

[274] Self-consistency in frozen-density embedding theory based calculations. F. Aquilante and T. A. Wesolowski. *J. Chem. Phys.* **135**, 084120 (2011).

[275] T. A. Wesolowski, One-electron equations for embedded electron density: challenge for theory and practical payoffs in multilevel modelling of complex polyatomic systems, in *Computational Chemistry: Reviews of Current Trends*, ed.J. Leszczynski, vol. 10, pp. 1–82 (World Scientific, Singapore, 2006).

[276] Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds. S. Fux, Ch. R. Jacob, J. Neugebauer, L. Visscher and M. Reiher. *J. Chem. Phys.* **132**, 164101 (2010).

[277] Kohn-Sham equations with constrained electron density: an iterative evaluation of the ground-state electron density of interaction molecules. T. A. Wesolowski and J. Weber. *Chem. Phys. Lett.* **248**, 71–76 (1996).

[278] Quantum mechanical embedding theory based on a unique embedding potential. C. Huang, M. Pavone and E. A. Carter. *J. Chem. Phys.* **134**, 154110 (2011).

[279] Embedding theory for excited states. Y. G. Khait and M. R. Hoffmann. *J. Chem. Phys.* **133**, 044107 (2010).

[280] Generalization of the Kohn-Sham Equations with Constrained Electron Density Formalism and Its Time-Dependent Response Theory Formulation. M. E. Casida and T. A. Wesolowski. *Int. J. Quantum Chem.* **96**, 577–588 (2004).

[281] Couplings between electronic transitions in a subsystem formulation of time-dependent density functional theory. J. Neugebauer. *J. Chem. Phys.* **126**, 134116 (2007).

[282] On the calculation of general response properties in subsystem density functional theory. J. Neugebauer. *J. Chem. Phys.* **131**, 084104 (2009).

[283] On the subsystem formulation of linear-response time-dependent DFT. M. Pavanello. *J. Chem. Phys.* **138**, 204118 (2013).

[284] Calculation of electronic excitations using wave-function in wave-function frozen-density embedding. S. Höfener and L. Visscher. *J. Chem. Phys.* **137**, 204120 (2012).

[285] Wave Function Frozen-Density Embedding: Coupled Excitations. S. Höfener and L. Visscher. *J. Chem. Theory Comput.* **12**, 549–557 (2016).

[286] Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. D. Bakowies and W. Thiel. *J. Phys. Chem.* **100**, 10580–10594 (1996).

[287] Theory for the Forces between Closed-Shell Atoms and Molecules. R. G. Gordon and Y. S. Kim. *J. Chem. Phys.* **56**, 3122–3133 (1972).

[288] Study of the electron gas approximation. Y. S. Kim and R. G. Gordon. *J. Chem. Phys.* **60**, 1842–1850 (1974).

[289] Orbital-free embedding applied to the calculation of induced dipole moments in $CO_2 \cdots X$ (X=He, Ne, Ar, Kr, Xe, Hg) van der Waals complexes. Ch. R. Jacob, T. A. Wesolowski and L. Visscher. *J. Chem. Phys.* **123**, 174104 (2005).

[290] Frozen density embedding with hybrid functionals. S. Laricchia, E. Fabiano and F. Della Sala. *J. Chem. Phys.* **133**, 164111 (2010).

[291] Kohn-Sham Equations with Constrained Electron Density: The Effect of Various Kinetic Energy Functiona Parametrizations on the Ground-State Molecular Properties. T. A. Wesolowski and J. Weber. *Int. J. Quantum Chem.* **61**, 303–311 (1997).

[292] Y. A. Wang and E. A. Carter, Orbital-free kinetic-energy density functional theory, in *Theoretical Methods in Condensed Phase Chemistry*, ed.S. D. Schwartz, pp. 117–184 (Kluwer, Dordrecht, 2000).

[293] Link between the Kinetic- and Exchange-Energy Functionals in the Generalized Gradient Approximation. F. Tran and T. A. Wesolowski. *Int. J. Quantum Chem.* **89**, 441 – 446 (2002).

[294] Orbital-free effective embedding potential at nuclear cusps. J. M. Garcia Lastra, J. W. Kaminski and T. A. Wesolowski. *J. Chem. Phys.* **129**, 074107 (2008).

[295] Performance of Kinetic Energy Functionals for Interaction Energies in a Subsystem Formulation of Density Functional Theory. A. W. Götz, S. M. Beyhan and L. Visscher. *J. Chem. Theory Comput.* **5**, 3161–3174 (2009).

[296] Generalized Gradient Approximations of the Noninteracting Kinetic Energy from the Semiclassical Atom Theory: Rationalization of the Accuracy of the Frozen Density Embedding Theory for Nonbonded Interactions. S. Laricchia, E. Fabiano, L. A. Constantin and F. Della Sala. *J. Chem. Theory Comput.* **7**, 2439–2451 (2011).

[297] Accuracy of approximate kinetic energy functionals in the model of Kohn-Sham equations with constrained electron density: The FH—NCH complex as a test case. T. A. Wesolowski, H. Chermette and J. Weber. *J. Chem. Phys.* **105**, 9182–9190 (1996).

[298] Density functional theory with approximate kinetic energy functionals applied to hydrogen bonds. T. A. Wesolowski. *J. Chem. Phys.* **106**, 8516–8526 (1997).

[299] Obtaining a gradient-corrected kinetic-energy functional from the Perdew-Wang exchange functional. A. Lembarki and H. Chermette. *Phys. Rev. A* **50**, 5328–5331 (1994).

[300] Topological analysis of electron densities from Kohn-Sham and subsystem density functional theory. K. Kiewisch, G. Eickerling, M. Reiher and J. Neugebauer. *J. Chem. Phys.* **128**, 044114 (2008).

[301] Analysis of electron density distributions from subsystem density functional theory applied to coordination bonds. S. Fux, K. Kiewisch, Ch. R. Jacob, J. Neugebauer and M. Reiher. *Chem. Phys. Lett.* **461**, 353–359 (2008).

[302] Exact functional derivative of the nonadditive kinetic-energy bifunctional in the long-distance limit. Ch. R. Jacob, S. M. Beyhan and L. Visscher. *J. Chem. Phys.* **126**, 234116 (2007).

[303] The weak covalent bond in NgAuF (Ng = Ar, Kr, Xe): A challenge for subsystem density functional theory. S. M. Beyhan, A. W. Götz, Ch. R. Jacob and L. Visscher. *J. Chem. Phys.* **132**, 044114 (2010).

[304] A subsystem density-functional theory approach for the quantum chemical treatment of proteins. Ch. R. Jacob and L. Visscher. *J. Chem. Phys.* **128**, 155102 (2008).

[305] Self-consistent calculations of total energies and charge densities of solids without solving the band-structure problem. P. Cortona and A. Villafiorita Monteleone. *Int. J. Quantum Chem.* **52**, 987–992 (1994).

[306] Direct calculations of charge densities of solids: Applications to the alkali-earth sulfides. P. Cortona, A. Villafiorita Monteleone and P. Becker. *Int. J. Quantum Chem.* **56**, 831–837 (1995).

[307] Analysis of the MgO structure factors. J.-M. Gillet and P. Cortona. *Phys. Rev. B* **60**, 8569–8574 (1999).

[308] A self consistent atomic deformation model for total energy calculations: Application to ferroelectrics. L. L. Boyer and M. J. Mehl. *Ferroelectrics* **150**, 13–24 (1993).

[309] Calculation of electronic, structural, and vibrational properties in alkali halides using a density-functional method with localized densities. W. N. Mei, L. L. Boyer, M. J. Mehl, M. M. Ossowski and H. T. Stokes. *Phys. Rev. B* **61**, 11425–11431 (2000).

[310] Lattice dynamics and elastic properties of corundum by the self-consistent atomic deformation method. M. M. Ossowski, L. L. Boyer, M. J. Mehl and H. T. Stokes. *Phys. Rev. B* **66**, 224302 (2002).

[311] Optical and vibrational properties of $MnF_6^{4-}$ complexes in cubic fluoroperovskites: insight through embedding calculations using Kohn–Sham equations with constrained electron density. J. M. García-Lastra, T. Wesolowski, M. T. Barriuso, J. A. Aramburu and M. Moreno. *J. Phys.: Condens. Matter* **18**, 1519–1534 (2006).

[312] Application of the density functional theory derived orbital-free embedding potential to calculate the splitting energies of lanthanide cations in chloroelpasolite crystals. M. Zbiri, M. Atanasov, C. Daul, J. M. Garcia-Lastra and T. A. Wesolowski. *Chem. Phys. Lett.* **397**, 441–446 (2004).

[313] Effect of the f-Orbital Delocalization on the Ligand-Field Splitting Energies in Lanthanide-Containing Elpasolites. M. Zbiri, C. A. Daul and T. A. Wesolowski. *J. Chem. Theory Comput.* **2**, 1106–1111 (2006).

[314] Hydrogen-Bonding-induced Shifts of the Excitaion Energies in Nucleic Acid Bases: An Interplay between Electrostatic and Electron Density Overlap Effects. T. A. Wesolowski. *J. Am. Chem. Soc.* **126**, 11444–11445 (2004).

[315] The merits of the frozen-density embedding scheme to model solvatochromic shifts. J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski. *J. Chem. Phys.* **122**, 094115 (2005).

[316] An Explicit Quantum Chemical Method for Modeling Large Solvation Shells Applied to Aminocoumarin C151. J. Neugebauer, Ch. R. Jacob, T. A. Wesolowski and E. J. Baerends. *J. Phys. Chem. A* **109**, 7805–7814 (2005).

[317] Photophysical Properties of Natural Light-Harvesting Complexes Studied by Subsystem Density Functional Theory. J. Neugebauer. *J. Phys. Chem. B* **112**, 2207–2217 (2008).

[318] Theoretical Spectroscopy of Astaxanthin in Crustacyanin Proteins: Absorption, Circular Dichroism, and Nuclear Magnetic Resonance. J. Neugebauer, J. Veldstra and F. Buda. *J. Phys. Chem. B* **115**, 3216–3225 (2011).

[319] Comparison of frozen-density embedding and discrete reaction field solvent models for molecular properties. Ch. R. Jacob, J. Neugebauer, L. Jensen and L. Visscher. *Phys. Chem. Chem. Phys.* **8**, 2349–2359 (2006).

[320] A Subsystem TDDFT Approach for Solvent Screening Effects on Excitation Energy Transfer Couplings. J. Neugebauer, C. Curutchet, A. Muñoz Losa and B. Mennucci. *J. Chem. Theory Comput.* **6**, 1843–1851 (2010).

[321] First-principles calculation of electronic spectra of light-harvesting complex II. C. König and J. Neugebauer. *Phys. Chem. Chem. Phys.* **13**, 10475 (2011).

[322] Fragment-Based Time-Dependent Density Functional Theory. M. A. Mosquera, D. Jensen and A. Wasserman. *Phys. Rev. Lett.* **111**, 023001–5 (2013).

[323] Current density partitioning in time-dependent current density functional theory. M. A. Mosquera and A. Wasserman. *J. Chem. Phys.* **140**, 18A525–9 (2014).

[324] Time-Dependent Electronic Populations in Fragment-Based Time-Dependent Density Functional Theory. M. A. Mosquera and A. Wasserman. *J. Chem. Theory Comput.* **11**, 3530–3536 (2015).

[325] Subsystem real-time time dependent density functional theory. A. Krishtal, D. Ceresoli and M. Pavanello. *J. Chem. Phys.* **142**, 154116–14 (2015).

[326] Periodic subsystem density-functional theory. A. Genova, D. Ceresoli and M. Pavanello. *J. Chem. Phys.* **141**, 174101–11 (2014).

[327] Exploiting the locality of periodic subsystem density-functional theory: efficient sampling of the Brillouin zone. A. Genova and M. Pavanello. *Journal of Physics: Condensed Matter* pp. 1–7 (2015).

[328] An accurate and linear-scaling method for calculating charge-transfer excitation energies and diabatic couplings. M. Pavanello, T. Van Voorhis, L. Visscher and J. Neugebauer. *J. Chem. Phys.* **138**, 054101 (2013).

[329] Quantifying Environmental Effects on the Decay of Hole Transfer Couplings in Biosystems. P. Ramos and M. Pavanello. *J. Chem. Theory Comput.* **10**, 2546–2556 (2014).

[330] Subsystem-Based Theoretical Spectroscopy of Biomolecules and Biomolecular Assemblies. J. Neugebauer. *ChemPhysChem* **10**, 3148–3173 (2009).

[331] Chromophore-specific theoretical spectroscopy: From subsystem density functional theory to mode-specific vibrational spectroscopy. J. Neugebauer. *Phys. Rep.* **489**, 1–87 (2010).

[332] Quantum Chemical Description of Absorption Properties and Excited-State Processes in Photosynthetic Systems. C. König and J. Neugebauer. *ChemPhysChem* **13**, 386–425 (2012).

[333] Accurate ab initio energetics of extended systems via explicit correlation embedded ia a density functional environment. N. Govind, Y. A. Wang, A. J. R. da Silva and E. A. Carter. *Chem. Phys. Lett.* **295**, 129–134 (1998).

[334] Prediction of Electronic Excited States of Adsorbates on Metal Surfaces from First Principles. T. Klüner, N. Govind, Y. A. Wang and E. A. Carter. *Phys. Rev. Lett.* **86**, 5954–5957 (2001).

[335] Periodic density functional embedding theory for complete active space self-consistent field and configuration interaction calculations: Ground and excited state. T. Klüner, N. Govind, Y. A. Wang and E. A. Carter. *J. Chem. Phys.* **116**, 42–54 (2002).

[336] Self-consistent embedding theory for locally correlated configuration interaction wave functions in condensed matter. P. Huang and E. A. Carter. *J. Chem. Phys.* **125**, 084102 (2006).

[337] A self-consistent density based embedding scheme applied to the adsorption of CO on Pd(111). D. Lahav and T. Klüner. *J. Phys.: Condens. Matter* **19**, 226001 (2007).

[338] Local Electronic Structure around a Single Kondo Impurity. P. Huang and E. A. Carter. *Nano Lett.* **6**, 1146–1150 (2006).

[339] Embedded Configuration Interaction Description of CO on Cu(111): Resolution of the Site Preference Conundrum. S. Sharifzadeh, P. Huang and E. Carter. *J. Phys. Chem. C* **112**, 4649–4657 (2008).

[340] Ab Initio Explanation of Tunneling Line Shapes for the Kondo Impurity State. P. Huang and E. A. Carter. *Nano Lett.* **8**, 1265–1269 (2008).

[341] All-electron embedded correlated wavefunction theory for condensed matter electronic structure. S. Sharifzadeh, P. Huang and E. A. Carter. *Chem. Phys. Lett.* **470**, 347–352 (2009).

[342] Quantum mechanical modeling of electronic excitations in metal oxides: Magnesia as a prototype. D. K. Kanan, S. Sharifzadeh and E. A. Carter. *Chem. Phys. Lett.* **519–520**, 18–24 (2012).

[343] Coupled-cluster frozen-density embedding using resolution of the identity methods. S. Höfener. *J. Comput. Chem.* **35**, 1716–1724 (2014).

[344] Wavefunction in Density Functional Theory Embedding for Excited States: Which Wavefunctions, which Densities? C. Daday, C. König, J. Neugebauer and C. Filippi. *ChemPhysChem* **15**, 3205–3217 (2014).

[345] Y. G. Khait and M. R. Hoffmann, *Chapter Three - On the Orthogonality of Orbitals in Subsystem Kohn-Sham Density Functional Theory*, vol. 8 (Elsevier, 2012).

[346] No need for external orthogonality in subsystem density-functional theory. J. P. Unsleber, J. Neugebauer and C. R. Jacob. *Phys. Chem. Chem. Phys.* pp. – (2016).

[347] Functional derivative of noninteracting kinetic energy density functional. S. Liu and P. W. Ayers. *Phys. Rev. A* **70**, 022501 (2004).

[348] From electron densities to Kohn-Sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies. Q. Zhao, R. C. Morrison and R. G. Parr. *Phys. Rev. A* **50**, 2138 (1994).

[349] A direct optimization method for calculating density functionals and exchange–correlation potentials from electron densities. Q. Wu and W. Yang. *J. Chem. Phys.* **118**, 2498–2509 (2003).

[350] Exact nonadditive kinetic potentials for embedded density functional theory. J. D. Goodpaster, N. Ananth, F. R. Manby and T. F. Miller. *J. Chem. Phys.* **133**, 084103 (2010).

[351] Embedded density functional theory for covalently bonded and strongly interacting subsystems. J. D. Goodpaster, T. A. Barnes and T. F. Miller. *J. Chem. Phys.* **134**, 164108 (2011).

[352] Excitation energies from frozen-density embedding with accurate embedding potentials. D. G. Artiukhin, C. R. Jacob and J. Neugebauer. *J. Chem. Phys.* **142**, 234101 (2015).

[353] An inversion technique for the calculation of embedding potentials. O. Roncero, M. P. de Lara-Castells, P. Villarreal, F. Flores, J. Ortega, M. Paniagua and A. Aguado. *J. Chem. Phys.* **129**, 184104 (2008).

[354] Comment on "Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds" [J. Chem. Phys. 132, 164101 (2010)]. T. A. Wesolowski. *J. Chem. Phys.* **135**, 027101 (2011).

[355] Response to "Comment on 'Accurate frozen-density embedding potentials as a first step towards a subsystem description of covalent bonds'" [J. Chem. Phys. 135, 027101 (2011)]. S. Fux, Ch. R. Jacob, J. Neugebauer, L. Visscher and M. Reiher. *J. Chem. Phys.* **135**, 027102 (2011).

[356] A density-division embedding potential inversion technique. O. Roncero, A. Zanchet, P. Villarreal and A. Aguado. *J. Chem. Phys.* **131**, 234110 (2009).

[357] Molecular binding energies from partition density functional theory. J. Nafziger, Q. Wu and A. Wasserman. *J. Chem. Phys.* **135**, 234101 (2011).

[358] Potential-functional embedding theory for molecules and materials. C. Huang and E. A. Carter. *J. Chem. Phys.* **135**, 194104 (2011).

[359] Time-dependent potential-functional embedding theory. C. Huang, F. Libisch, Q. Peng and E. A. Carter. *J. Chem. Phys.* **140**, 124113–13 (2014).

[360] Can optimized effective potentials be determined uniquely? S. Hirata, S. Ivanov, I. Grabowski, R. J. Bartlett, K. Burke and J. D. Talman. *J. Chem. Phys.* **115**, 1635 (2001).

[361] Optimized effective potentials yielding Hartree–Fock energies and densities. V. N. Staroverov, G. E. Scuseria and E. R. Davidson. *J. Chem. Phys.* **124**, 141103 (2006).

[362] Numerically stable optimized effective potential method with balanced Gaussian basis sets. A. Heßelmann, A. W. Götz, F. Della Sala and A. Görling. *J. Chem. Phys.* **127**, 054102 (2007).

[363] Optimized Effective Potentials in Finite Basis Sets. T. Heaton-Burgess, F. A. Bulat and W. Yang. *Phys. Rev. Lett.* **98**, 256401 (2007).

[364] Optimized effective potentials from arbitrary basis sets. T. Heaton-Burgess and W. Yang. *J. Chem. Phys.* **129**, 194102 (2008).

[365] Unambiguous optimization of effective potentials in finite basis sets. Ch. R. Jacob. *J. Chem. Phys.* **135**, 244102 (2011).

[366] The electronic structure of the uranyl ion. Part I. The electronic spectrum of $Cs_2UO_2Cl_4$. R. G. Denning, T. R. Snellgrove and D. R. Woodwark. *Mol. Phys.* **32**, 419–442 (1976).

[367] Applications of two-photon spectroscopy to inorganic compounds. 1. Spectrum and electronic structure of dicesium tetrachlorodioxouranate. T. J. Barker, R. G. Denning and J. R. G. Thorne. *Inorg. Chem.* **26**, 1721 (1987).

[368] The electronic structure of actinyl ions: the excited-state absorption spectrum of $Cs_2UO_2Cl_4$. R. G. Denning and I. D. Morrison. *Chem. Phys. Lett.* **180**, 101 (1991).

[369] Polarization Dependent High Energy Resolution X-ray Absorption Study of Dicesium Uranyl Tetrachloride. T. Vitova, J. C. Green, R. G. Denning, M. Löble, K. Kvashnina, J. J. Kas, K. Jorissen, J. J. Rehr, T. Malcherek and M. A. Denecke. *Inorg. Chem.* **54**, 174–182 (2015).

[370] Actinyl Ions in $Cs_2UO_2Cl_4$. S. Matsika and R. M. Pitzer. *J. Phys. Chem. A* **105**, 637–645 (2001).

[371] Electronic structure and bonding in actinyl ions. R. G. Denning. *Struct. Bond.* **79**, 215–276 (1992).

[372] Photoelectron spectroscopy and the electronic structure of the uranyl tetrachloride dianion: $UO_2Cl_4^{2-}$. P. D. Dau, J. Su, H.-T. Liu, D.-L. Huang, J. Li and L.-S. Wang. *J. Chem. Phys.* **137**, 064315–9 (2012).

[373] On the electron leak problem in orbital-free embedding calculations. M. Dułak and T. A. Wesołowski. *J. Chem. Phys.* **124**, 164101 (2006).

[374] Analysis of electron density distributions from subsystem density functional theory applied to coordination bonds. S. Fux, K. Kiewisch, C. R. Jacob, J. Neugebauer and M. Reiher. *Chem. Phys. Lett.* **461**, 353 (2008).

[375] The energy-differences based exact criterion for testing approximations to the functional for the kinetic energy of non-interacting electrons. Y. A. Bernard, M. Dułak, J. W. Kamiński and T. A. Wesołowski. *J. Phys A-Mat. Theor.* **41**, 055302 (2008).

[376] Unravelling the quantum-entanglement effect of noble gas coordination on the spin ground state of CUO. P. Tecmer, K. Boguslawski, Ö. Legeza and M. Reiher. *Phys. Chem. Chem. Phys.* **16**, 719–727 (2014).

[377] Calculation of nuclear magnetic resonance shieldings using frozen-density embedding. C. R. Jacob and L. Visscher. *J. Chem. Phys.* **125**, 194104 (2006).

[378] NMR Solvent Shifts of Acetonitrile from Frozen Density Embedding Calculations. R. E. Bulo, C. R. Jacob and L. Visscher. *J. Phys. Chem. A* **112**, 2640–2647 (2008).

[379] Calculation of nuclear spin-spin coupling constants using frozen density embedding. A. W. Götz, J. Autschbach and L. Visscher. *J. Chem. Phys.* **140**, 104107 (2014).

[380] Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spin–Spin Coupling Constants. T. Helgaker, M. Jaszuński and K. Ruud. *Chem. Rev.* **99**, 293–352 (1999).

[381] Conventions for the reporting of nuclear magnetic shielding (or shift) tensors suggested by participants in the NATO ARW on NMR shielding constants at the University of Maryland, College Park, July 1992. J. Mason. *Solid State Nuclear Magnetic Resonance* **2**, 285 – 288 (1993).

[382] 4-Component relativistic calculation of the magnetically induced current density in the group 15 heteroaromatic compounds. R. Bast, J. Jusélius and T. Saue. *Chem. Phys.* **356**, 187–194 (2009).

[383] Molecular electronic property density functions: The nuclear magnetic shielding density. C. J. Jameson and A. D. Buckingham. *J. Chem. Phys.* **73**, 5684–5692 (1980).

[384] 4-Component relativistic magnetically induced current density using London atomic orbitals. D. Sulzer, M. Olejniczak, R. Bast and T. Saue. *Phys. Chem. Chem. Phys.* **13**, 20682–20689 (2011).

[385] J. Autschbach, Chapter 4 - Relativistic Effects on NMR Parameters, in *High Resolution NMR Spectroscopy. Understanding Molecules and their Electronic Structures*, ed.R. H. Contreras, vol. 3 of *Science and Technology of Atomic, Molecular, Condensed Matter & Biological Systems*, pp. 69–117 (Elsevier, 2013).

[386] Magnetizability of Hydrocarbons. K. Ruud, H. Skaane, T. Helgaker, K. L. Bak and P. Jorgensen. *J. Am. Chem. Soc.* **116**, 10135–10140 (1994).

[387] Atomic magnetizability tensors of benzene and fluoro- and chlorobenzenes. P.-O. Åstrand and K. V. Mikkelsen. *Magnetic Resonance in Chemistry* **36**, 92–97 (1998).

[388] Properties of atoms in molecules: Magnetic susceptibilities. R. F. W. Bader and T. A. Keith. *J. Chem. Phys.* **99**, 3683–3693 (1993).

[389] Breakdown of the pseudopotential approximation for magnetizabilities and electric multipole moments: Test calculations for Au, AuF, and Snn cluster (n $\geq$ 20). P. Schwerdtfeger, B. Assadollahzadeh, U. Rohrmann, R. Schafer and J. R. Cheeseman. *J. Chem. Phys.* **134**, 204102 (2011).

[390] Recherches magnéto-chimiques. P. Pascal. *Ann. Chim. Phys.* **19**, 5 (1910).

[391] P. Pascal. *Rev. Sci. Instrum.* **86**, 38 (1948).

[392] Magnetic interactions in molecules and an analysis of molecular electronic charge distribution from magnetic parameters. W. H. Flygare. *Chem. Rev.* **74**, 653–687 (1974).

[393] Molecular Magnetizabilities Zero-Point Vibrational Effects and the Breakdown of Pascals Rule. K. Ruud, P.-O. Astrand and P. R. Taylor. *J. Phys. Chem. A* **105**, 9926–9930 (2001).

[394] Theoretical studies of the benzene molecule: Magnetic susceptibility and nuclear shielding constants. P. Lazzeretti and R. Zanasi. *J. Chem. Phys.* **75**, 5019–5027 (1981).

[395] Inconsistency of the ring-current model for the cyclopropenyl cation. P. Lazzeretti and R. Zanasi. *Chem. Phys. Lett.* **80**, 533–536 (1981).

[396] Fock-space coupled-cluster method: The (1,2) sector. S. R. Hughes and U. Kaldor. *Phys. Rev. A* **47**, 4705 (1993).

[397] Equation-of-motion coupled cluster method for high spin double electron attachment calculations. M. Musiał, Ł. Lupa and S. A. Kucharski. *J. Chem. Phys.* **140**, 114107–8 (2014).

[398] The Coupled-Cluster Method in High Sectors of the Fock Space. S. R. Hughes and U. Kaldor. *International Journal of Quantum Chemistry* **55**, 127–132 (1995).

[399] Perturbative triples correction for the equation-of-motion coupled-cluster wave functions with single and double substitutions for ionized states: Theory, implementation, and examples. P. U. Manohar, J. F. Stanton and A. I. Krylov. *J. Chem. Phys.* **131**, 114112–14 (2009).

[400] Communication: Four-component density matrix renormalization group. S. Knecht, Ö. Legeza and M. Reiher. *J. Chem. Phys.* **140**, 041101–5 (2014).

[401] Further insights in the ability of classical non-additive potentials to model actinide ions in aqueous solution. F. Réal, M. Trumm, B. Schimmelpfennig, M. Masella and V. Vallet. *J. Comput. Chem.* **34**, 707–719 (2013).

[402] Quantum mechanical embedding theory based on a unique embedding potential. C. Huang, M. Pavone and E. A. Carter. *J. Chem. Phys.* **134**, 154110 (2011).

[403] Evaluation of $^{95}$Mo Nuclear Shielding and Chemical Shift of $[Mo_6X_{14}]^{2-}$ Clusters in the Liquid Phase. T. T. Nguyen, J. Jung, X. Trivelli, J. Trébosc, S. Cordier, Y. Molard, L. L. Pollès, C. J. Pickard, J. Cuny and R. Gautier. *Inorg. Chem.* **54**, 7673–7683 (2015). PMID: 26208250.

# Appendix A

# Curriculum Vitae


| | |
|---|---|
| **Last, First name** | Severo Pereira Gomes, André |
| **Date and place of birth** | 14-08-1976, Brasília, DF, Brazil |
| **Professional address** | Laboratoire de Physique des Lasers, Atomes et Molécules (PhLAM)<br>CNRS UMR 8523 Université de Lille 1, Bât. P5, bureau 061<br>F-59655 Villeneuve d'Ascq cedex, France<br>tel: +33 (0)3 20 43 41 63<br>fax: +33 (0)3 20 33 70 20<br>e-mail: andre.gomes@univ-lille1.fr |


## Academic degrees

| | |
|---|---|
| **PhD** | Universidade Estadual de Campinas, Campinas, Brazil, 2005 |
| **Thesis** | "Computational and Theoretical Developments of the Density Matrix Methods Applied to the Relativistic Generator Coordinate Method" |
| **Supervisor** | Prof. Dr. Rogério Custodio |
| **Funding agency** | Fundação de Amparo à Pesquisa de São Paulo (FAPESP), Brazil |
| **Master in Physical Chemistry** | Universidade Estadual de Campinas, Campinas, Brazil, 2001 |
| **Dissertation** | "Evaluation of Numerical Integration Methods in the Discretization of Integral Transforms" |
| **Supervisor** | Prof. Dr. Rogério Custodio |
| **Funding agency** | Fundação de Amparo à Pesquisa de São Paulo (FAPESP), Brazil |
| **Bachelor in Chemistry** | Universidade Estadual de Campinas, Campinas, SP, Brazil, 1998 |

## Scientific Employment

| | |
|---|---|
| **01-10-2013 – present** | Chargé de Recherche 1e Classe au CNRS. |
| **01-10-2009 – 30-09-2013** | Chargé de Recherche 2e Classe (2009: stagiaire, 2010: titulaire) au CNRS. |
| **01-04-2009 – 30-09-2009** | Postdoctoral researcher, Université de Lille 1. |
| **01-04-2007 – 31-03-2009** | Postdoctoral researcher, VU University Amsterdam, The Netherlands. |
| **01-04-2006 – 31-03-2007** | Postdoctoral researcher, Université de Strasbourg (Univ. Louis Pasteur) |
| **01-05-2005 – 31-03-2006** | Postdoctoral researcher, VU University Amsterdam, The Netherlands. |

## Additional training

| | |
|---|---|
| **17-07-2013 – 19-07-2013** | "EURACT-NMR Workshop 2013", Karlsruhe, Allemagne |
| **14-01-2013 – 18-01-2013** | "The ThUL School in Actinide Chemistry", Karlsruhe, Allemagne |
| **25-06-2012 – 29-06-2012** | "Initiation à MPI et OpenMP (dispensé par l'IDRIS)", Centre de Ressources Informatiques, Université de Lille 1, France |
| **16-06-2009 – 19-06-2009** | "5th Summer School on Actinide Science and Applications", Institute for Transuranium Elements (ITU), Karlsruhe, Allemagne |
| **29-06-2008 – 11-07-2008** | "The 10th Sostrup Summer School on Quantum Chemistry and Molecular Properties.", Himmelbjergegnens Natur- og Idraetsefterskole, Danemark |
| **14-12-2007 – 17-12-2007** | "Helsinki Winter School in Theoretical Chemistry", University of Helsinki, Department of Chemistry, Finland |

## Language proficiency

French (B1), Dutch (A1), English (fluent), Portuguese (native speaker).

## Adminstrative and teaching activities

### Teaching

3 "Informatique d'Instrumentation", 54h TP, 14h TD, DUT Mesures Physiques, IUT A de Lille, 04-2016 – 6-2016

2 "Initiation à la programmation", 58h equiv. TD, DUT Chimie, IUT A de Lille, 09-2012 – 12-2012

1 "Programmeren voor Chemici" (Programming for Chemists), 98h, VU University Amsterdam, 01-2009 – 03-2009

### Participation in recruitment committees

1 Assistant professor (MdC) position "Modélisation des propriétés optiques des fibres à cristal photonique innovantes" (section 30), UFR de Physique, Université de Lille 1, 2012

### Supervision

4 **10-2016 – 11-2019** Yassine BOUCHAFRA (doctoral co-supervision, supervisor: Dr. Valérie Vallet)
"Approches de sous-systèmes pour la modélisation de molécules contenant des éléments lourds dans solides et surfaces"
Funding: Ministère de L'enseignement supérieur et de la recherche

3 **02-2016 – 09-2016** Dr. Avijit SHEE (postdoctoral supervision)
"Accurate wavefunction approaches for the properties of heavy element compounds"
Funding: LABEX CaPPA

2 **10-2015 – 11-2018** Sophie KERVAZO (doctoral co-supervision, supervisor: Dr. Valérie Vallet)
"Modélisation de la thermodynamique des radioéléments volatiles par des nouvelles approches quantiques adaptées aux systèmes moléculaires fortement corrélés"
Funding: LABEX CaPPA

1 **02-2014 – 01-2016** Dr. Malgorzata OLEJNICZAK (postdoctoral supervision, co-supervisor: Dr. Valérie Vallet) "Theoretical modelling of optical and thermodynamical properties of actinide (U, Np, Pu,...) complexes"
Funding: LABEX CaPPA

# Scientific activities

## Code development

4 "DIRAC A relativistic electronic structure code"
`http://diracprogram.org/`

3 "XCFUN, "A library of approximate exchange-correlation functionals"
`https://github.com/dftlibs/xcfun`

2 "PYADF, A scripting framework for multiscale quantum chemistry
`http://www.pyadf.org`

1 "(PY)EMBED, A module to calculate embedding potentials based on the frozen density ansatz"
`http://gitlab.pyadf.org/workflows/embed`
`http://gitlab.pyadf.org/workflows/PyEmbed`

## Organisation of conferences

3 Local organizing committee, "46th EGAS (European Group on Atomic Systems)", Lille, France, 2014

2 Local organizing committee, "12th International Conference on Density Functional Theory and its Applications (DFT2007)", Amsterdam, The Netherlands, 2007.

1 Local organizing committee, "International Conference on Relativistic Effects in Heavy-Element Chemistry and Physics", Ottrott, France, 2007.

## Peer-reviewing activities

2 Peer-reviewing for "Structural Chemistry", "Inorganic Chemistry" and "The Journal of Chemical Physics".
Details at **Publons.com**

1 Peer-reviewing for funding proposals, "Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco" (FACEPE)

## Stays in other institutions

1 Visiting Faculty, Department of Chemistry, Rutgers University Newark, Newark, USA, 25-08-2014 – 11-09-2014
Contact: Prof. Dr. Michele Pavanello

## Participation in research projects

9 **CPER CLIMBIO: Physiques et Chimie de l'Environnement Atmosphérique (2012 – 2020)**
Funding: Agente Nationale de la Recherche (France) / Investissements d'avenir
PI: PRES Université Lille Nord de France / LOA – Laboratoire d'Optique Atmosphérique et PC2A – Laboratoire de Physiochimie des Processus de Combustion et de l'Atmosphère

8 **Novel theoretical approaches to the electronic structure of actinides compounds**
Funding: Fonds France Canada pour la Recherche
PIs: Valérie Vallet (CNRS / Université de Lille), Paul W. Ayers (McMaster University, Canada)

7 **Vers une modélisation théorique précise des effets d'environnement sur des spectres éléctroniques des actinides en phase condensée (2014-2016)**
Funding: CNRS (Campagne PICS 2013, PICS 6386) PI: André Severo Pereira Gomes (CNRS / Université de Lille)

6 **Talisman: Transnational Access to Large Infrastructure for a Safe Management of ActiNide (2013-2016)**
Funding: European Union
PI: Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), France

5 **LABEX CaPPA: Physiques et Chimie de l'Environnement Atmosphérique (2012 – 2020)**
Funding: Agente Nationale de la Recherche (France) / Investissements d'avenir
PI: PRES Université Lille Nord de France / LOA – Laboratoire d'Optique Atmosphérique et PC2A – Laboratoire de Physiochimie des Processus de Combustion et de l'Atmosphère

4 **Towards an accurate description of environment effects on the electronic spectra of actinides in condensed phase (2011 – 2012)**
Funding: ACTINET Integrated Infrastructure Initiative (European Union)
PI: André Severo Pereira Gomes (CNRS / Université de Lille)

3 **Highly accurate potential surfaces for quantum dynamics simulations of triiodide $I_3^-$ (2007)**
Funding: DEISA Extreme Computing Initiative (European Union).
PI: Lucas Visscher (VU University Amsterdam)

2 **Seamless Computational Chemistry (2006 – 2010)**
Funding: Nederlandse Organisatie voor Wetenschappelijke Onderzoek (The Netherlands).
PI: Lucas Visscher (VU University Amsterdam)

1 **ANR NCPMOL "Non-conservation de la parité moléculaire" (2005 — 2008)**.
Funding: Agente Nationale de la Recherche (France)
PI: Christian Chardonnet (CNRS/Université Paris XIII)

## Talks in (inter)national conferences and workshops

9 **"Calculation of Second-order Magnetic Properties with Frozen Density Embedding in the Relativistic Framework", ISTCP 2016 Conference, Symposium "Subsystems in Density Functional Theory" (invité), North Dakota, USA, 22-07-2016**

8 "WFT-in-DFT Embedding with Coupled-Cluster Wavefunctions", CFCAM discussion meeting"Excited states from small molecules to more complex systems", Marne-la-Vallée, 15-04-2014

7 **"WFT-in-DFT Embedding with Coupled-Cluster Wavefunctions", CECAM Workshop "Density-based embedding for multiscale simulations" (invité), Lausanne, 26-03-2014.**

6 "Towards Systematically Improvable Models for Actinides in Condensed Phase", Réunion plénière GDR CORREL "Méthodes corrélées pour la structure éléctronique", Paris, 29-11-2013

5 "Towards Systematically Improvable Models for Actinides in Condensed Phase", Excited states and complex environments – ESCE2013, Münster, Allemagne, 08-10-2013.

4 "Le spectre électronique des actinyls en phase condensée – explorations avec frozen density embedding" Réunion plénière GDR CORREL "Méthodes corrélées pour la structure éléctronique", Toulouse, 02-12-2011

3 **"Molecules in condensed phase, the ab initio way. Using embedding methods to calculated local properties of large-scale systems" (Invité), Réunion plénière GDR "Thermodynamique, Fragmentation et Agrégation de systèmes moléculaires complexes isolés", Paris, 23-11-2009.**

2 "Using WFT-in-DFT and DFT-in-DFT embedding to model actinide spectra in condensed phase", Very Heavy Metals 2009, Canet-Plage, 28-05-2009.

1 "A Framework for Execution of Computational Chemistry Codes in Grid Environments", International Conference on Computational Science 2006 (ICCS 2006), Reading, UK, 31-05-2006.

# Appendix B

# List of Publications

For further details please consult the profiles **ResearcherID**, **Google Scholar** ou **HAL**.

## Peer-reviewed journals

30 C Daday, ASP Gomes, C Filippi, CR Jacob, "Wavefunction embedding using exact freeze-thaw cycles", Journal of Chemical Physics, in preparation, 2017

29 A Shee, T Saue, L Visscher, ASP Gomes, "Four-component Equation-of-Motion Coupled Cluster excited state first-order properties", Journal of Chemical Physics, in preparation, 2017

28 A Shee, T Saue, L Visscher, ASP Gomes, "Four-component Equation-of-Motion Coupled Cluster for excitation energies and ionic states", Journal of Chemical Physics, in preparation, 2017

27 M Olejniczak, R Bast, ASP Gomes, "On the calculation of second-order magnetic properties using subsystem approaches in the relativistic framework", Physical Chemistry Chemical Physics, **19**, 8400, 2017.
http://dx.doi.org/10.1039/c6cp08561j

26 K Bogulslawski, F Réal, P Tecmer, C Duperrouzel, ASP Gomes, O Legeza, PW Ayers, V Vallet, "On the Multi-Reference Nature of Plutonium Oxides: $PuO_2^{2+}$, $PuO_2$, $PuO_3$ and $PuO_2(OH)_2$", Physical Chemistry Chemical Physics, **19**, 4317, 2017.
http://dx.doi.org/10.1039/c6cp05429c

25 F Réal, ASP Gomes, YOG Martínez, T Ayed, N Galland, M Masella, V Vallet, "Structural, dynamical and transport properties of the hydrated halides: How do $At^-$ bulk properties compare with those of the other halides, from $F^-$ to $I^-$?", Journal of Chemical Physics, **144**, 124513, 2016.
http://dx.doi.org/10.1063/1.4944613

24 M Denis, MS Norby, HJAa Jensen, ASP Gomes, MK Nayak, S Knecht, T Fleig, "Theoretical study on $ThF^+$, a prospective system in search of time-reversal violation", New Journal of Physics, **17**, 043005, 2015.
http://dx.doi.org/10.1088/1367-2630/17/4/043005

23 R Maurice, F Réal, ASP Gomes, V Vallet, G Montavon, N Galland, "Effective bond orders from two-step spin-orbit coupling approaches: The I2, At2, IO+, and AtO+ case studies", Journal of Chemical Physics, **142**, 094305, 2015.
http://dx.doi.org/10.1063/1.4913738

22 P Tecmer, ASP Gomes, S Knecht, L Visscher, "Communication: Relativistic Fock-space coupled cluster study of small building blocks of larger uranium complexes", Journal of Chemical Physics, **141**, 041107, 2014.
http://dx.doi.org/10.1063/1.4891801

21 ASP Gomes, F Real, N Galland, C Angeli, R Cimiraglia, V Vallet, "Electronic structure investigation of the evanescent $AtO^+$ ion", Physical Chemistry Chemical Physics, **16**, 9238, 2014.
http://dx.doi.org/10.1039/c3cp55294b

20 S Höfener, ASP Gomes, L Visscher, "Towards an efficient first-principles description of environment effects: solvatochromic shifts from coupled-cluster theory embedded in density functional theory", Journal of Chemical Physics, **139**, 104106, 2013.
http://dx.doi.org/10.1063/1.4820488

19 ASP Gomes, CR Jacob, F Real , L Visscher, V Vallet "Towards systematically improvable models for actinides in condensed phase: the electronic spectrum of uranyl in $Cs_2UO_2Cl_4$ as a test case", Physical Chemistry Chemical Physics, **15**, 15153, 2013.
http://dx.doi.org/10.1039/C3CP52090K

18 P Tecmer, H van Lingen, ASP Gomes, L Visscher, "The electronic spectrum of CUONg4 (Ng = Ne, Ar, Kr, Xe): New insights in the interaction of the CUO molecule with noble gas matrices", Journal of Chemical Physics **137**, 084308, 2012.
http://dx.doi.org/10.1063/1.4742765

17 S Höfener, ASP Gomes and L Visscher "Molecular properties via a subsystem density functional theory formulation: A common framework for electronic embedding", Journal of Chemical Physics, **136**, 044104 , 2012.
http://dx.doi.org/10.1063/1.3675845

16 CR Jacob, SM Beyhan, RE Bulo, ASP Gomes, AW Goetz, K Kiewisch, W Sikkema and L Visscher, "PyADF – A scripting framework for multiscale quantum chemistry", Journal of Computational Chemistry, **32**, 2328, 2011.
http://dx.doi.org/10.1002/jcc.21810

15 P Tecmer, ASP Gomes, Ekstrom U and L Visscher, "Electronic spectroscopy of $UO_2^{2+}$, $NUO^+$ and NUN: An Evaluation of Time-Dependent Density Functional Theory for actinides.", Physical Chemistry Chemical Physics, **13**, 6249, 2011.
http://dx.doi.org/10.1039/C0CP02534H

14 R Bast, A Koers, ASP Gomes, M Ilias, L Visscher, P Schwerdtfeger, T Saue. "Analysis of parity violation in chiral molecules", Physical Chemistry Chemical Physics, **13**, 864, 2010.
http://dx.doi.org/10.1039/c0cp01483d

13 ASP Gomes, L Visscher, H Bolvin, T Saue, T Fleig and E Eliav, "The Electronic Structure of the Triiodide ion from Relativistic Correlated Calculations: A Comparison of Different Methodologies", Journal of Chemical Physics, **133**, 064305, 2010.
http://dx.doi.org/10.1063/1.3474571

12 F De Montigny, R Bast, ASP Gomes, G Pilet, N Vanthuyne, C Roussel, L Guy, P Schwerdtfeger, T Saue and J Crassous, "Chiral oxorhenium(v) complexes as candiaus for the expérimental observation of molecular parity violation: a structural, synthetic and theoretical study", Physical Chemistry Chemical Physics, **12**, 8792, 2010.
http://dx.doi.org/10.1039/b925050f

11 ASP Gomes, KG Dyall and L Visscher, "Relativistic double-zeta, triple-zeta end quadruple-zeta basis sets for the lanthanide elements La-Lu", Theoretical Chemistry Accounts, **127**, 369, 2010.
http://dx.doi.org/10.1007/s00214-009-0725-7

10 ASP Gomes and KG Dyall "Revised double-zeta, triple-zeta, and quadruple-zeta basis sets for the 5d elements Hf-Hg." Theoretical Chemistry Accounts, **125**, 97, 2010.
http://dx.doi.org/10.1007/s00214-009-0717-7

9 F Real, ASP Gomes, V Vallet, L Visscher and E Eliav, "Benchmarking electronic structure calculations on the bare $UO_2^{2+}$ ion: How Different are Single and Multireference Electron Correlation Methods?", Journal of Physical Chemistry A, **113**, 12504, 2009.
http://dx.doi.org/10.1021/jp903758c

8 ASP Gomes, CR Jacob, and L Visscher, "Calculation of Excitation Energies in Large Systems by Embedding Wave Function Theory in Density Functional Theory", Physical Chemistry Chemical Physics, **10**, 5353, 2008
http://dx.doi.org/10.1039/b805739g.

7 I Infante, ASP Gomes, and L Visscher, "On the Performance of the Intermediate Hamiltonian Fock-Space Coupled-Cluster Method on Linear Triatomic Molecules: the Electronic Spectra of $NpO_2^+$, $NpO_2^{+2}$ and $PuO_2^{+2}$", Journal of Chemical Physics, **125**, 074301, 2006.
http://dx.doi.org/10.1063/1.2244564

6 R Custodio, ASP Gomes, FR Sensato, and JMS Trevas, "Analysis of Segmented Contraction of Basis Functions Using Density Matrix Theory", Journal of Computational Chemistry, **27**, 1822, 2006.
http://dx.doi.org/10.1002/jcc.20514

5 ASP Gomes, R Custodio, and L Visscher, "Systematic Sequences of Geometric Relativistic Basis Sets. I. s and p-Block Elements up to Xe". Theoretical Chemistry Accounts, **115**, 398, 2006.
http://dx.doi.org/10.1007/s00214-006-0120-6

4 ASP Gomes and R Custodio, "Polynomial Expansions of Basis Sets: Alternatives to Fully Optimized Exponents". Journal of Molecular Structure (THEOCHEM), **760**, 39, 2006.
http://dx.doi.org/10.1016/j.theochem.2005.11.002

3 ASP Gomes and L Visscher, "The Influence of Core Correlation on the Spectroscopic Constants of HAt", Chemical Physics Letters, **399**, 1, 2004.
http://dx.doi.org/10.1016/j.cplett.2004.09.132

2 ASP Gomes and R Custodio, "Exact Gaussian Expansions of Slater-Type Atomic Orbitals", Journal of Computational Chemistry, **23**, 1007, 2002.
http://dx.doi.org/10.1002/jcc.10090

1 ASP Gomes, PAM Vazquez, and LR Martins, "Técnicas de Análise do Perfil de Execução e Otimização de Programas em Química Computacional", Química Nova, **25**, 465, 2002
http://dx.doi.org/10.1590/S0100-40422002000300020

## Non peer-reviewed journals

1 ASP Gomes, CR Jacob, "Quantum-chemical embedding methods for treating local electronic excitations in complex chemical systems", Annual Reports Section C (Physical Chemistry) **108**, 222, 2012.
http://dx.doi.org/10.1039/C2PC90007F

## Communications in international conferences

1 ASP Gomes, A Merzky, and L Visscher, "A Framework for Execution of Computational Chemistry Codes in Grid Environments", Lecture Notes in Computer Science, **3993**, 97, 2006.
http://dx.doi.org/10.1007/11758532_15

## Book chapters

1 ASP Gomes, F Real, B Schimmelpfennig, U Wahlgren, V Vallet, "Applied Computational Actinide Chemistry", in "Computational Methods in Lanthanide and Actinide Chemisty", 1st edition, Michael Dolg (Ed.), 2015, John Wiley and Sons.

# Appendix C

# Calculations on Gas-phase species

This chapter presents the papers dealing with the application of electronic structure methods such as IHFSCC, CASPT2 or MRCI in order to understand their accuracy with respect to experimental results and their relative accuracy, on the basis of alculations on relatively small species in the gas-phase. In C.1 we disciss the use of IHFSCC to calculate the $f - f$ spectra of plutonyl and neptunyl with the DC Hamiltonian for the first time in the literature.

The same approach is used to look at the other actinyl species in C.3 and investigate the differences of IHFSCC and LRCC, or in C.2 to compare IHFSCC, CASPT2 and MRCI the $I_3^-$.

The performance of different density functional approximations for excitation energies and ionization potentials for uranyl and two of its isoelectronic species (NUN and $NUO^+$)is discussed in C.4 with the spin-free DC Hamiltonian and, later, with the advent of accurate 2-component approaches, this question is revisited in C.5. The performance of DFT and correlated wavefunctions is again discussed in C.6 for halide oxides species IO and AtO and the respective cations.

Finally, the question of accurately establishing the nature of the ground-state for the $ThF^+$ molecule, a prospective system for the experimental determinantion of the electron electric dipole moment, is addressed in C.7.

## C.1 Paper I

THE JOURNAL OF CHEMICAL PHYSICS **125**, 074301 (2006)

# On the performance of the intermediate Hamiltonian Fock-space coupled-cluster method on linear triatomic molecules: The electronic spectra of $NpO_2^+$, $NpO_2^{2+}$, and $PuO_2^{2+}$

Ivan Infante, André Severo Pereira Gomes, and Lucas Visscher[a)]
*Section Theoretical Chemistry, Faculty of Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*

(Received 23 May 2006; accepted 6 July 2006; published online 16 August 2006)

In this paper we explore the use of the novel relativistic intermediate Hamiltonian Fock-space coupled-cluster method in the calculation of the electronic spectrum for small actinyl ions ($NpO_2^+$, $NpO_2^{2+}$, and $PuO_2^{2+}$). It is established that the method, in combination with uncontracted double-zeta quality basis sets, yields excitation energies in good agreement with experimental values, and better than those obtained previously with other theoretical methods. We propose the reassignment of some of the peaks that were observed experimentally, and confirm other assignments. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2244564]

## I. INTRODUCTION

Considerable attention has been paid to actinide chemistry in recent years, due to the need to find new techniques for storage and reprocessing of spent nuclear fuel.[1–3] One of the most important steps of the plutonium uranium extraction (PUREX) process remains the separation of uranium(VI), plutonium(IV), and neptunium(VI) from fission products with aid of the tributylphosphate (TBP) extractant. In this process $Pu^{4+}$ is complexed with two nitrate ions and two TBP ligands, while the other two elements are extracted in the form of the triatomic actinyls $UO_2^{2+}$ and $NpO_2^{2+}$.[4]

The small size of these actinyls makes calculations feasible, and their energetical and structural parameters are reasonably well characterized.[5,6] Studies regarding the spectroscopic properties of actinyls focused mainly on the uranyl ion, but some studies have also been performed on neptunyl and plutonyl.[7–9] All three actinyl molecules have rather dense spectra due to the low-lying $5f$ and $6d$ orbitals localized on the metal. This characteristic poses a challenge to the currently available theoretical models, as they should describe the manifestation of relativistic effects as well as the multireference character of many of the states that significantly mix under the influence of spin-orbit coupling (SOC).

Among the theoretical methods that have already been used to investigate the spectra of small actinide compounds, single-reference coupled-cluster (CC) theory, both in its nonrelativistic and relativistic formulations, is arguably the most accurate method to calculate dynamic correlation energy. Its applicability is, however, severely limited due to its inability to handle states which have a considerable multireference character.[10] This has up to now left the spin-orbit complete-active-space second-order perturbation theory (SO-CASPT2) or spin-orbit configuration-interaction (SO-CI) methods as the only choices for qualitative or quantitative determination of spectra of neptunyl and plutonyl.

These methods, however, are not without important drawbacks, namely, (a) SO-CI methods are reasonably good for qualitative studies, but have difficulties in attaining quantitative agreement with experiment due to the lack of size extensivity in the electron correlation treatment and to restrictions on the number of configurations that can be included; and (b) SO-CASPT2, at present found to be the most accurate method employed for these systems, due to its ability to handle the use of larger basis sets, has a steep computational scaling with active space size. This limits the flexibility in choosing a suitable reference space and negatively affects the quality of the calculated spectra.

It is therefore of interest to assess new methods which could describe the electronic spectra of small actinyl ions accurately while still possessing a reasonably low scaling behavior. In this paper we explore the use of the intermediate Hamiltonian Fock-space coupled-cluster (IHFSCC) method[11–14] as an alternative to SO-CASPT2 and SO-CI. This method, while well established and routinely applied in high accuracy calculations of atomic transition energies, has scarcely been applied to molecular systems.[15]

The outline of the paper is as follows: in the Methodology section we outline the characteristics of the IHFSCC method and the computational procedure followed; in Results and Discussion, we first present a short analysis of the $f^1$ configurations, before moving on to our primary interest, the study of the electronic spectrum of the $f^2$ systems. The computed bond lengths and symmetric stretch frequencies for these gas phase model systems are also discussed, and we conclude by comparing the results of the IHFSCC method to other type of theoretical methods and experiments and discuss the merits and drawbacks of this method.

## II. METHODOLOGY

Fock-space coupled-cluster (FSCC) methods[16] have been quite successful in computing the excitation energies of atoms and molecules with very high accuracy.[17] The methods

[a)]Electronic mail: visscher@chem.vu.nl

scale like standard CC calculation [e.g., $O(N^6)$ for CCSD], but produce an effective Hamiltonian that, upon diagonalization, yields the energy of several states at once. In this family of methods, the IHFSCC approach[18] represents a breakthrough, as it greatly reduces the likelihood of intruder states and associated convergence problems in the solution of the CC equations.

The IHFSCC implementation used is that of the molecular four-component code DIRAC,[19,20] in which the $T_1$ and $T_2$ excitation operators are included, giving an IHFSCCSD approach that allows for creation of at most two holes and/or electrons outside the reference closed-shell system. While DIRAC can work with various relativistic Hamiltonians, in this application we have used the standard Dirac-Coulomb Hamiltonian, which is capable of describing the strong relativistic and SOC effects in actinyls. As has become common practice in the usage of DIRAC, we neglect contributions from the $(SS|SS)$-type integrals, replacing them by a simple correction,[21] and employ a Gaussian finite nucleus model[22] with this Hamiltonian. As with all FSCC methods, for the IHFSCC approach the reference state must be a single determinant. This means, in the case of $PuO_2^{2+}$ and $NpO_2^{+}$, that we start, respectively, from the $PuO_2^{4+}$ and $NpO_2^{3+}$ species and add the missing two open-shell electrons in the IHFSCC step. This amounts to the addition of two particles in the $P$ space, following the sequence

$$\left.\begin{matrix} NpO_2^{3+} \\ PuO_2^{4+} \end{matrix}\right\}(0h,0p) \rightarrow \left.\begin{matrix} NpO_2^{2+} \\ PuO_2^{3+} \end{matrix}\right\}(0h,1p) \rightarrow \left.\begin{matrix} NpO_2^{+} \\ PuO_2^{2+} \end{matrix}\right\}(0h,2p), \tag{1}$$

which is equivalent to calculating the first and second electron affinities for these highly charged systems. The restriction to two creation operators means that quintet states, important at higher energies, are not included. Such states belong to the $(1h,3p)$ sector of Fock space not yet available in the currently CCSD-based implementation.

The equilibrium geometries and harmonic frequencies for the ground and some of the excited states were determined by fitting tenth-order polynomials on discrete representations of the potential energy surfaces. As these molecules are known to be linear in their ground state, we have only considered the displacements along the symmetric stretch of the An–O bonds (An=Pu,Np). Under these circumstances, it is possible to exploit linear symmetry ($D_{\infty h}$) with the DIRAC program. Due to limitations in the computational resources, we did not explore displacements along the other vibrational modes, as the lowering of symmetry brought about by displacements along the asymmetric ($C_{\infty v}$) and bending modes ($C_{2v}$) would have increased the computational costs significantly.

Since the starting point in the self-consistent-field (SCF) calculations were ions with a +2 charge higher than the actual species, it was necessary to ensure that the ordering of the spinors, particularly in the highest occupied molecular orbital-lowest unoccupied molecular orbital (HOMO-LUMO) region, was consistent and adequate for the subsequent correlation treatment. This made us reorder the spinors in some cases, particularly for bond lengths larger than the equilibrium distance, in such a manner that the $5f$ shells in the starting species were always left empty.

The number of electrons correlated, apart from the two electrons that are included during the IHFSCC treatment, is 24; 10 in spinors of $g$ symmetry and 14 electron in spinors of $u$ symmetry. The virtual space was truncated by excluding spinors with energies larger than 6 a.u. This cutoff is consistent with our more extensive work on $UO_2$ that will be reported separately.[23]

For the IHFSCC calculations, a partitioning of the $P,Q$ spaces, hereby named "IH-u," was employed for all systems considered. In this partitioning 25 spinors of $u$ symmetry were included in the $P$ space. The $P$ space was further partitioned as follows: the six lowest-lying unoccupied $5f$ spinors from the Pu and Np atoms were included in the $P_m$ space, and the remaining $7p$ and $5f^{\sigma}$ spinors placed in the $P_i$ space. In the orthogonal $Q$ space all virtual spinors of $g$ symmetry, and the spinors of $u$ symmetry not included in the $P=P_m+P_i$, were included. To check convergence with active space size, for $PuO_2^{2+}$ and $NpO_2^{+}$ a second partitioning, hereby named "IH-u+g," was also explored. This consisted of the same $P_m$ space as in IH-u, but with 20 spinors of $g$ symmetry added to the $P_i$ space, in order to have a more balanced description of the $P$ space. The calculations with the latter are substantially more demanding and turned out to give negligible differences in excitation energies for states up to 30 000 $cm^{-1}$ relative to the IH-u space.

The basis set employed for the actinides was that of double-zeta quality developed by Dyall.[24] These sets were used in their uncontracted form and are of size $26s23p17d10f1g1h$. For oxygen the valence correlation-consistent triple-zeta (cc-pVTZ) set of Dunning[25] was used, also in uncontracted form. It should be noted that the TZ set was used here instead of the DZ set due to the need to add additional tight functions in a relativistic calculation that uses a nonrelativistic basis set. We also performed exploratory runs using a TZ quality basis set on the actinide element $(33s29p20d12f3g2h)$ on the equilibrium bond distances. As there was little variation upon enlargement (the excited states are shifted at most by about 200 $cm^{-1}$ for each of the excited states) and the computational cost for each point on the potential energy surface scan was greatly increased, we have opted to employ the uncontracted DZ set on the heavy element.

## III. RESULTS AND DISCUSSION

### A. Electronic structure of the ground state of $NpO_2^{3+}$ and $PuO_2^{4+}$

As previously noted, to study the electronic spectrum of $NpO_2^{+}$ and $PuO_2^{2+}$ using the IHFSCC approach one has to start from a closed-shell model molecule, and then proceed by computing the first and second electron affinities successively. Before discussing the results for $NpO_2^{+}$ and $PuO_2^{2+}$, we first analyze the relative ordering of the virtual orbitals of $NpO_2^{3+}$ and $PuO_2^{4+}$, as these give a first indication of the expected low-lying states of the $f^2$ ions. The ordering of the $5f$ orbitals is presented in Table I. It should be noted that the

TABLE I. Correlated electron affinities of the 5*f*, 7*s*, and 6*d* orbitals for the $UO_2^{2+}$, $NpO_2^{3+}$, and $PuO_2^{4+}$ molecules computed from sector (0,1) of the IHFSCC calculations. The correlation space was (24*e*/6 a.u.), with a (17*g*,20*u*) *P* model space. All calculations were performed at the equilibrium bond distance of 1.770 Å for $UO_2^+$, 1.701 Å for $NpO_2^{2+}$, and 1.645 Å for $PuO_2^{3+}$. The assignment of the experimental transitions (Ref. 37) in parentheses is uncertain. The excitation energies are given in $cm^{-1}$.

| | | | | | Experiment on $NpO_2^{2+}$ | | |
|---|---|---|---|---|---|---|---|
| Spin | $UO_2^+$ | $NpO_2^{2+}$ | $PuO_2^{3+}$ | $NpO_2^{2+}$ SO-CI[a] | $H_2O$[b] | $Cs_2UO_2Cl_4$[c] | $CsUO_2(NO_3)_3$[d] |
| $5f^{\phi}_{5/2u}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $5f^{\delta}_{5/2u}$ | 2 376 | 3 544 | 5 746 | 447 | ⋯ | 1000 | ⋯ |
| $5f^{\phi}_{7/2u}$ | 5 736 | 7 227 | 8 990 | 5515 | 6 760 | 6880 | 6459 |
| $5f^{\delta}_{5/2u}$ | 6 843 | 8 929 | 11 907 | 6565 | 81 80 | 7990 | 9420 |
| $6d^{\delta}_{3/2g}$ | 16 820 | 42 524 | 68 400 | ⋯ | ⋯ | ⋯ | ⋯ |
| $7s^{\sigma}_{1/2g}$ | 18 479 | 51 276 | 84 179 | ⋯ | ⋯ | ⋯ | ⋯ |
| $6d^{\delta}_{5/2g}$ | 20 642 | 47 443 | 74 493 | ⋯ | ⋯ | ⋯ | ⋯ |
| $5f^{\pi}_{1/2u}$ | 20 764 | 29 441 | 41 312 | ⋯ | (17 990) | ⋯ | ⋯ |
| $5f^{\pi}_{3/2u}$ | 24 535 | 33 856 | 45 747 | ⋯ | (21 010) | ⋯ | ⋯ |
| $7p^{\pi}_{1/2u}$ | 59 179 | 103 844 | 146 927 | ⋯ | ⋯ | ⋯ | ⋯ |
| $7p^{\pi}_{3/2u}$ | 62 170 | 107 330 | 152 277 | ⋯ | ⋯ | ⋯ | ⋯ |
| $5f^{\sigma}_{1/2u}$ | 69 317 | 85 553 | 107 476 | ⋯ | ⋯ | ⋯ | ⋯ |
| $7p^{\sigma}_{1/2u}$ | 74 148 | 119 663 | 164 003 | ⋯ | ⋯ | ⋯ | ⋯ |

[a]Reference 7.
[b]Reference 37.
[c]Reference 34.
[d]Reference 33.

energies, given in relative energies for ease of comparison with experimental results, are obtained not from the energies of the virtual Hartree-Fock orbitals of the closed-shell systems, but rather from the excitation energies, calculated with the IHFSCC method for the $f^1$ molecules. These energies are thus relative to the $^2\Phi_u$ ground state that results from occupying the $5f^{\phi}_{5/2}$ orbital. For comparison we also show the first electron affinities for the $UO_2^{2+}$ molecule, which is isoelectronic with $NpO_2^{3+}$ and $PuO_2^{4+}$.

A qualitative model for the electronic structure of actinyl ions has been given by Matsika and co-workers.[7,26] The lowest two unoccupied orbitals are the nonbonding $5f^{\phi}$ and $5f^{\delta}$ orbitals, which are fully localized on the metal ion. The higher unoccupied orbitals are the antibonding combinations of the oxygen 2*p* and metal $5f^{\pi}$ and $5f^{\sigma}$ orbitals that are also mainly localized on the metal. The nonbonding $6d^{\delta}$ orbital lies relatively high in neptunyl and plutonyl and only plays a role in explaining the observed spectral intensities.[26] In Fig. 1 we show the $f^1$ affinities (in absolute values) computed with the IHFSCC method. The increasing nuclear charge, going from U to Pu, leads to a stronger binding of the electrons in these isoelectronic systems. The lowering of the 5*f* level is the most pronounced and results in bringing the less shifted 6*d* and, in particular, the 7*s* and 7*p* levels, at too high energies to be of relevance in Np(V) and Pu(VI). All low-lying excited states are thus well described in terms of the $5f^2$ configuration only. In $UO_2$ the situation is different, since for U(IV) the 6*d*, 7*s*, and 7*p* orbitals all lie at similar energies as the 5*f*'s, yielding a denser and more complicated spectrum more easily perturbed by intermolecular interactions. The $UO_2$ molecule has attracted much attention lately since matrix spectroscopy studies indicate that the ground state in a neon matrix could differ from that in an argon matrix.[23,27–30] An investigation of $UO_2$ employing the IHFSCC method and including all relevant configurations ($5f^17s^1$, $5f^2$, $5f^16d^1$, and $5f^17p^1$) will be reported separately.[23]

In Table I, the available experimental values for $f^1$ $NpO_2^{2+}$ are also included. It should be noted that these absorption spectra were measured in water,[31] but studies in other polar solvents and crystals indicate that the transition energies are not much affected by the environment.[32–34] This

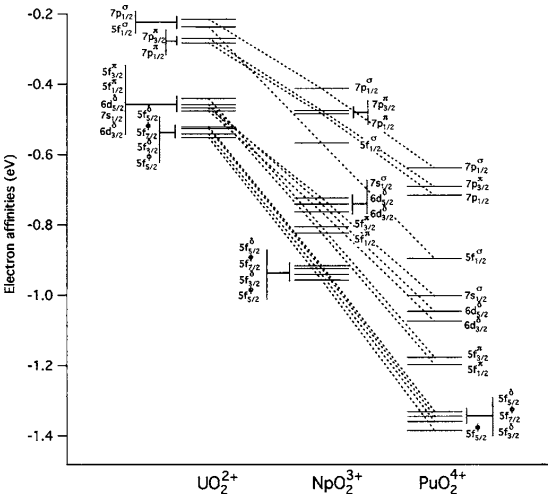



FIG. 1. Orbital energy diagrams including the 5*f*, 7*s*, 6*d*, and 7*p* orbitals for the $UO_2^{2+}$, $NpO_2^{3+}$, and $PuO_2^{4+}$ molecules. The correlated values from sector (0,1) of the IHFSCC calculations using the "IH-u" model space are shown. The bond distances are 1.770, 1.701, and 1.645 Å, respectively.

TABLE II. Vertical excitation energies ($E$, in $cm^{-1}$) for $NpO_2^+$, calculated with the IHFSCC method using the "IH-u" model space. The computed energies are evaluated at 1.701 Å (equilibrium bond length). Previous assignment based on SO-CI results from Matsika and Pitzer and the experimental data are also given (Refs. 8 and 37). The assignment of the experimental transitions in parentheses is uncertain.

| SO-CI Matsika and Pitzer[a] $r_e$=1.720 Å | | IHFSCC This work $r_e$=1.701 Å | | Expt.[b] | | Expt.[c] | |
|---|---|---|---|---|---|---|---|
| State | $E$ ($cm^{-1}$) | State | $E$ ($cm^{-1}$) | State | $E$ ($cm^{-1}$) | State | $E$ ($cm^{-1}$) |
| $4_g$ | 0 | $4_g$ | 0 | $^3H_4$ | 0 | 0 | |
| $0_g$ | 3 366 | $0_g$ | 2 527 | $\Sigma_0$ | ⋯ | $\Sigma_{0g}$ | |
| $5_g$ | 4 721 | $1_g$ | 4 102 | $\Pi_1$ | ⋯ | $\Pi_{1g}$ | ⋯ |
| $1_g$ | 4 938 | $5_g$ | 5 379 | $^3H_5$ | 6 173 | $^3H_{5g}$ | ⋯ |
| $6_g$ | 8 867 | $0_g$ | 8 628 | $^3\Pi_0$ | 8 953 | $^3\Pi_{0g}$ | 8 953 |
| $1_g$ | 9 076 | $1_g$ | 8 929 | $^3\Sigma_1$ | 9 146 | $^3\Sigma_{1g}$ | 9 116 |
| $0_g$ | 9 537 | $0_g$ | 9 378 | $^3\Pi_0$ | 9 780 | $^3\Pi_{0g}$ | 9 777 |
| $0_g$ | 9 708 | $6_g$ | 9 690 | $^3H_6$ | ⋯ | $^3H_{6g}$ | ⋯ |
| $2_g$ | 11 187 | $2_g$ | 10 056 | $^3\Pi_2$ | 10 208 | $^3\Pi_{2g}$ | 10 202 |
| | | | | vib. | 11 160 | vib. | 10 952 |
| $0_g$ | 14 415 | $0_g$ | 14 105 | $^3\Gamma_3$ | 13 020 | $^3\Phi_{2g}$ | 12 995 |
| $4_g$ | 15 249 | $4_g$ | 14 422 | $\Sigma_0$ | 13 824 | ⋯ | ⋯ |
| $1_g$ | 16 156 | $1_g$ | 15 031 | $^1\Gamma_4$ | 14 577 | $^1\Pi_{1g}$ | 14 558 |
| $0_g$ | 19 647 | $0_g$ | 16 551 | $^3\Sigma_1$ | 16 220 | $^3\Delta_{2g}$ | 16 221 |
| $1_g$ | 21 672 | $1_g$ | 18 992 | $^3\Phi_2$ | 16 100 | ⋯ | ⋯ |
| $5_g$ | 22 031 | $3_g$ | 19 735 | $\Delta_2$ | 16 906 | ⋯ | ⋯ |
| $1_g$ | 23 079 | $5_g$ | 19 761 | $\Sigma_0$ | ⋯ | ⋯ | ⋯ |
| $6_g$ | 23 327 | $6_g$ | 20 035 | $^3\Delta_1$ | 18 116 | ⋯ | ⋯ |
| $2_g$ | 23 649 | $2_g$ | 23 1877 | $\Pi_1$ | (19 360) | ⋯ | ⋯ |
| $3_g$ | 24 834 | $2_g$ | 23 322 | $^2I_6$ | 21 008 | ⋯ | 21 004 |
| $4_g$ | 26 592 | $4_g$ | 25 119 | $^3\Delta_3$ | (21 700) | ⋯ | ⋯ |
| ⋯ | ⋯ | $1_g$ | 25 436 | $^3\Pi_0$ | 22 600 | ⋯ | ⋯ |

[a]Reference 7.
[b]Reference 37.
[c]Reference 8.

suggests that the gas phase excitation energies should be a good approximation for the excited states occurring in solution. From our calculations, the energies of the third and fourth levels are in very good agreement with the excitation energies measured in water. The first excited state lies too low to be seen experimentally. The maximum error for these first two transition energies is 749 $cm^{-1}$. Comparing these values to those of Matsika and Pitzer,[7] obtained from SO-CI calculations, our results appear to be more accurate, as Matsika and Pitzer's show larger (about 1200–1500 $cm^{-1}$) discrepancies with respect to the experimental values.

Eisenstein and Pryce[35–37] interpreted the transitions at 17 990 and 21 010 $cm^{-1}$ as belonging to $f$-$f$ type excitations, occupying the $5f^{\pi}$ orbital. As our computed $5f^{\pi}$ energy is much higher, we believe that these transitions are more likely to be due to charge transfer states in which one of the $\sigma_u$ electrons in $NpO_2^{2+}$ is excited to a higher level. These transitions were not accessible in our calculations, as the current implementation of the method only considers Fock-space sectors that differ by two creation or annihilation operations from the reference space. Matsika and Pitzer[7] have computed energies of such charge transfer states and found them to lie within this experimental range.

### B. Electronic spectrum of $NpO_2^{2+}$ and $PuO_2^{2+}$

In $NpO_2^{2+}$ and $PuO_2^{2+}$ both open-shell electrons occupy the $5f^{\delta}_{3/2u}$ and the $5f^{\phi}_{5/2u}$ orbitals, resulting in a $4_g$ ground state. Since all lower excited states also belong to the $5f^2$ configuration, transitions between the ground and these excited states are electric dipole forbidden. The experimental spectra[37,38] are consistent with this picture, since most of the measured peaks have low intensity.[26,36,37,39–42] There is one intense peak at 10 204 $cm^{-1}$ for $NpO_2^+$ and at 12 037 $cm^{-1}$ for $PuO_2^{2+}$. The assignment of the spectrum is relatively easier for $PuO_2^{2+}$ than for $NpO_2^{2+}$, due to the larger splitting of the $5f$ orbitals in the former. Above 20 000 $cm^{-1}$ the assignment becomes less certain for both $NpO_2^{2+}$ and $PuO_2^{2+}$, as quintet charge transfer states also appear in this region.[7]

In Table II we present all the excitations up to 26 000 $cm^{-1}$ for $NpO_2^+$ and, in Table III, the excitations up to 34 000 $cm^{-1}$ for $PuO_2^{2+}$. The experimental spectra were originally interpreted by Eisenstein and Pryce[37] on the basis of semiempirical ligand field calculations. These assignments were later reconsidered on the basis of more accurate calculations.[7,9] However, even in these recent results, there were typical errors of a few thousand wave numbers, making some of the assignments still uncertain. Our new calculations

TABLE III. Vertical excitation energies ($E$, in $cm^{-1}$) for $PuO_2^{2+}$, calculated with IHFSCC method using the "IH-u" model space. The computed energies are evaluated at 1.645 Å (equilibrium bond length). For comparison the results of Maron *et al.* (Ref. 43) and Clavaguera-Sarrio *et al.* (Ref. 9) are shown, along with the experimental data (Reference 38).

| SDCI+Q+SO Maron *et al.*[a] $r_e$=1.699 Å | | CASPT2+SO Clavaguera-Sarrio *et al.*[b] $r_e$=1.677 Å | | IHFSCC This work $r_e$=1.645 Å | | Expt.[c] | |
|---|---|---|---|---|---|---|---|
| State | $E$ ($cm^{-1}$) | State | $E$ ($cm^{-1}$) | State | $E$ ($cm^{-1}$) | State | $E$ ($cm^{-1}$) |
| $4_g$ | 0 | $4_g$ | 0 | $4_g$ | 0 | $^3H_4$ | 0 |
| $0_g$ | 4 295 | $0_g$ | 4 190 | $0_g$ | 2 530 | $\Sigma_0$ | ⋯ |
| $5_g$ | 6 593 | $1_g$ | 6 065 | $1_g$ | 4 870 | $\Pi_1$ | ⋯ |
| $1_g$ | 7 044 | $5_g$ | 8 034 | $5_g$ | 6 700 | $^3H_5$ | ⋯ |
| $0_g$ | 7 393 | $0_g$ | 12 874 | $0_g$ | 10 334 | $^3\Pi_0$ | 10 185 |
| $6_g$ | 7 848 | $1_g$ | 12 906 | $1_g$ | 10 983 | $\Sigma_1$ | 10 500 |
| $0_g$ | 9 4115 | $6_g$ | 14 326 | $0_g$ | 11 225 | $^3\Pi_0$ | 10 700 |
| $1_g$ | 12 874 | $0_g$ | 14 606 | $6_g$ | 11 651 | $^3H_6$ | ⋯ |
| $2_g$ | 14 169 | $2_g$ | 14 910 | $0_g$ | 12 326 | $^3\Pi_2$ | 12 037 |
| | | | | | | vib. | 12 660 |
| $5_g$ | 16 984 | ⋯ | ⋯ | $0_g$ | 16 713 | $^1\Gamma_4$ | 15 420 |
| $4_g$ | 23 091 | ⋯ | ⋯ | $1_g$ | 17 737 | $\Sigma_0$ | 16 075 |
| $1_g$ | 27 005 | ⋯ | ⋯ | $4_g$ | 18 565 | $\Sigma_1$ | 17 800 |
| $6_g$ | 30 254 | ⋯ | ⋯ | $0_g$ | 20 029 | $^3\Gamma_3$ | 19 100 |
| $3_g$ | 33 164 | ⋯ | ⋯ | $1_g$ | 22 703 | $\Sigma_0$ | 19 810 |
| $0_g$ | 33 314 | ⋯ | ⋯ | $6_g$ | 22 889 | $^3\Phi_2$ | 22 200 |
| $4_g$ | 33 318 | ⋯ | ⋯ | $5_g$ | 23 022 | $^1H_5$ | 21 840 |
| $3_g$ | 33 366 | ⋯ | ⋯ | $3_g$ | 29 710 | $\Pi_1$ | ⋯ |
| $2_g$ | 33 388 | ⋯ | ⋯ | $2_g$ | 32 198 | $\Delta_2$ | ⋯ |
| $1_g$ | 34 520 | ⋯ | ⋯ | $0_g$ | 32 759 | $^3\Gamma_4$ | ⋯ |
| $0_g$ | 35 210 | ⋯ | ⋯ | $1_g$ | 34 080 | $^1I_6$ | ⋯ |
| $2_g$ | 35 670 | ⋯ | ⋯ | $4_g$ | 34 702 | $^1I_6$ | ⋯ |
| $1_g$ | 36 703 | ⋯ | ⋯ | $2_g$ | 34 982 | $^3\Delta_1$ | ⋯ |

[a]Reference 43.
[b]Reference 9.
[c]Reference 38.

improve upon the excitation energies computed previously since we include all relativistic effects from the start and could also correlate more electrons, but a shortcoming is that we are not yet able to calculate oscillator strengths with the current IHFSCC implementation. We therefore resorted to estimating the shape and intensities of the expected peaks on the basis of the composition of the excited states, shown in Table IV. Eisenstein and Pryce[36] have previously argued that transitions between states that differ only on the sign of the $z$ component of the angular momentum, $L_z$, of one of the two unpaired electrons have to be narrow. This occurs because the charge distribution remains basically unchanged when going from the ground to the excited state. For excitations that involve a change of the absolute value of $L_z$, the peaks are broader due to vibrational excitations. This type of reasoning, combined with the fact transitions to doubly excited states should have a low intensity, gives sufficient information to assign the spectra of $NpO_2^+$ and $PuO_2^{2+}$ on the basis of our data.

From Table IV one can furthermore see the high degree of similarity of the two isoelectronic actinyl ions. There are in general only slight differences in the values of the contributions from different configurations (for instance, the ground state of $PuO_2^{2+}$ is more mixed than the $NpO_2^+$ ion, with more contribution of the higher-lying $5f^{\pi}_{3/2u}$ orbital), so the two spectra can be discussed together. To better structure the discussion about the assignments, we have divided the spectra into three regions, each possessing some features that are used for the interpretation of the experiment.

### 1. *Region I: from* 0 to 7000 $cm^{-1}$

These three excited states differ by a single excitation from the ground state. In all cases there is a dominant determinant in which one electron is found in either the $5f^{\delta}_{3/2u}$ or $5f^{\phi}_{5/2u}$ orbital, both of which are occupied in the ground state. This region is not well sampled experimentally and therefore a clear comparison with our calculated data cannot be given. We confirm the original assignment of the peak at 6173 $cm^{-1}$ for the $NpO_2^+$ ion as a $5_g$ state[37] with a slightly lower computed energy of 5379 $cm^{-1}$ for this $4_g \rightarrow 5_g$ transition.

### 2. *Region II: from* 7000 to 13 000 $cm^{-1}$

In this region we find an excellent agreement with the experimental transition energies for both the neptunyl and plutonyl ions, with errors of about few hundred wave num-

TABLE IV. Composition (in %) of the ground and some of the lowest excited states for $NpO_2^+$ and $PuO_2^{+2}$, together with the spinors occupied in the different IHFSCC sectors with respect to the closed shell species $NpO_2^{3+}$ and $PuO_2^{4+}$. All values are obtained at the calculated equilibrium geometries ($r_e$=1.701 and 1.645 Å, respectively) for the "IH-u" model space.

| | IHFSCC configuration | | Weight (%) | |
|---|---|---|---|---|
| State | $(0h,2p)$ | $(0h,1p)$ | $NpO_2^{+2}$ | $PuO_2^{+2}$ |
| $4_g$ | $5f^{\delta}_{3/2u}$ | $5f^{\phi}_{5/2u}$ | 94 | 81 |
| | $5f^{\pi}_{3/2u}$ | $5f^{\phi}_{5/2u}$ | 4 | 16 |
| $0_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\phi}_{-5/2u}$ | 59 | 70 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\delta}_{-3/2u}$ | 32 | 18 |
| $1_g$ | $5f^{\delta}_{-3/2u}$ | $5f^{\phi}_{5/2u}$ | 80 | 71 |
| | $5f^{\phi}_{-5/2u}$ | $5f^{\phi}_{7/2u}$ | 11 | 12 |
| | $5f^{\delta}_{-3/2u}$ | $5f^{\delta}_{5/2u}$ | 4 | 13 |
| $5_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\delta}_{5/2u}$ | 55 | 56 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\phi}_{7/2u}$ | 43 | 36 |
| $0_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\delta}_{-5/2u}$ | 49 | 49 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{-5/2u}$ | 49 | 49 |
| $1_g$ | $5f^{\delta}_{-3/2u}$ | $5f^{\delta}_{5/2u}$ | 55 | 41 |
| | $5f^{\phi}_{-5/2u}$ | $5f^{\phi}_{7/2u}$ | 28 | 37 |
| | $5f^{\delta}_{-3/2u}$ | $5f^{\phi}_{5/2u}$ | 12 | 10 |
| $0_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\delta}_{-5/2u}$ | 29 | 27 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{-5/2u}$ | 29 | 27 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\delta}_{-3/2u}$ | 24 | 18 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\delta}_{-5/2u}$ | 6 | 6 |
| $6_g$ | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{7/2u}$ | 67 | 57 |
| | $5f^{\phi}_{5/2u}$ | $5f^{\phi}_{7/2u}$ | 33 | 43 |
| $2_g$ | $5f^{\delta}_{-3/2u}$ | $5f^{\phi}_{7/2u}$ | 93 | 82 |
| | $5f^{\mathrm{H}}_{-3/2u}$ | $5f^{\phi}_{7/2u}$ | 4 | 16 |
| $0_g$ | $5f^{\delta}_{5/2u}$ | $5f^{\delta}_{-5/2u}$ | 31 | 20 |
| | $5f^{\phi}_{5/2u}$ | $5f^{\phi}_{-5/2u}$ | 25 | 23 |
| | $5f^{\phi}_{5/2u}$ | $5f^{\delta}_{-5/2u}$ | 14 | 16 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{-5/2u}$ | 14 | 16 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\delta}_{-3/2u}$ | 11 | 15 |
| $4_g$ | $5f^{\phi}_{3/2u}$ | $5f^{\delta}_{5/2u}$ | 83 | 72 |
| | $5f^{\pi}_{3/2u}$ | $5f^{\delta}_{5/2u}$ | 5 | 18 |
| $1_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\phi}_{-7/2u}$ | 43 | 43 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{-7/2u}$ | 38 | 28 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\delta}_{-5/2u}$ | 11 | 19 |
| $0_g$ | $5f^{\phi}_{7/2u}$ | $5f^{\phi}_{-7/2u}$ | 35 | 47 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\delta}_{-5/2u}$ | 27 | 33 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\delta}_{-3/2u}$ | 21 | 12 |
| $1_g$ | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{-7/2u}$ | 56 | 65 |
| | $5f^{\phi}_{5/2u}$ | $5f^{\phi}_{-7/2u}$ | 23 | 14 |
| | $5f^{\delta}_{3/2u}$ | $5f^{\delta}_{-5/2u}$ | 20 | 19 |
| $3_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\pi}_{1/2u}$ | 96 | 97 |
| $5_g$ | $5f^{\delta}_{3/2u}$ | $5f^{\phi}_{7/2u}$ | 55 | 50 |
| | $5f^{\phi}_{5/2u}$ | $5f^{\delta}_{5/2u}$ | 44 | 43 |
| $6_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\phi}_{7/2u}$ | 67 | 56 |
| | $5f^{\delta}_{5/2u}$ | $5f^{\phi}_{7/2u}$ | 33 | 43 |
| $2_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\pi}_{-1/2u}$ | 91 | 96 |
| $2_g$ | $5f^{\delta}_{3/2u}$ | $5f^{\pi}_{1/2u}$ | 93 | 80 |
| | $5f^{\pi}_{3/2u}$ | $5f^{\pi}_{1/2u}$ | 2 | 19 |
| $4_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\pi}_{3/2u}$ | 58 | 62 |
| | $5f^{\phi}_{7/2u}$ | $5f^{\pi}_{1/2u}$ | 37 | 20 |
| $1_g$ | $5f^{\phi}_{5/2u}$ | $5f^{\pi}_{-3/2u}$ | 89 | 73 |
| | $5f^{\phi}_{5/2u}$ | $5f^{\delta}_{-3/2u}$ | 2 | 21 |

TABLE IV. (*Continued.*)

bers. The characteristic feature in both spectra is the intense peak that appears 10 204 $cm^{-1}$ for $NpO_2^+$ and at 12 037 $cm^{-1}$ for $PuO_2^{2+}$. A mechanism that can explain the intensity of this dipole-forbidden transition is described in detail by Matsika *et al.*,[26] who considered systems with one, three, and five chloride ions in the equatorial plane. Their calculations show that the ligand field from the latter arrangement gives sufficient mixing of the $5f^{\phi}$ and $6d^{\delta}$ to cause an intense $^3H_{4g} \rightarrow {}^3\Pi_{2g}$ transition.

From the decomposition given in Table IV it is clear that this $2_g$ state for $NpO_2^+$ is dominated by a single determinant, accounting for 93% of the total wave function. With respect to the ground state configuration, this state corresponds to the excitation of an electron from the $5f^{\phi}_{5/2u}$ to the $5f^{\phi}_{-7/2u}$ orbital. This is also the case for $PuO_2^{2+}$, for which the weight of the relevant determinant in the $2_g$ state is slightly smaller (at 83% of the total wave function), corroborating the assignment of Matsika *et al.*.

The $6_g$ state is found close to the $2_g$ state, but it is unclear whether transitions to this state have enough intensity to be detected. Eisenstein and Pryce[37] suggested that the peak at 11 160 $cm^{-1}$ is either due to this state or to vibrational progression of the $2_g$ transition. Our analysis shows a $6_g$ wave function dominated by two determinants, where one with the highest weight corresponds to a double excitation from the ground state. Combined with the fact that a transition energy below 10 000 $cm^{-1}$ was obtained, we conclude that the assignment of the 11 160 $cm^{-1}$ peak to $6_g$ is unlikely, and that the interpretation as a vibrational band is probably correct.

Regarding the assignment of the remaining peaks in region II, there are three other excited states, namely, $0_g$, $1_g$, and $0_g$, that could be contributing. They all arise from orbitals that have the same $\delta$ and $\phi$ characters as the ground state, but with different signs of the $L_z$ component (see Tables II and III). All the peaks should be narrow but differ in intensities. The calculations by Matsika *et al.* show that the transition to the $1_g$ state is more intense than the ones to the $0_g$ states. This leads to the conclusion that the peaks (at 9146 $cm^{-1}$ for $NpO_2^+$ and at 10 500 $cm^{-1}$ for $PuO_2^{2+}$) should be assigned to the $1_g$ state. While this interpretation had already been put forward with a good deal of certainty in previous works,[26,37] the IHFSCC results serve as a litmus test for this assignment as we can compare the spacing of the computed and observed peaks.

In the experimental spectrum of $NpO_2^+$ (Fig. 1 of Ref. 26) three narrow peaks are visible to the right (higher wave length) of the strong $2_g$ transition. The lowest energy transitions are separated by only 163 $cm^{-1}$, while the higher energy transitions appear as a well resolved shoulder on the $2_g$ transition at 9780 $cm^{-1}$. Of the three peaks, the middle one is clearly the most intense. The relative energies of the $0_g$, $1_g$, and $0_g$ states are indeed consistent with this spectrum, with the $1_g$ appearing in the middle separated by 301 $cm^{-1}$ from the lower $0_g$ state and by 449 $cm^{-1}$ from the higher $0_g$ state. The deviations from the experimental peak positions are thus maximally 400 $cm^{-1}$, which should be considered very good agreement for a gas phase model. In the less resolved $PuO_2^{2+}$ spectrum,[37] the $0_g$, $1_g$, and $0_g$ states lie practically in the same band, with the $1_g$ peak at 10 500 $cm^{-1}$. This peak has one left shoulder, almost completely resolved at 10 185 $cm^{-1}$, and one right shoulder, hidden in the $1_g$ at 10 700 $cm^{-1}$. In our calculations a similar trend is found, with the lower $0_g$ and $1_g$ states again separated by a somewhat larger value (649 $cm^{-1}$) than the spacing that is experimentally observed (315 $cm^{-1}$). The calculated upper $0_g$ state is only 240 $cm^{-1}$ higher than the $1_g$, which is in very good agreement with the fit of the experimental data (where a distance of about 200 $cm^{-1}$ is given).

#### 3. Region III: above 13 000 $cm^{-1}$

For the higher excited state agreement with experiment cannot be expected to be as good, as there are larger effects due to the surroundings, and the possible presence of charge transfer states. Looking at the experimental spectra,[37,38] in the region we find for both neptunyl and plutonyl peaks with qualitatively similar shapes, with the most intense transition at about 16 000 $cm^{-1}$ surrounded by satellite shoulders. For $NpO_2^+$ these shoulders are resolved and narrow, while for $PuO_2^{2+}$ they are quite broad.

In our calculations we find five excited states ($0_g$, $4_g$, $1_g$, $0_g$, and $1_g$,) in this region, mainly made up by determinants containing $\delta$ and $\phi$ electrons in open shells. Based on the arguments put forth at the beginning of this section, this means that the associated peaks should be narrow. The oscillator strengths calculated by Matsika *et al.*[26] indicate that the most intense of these peaks should be the $1_g$ state. Our calculations place this state at 15 031 $cm^{-1}$ for $NpO_2^+$ and at 17 737 $cm^{-1}$ for $PuO_2^{2+}$, whereas the experimental positions are almost the same for both ions (16 220 and 16 075 $cm^{-1}$, respectively).

Matsika *et al.*, however, suggested that this peak results from a transition to a $^3\Delta_{2g}$ state arising from occupation of the $5f^{\pi}$ orbital. As already discussed in the previous section on $NpO_2^{2+}$, the $5f^{\pi}$ orbital is at a rather high energy relative to the $5f^{\phi}$ and $5f^{\delta}$. Consequently, all states with significant $5f^{\pi}$ character are found too high in energy (around 20 000 $cm^{-1}$ for $NpO_2^+$ and 30 000 $cm^{-1}$ for $PuO_2^{2+}$) to be associated with transitions at 16 000 $cm^{-1}$. While this may be an artifact of our gas phase model, it could also be that the the observed transition is to the $1_g$ state, rather than the $2_g$ state. This is particularly the case for $PuO_2^{2+}$, where it does not seem probable that the the surrounding water molecules lower this metal-to-metal transition to half the gas phase value. We therefore propose to reassign this transition to the $1_g$ state.

Another reassignment may be necessary for the experimental peak at 13 020 $cm^{-1}$ for $NpO_2^+$. This peak was previously assigned to a $3_g$ state by Eisenstein and Pryce,[37] and later to a $2_g$ state by Matsika *et al.*[26] In both cases, the composition of this excited state included a $5f^{\pi}$ orbital that we anticipate to get occupied only at much higher energies. It is difficult to assign these peaks with certainty, because the differences in energies involved are rather small. We notice, however, that the calculated $4_g$ state lies at lower energies than the more intense $1_g$ for $NpO_2^+$, while appearing at higher energies for $PuO_2^{2+}$. This agrees with the experimental spectra, where one small peak at lower energies than the state we assigned as the $1_g$ is found in the plutonyl spectrum, whereas two peaks are found for neptunyl.

Given the uncertainties related to the position of charge transfer peaks (found slightly above 20 000 $cm^{-1}$ in the calculations of Matsika and Pitzer[7]), we do not attempt to match our computed excitation energies at higher energies with the experimental data.

#### 4. Comparison with previous calculations

Comparing our computed excitation energies for $NpO_2^+$ with those of Matsika and Pitzer,[7] we see that a more rigorous treatment of electron correlation and relativistic effects indeed results in smaller deviations from experiment. This is so for the lower excited states (below 10 000 $cm^{-1}$), but also for most of the higher states, especially the important $2_g$ state, which differs from experiment by less than 200 $cm^{-1}$, compared to over 1500 $cm^{-1}$ for previous calculations.

More theoretical calculations are available for the plutonyl ion, so the relative accuracy of our results and the strenghts and weaknesses of the IHFSCC method can better be assessed. The calculations of Maron *et al.* [43] and of Clavaguéra-Sarrio *et al.*[9] give rise to a rather similar assignment of the lower excited states, but report excitation energies quite different from ours and from experiment. For instance, the results of Maron *et al.*[43] underestimate the low-lying transitions (region I) and strongly overestimate the upper states (region III), with discrepancies with respect to the experimental transitions of more than 10 000 $cm^{-1}$. Our calculations show errors on the 1000–2000 $cm^{-1}$ range for these states. The later calculations of Clavaguéra-Sarrio *et al.* are better than those of Maron *et al.*[43] for region II states, but their errors are still quite large (more than 2000 $cm^{-1}$) when compared to what can be achieved with the IHFSCC method that shows deviations of about 500 $cm^{-1}$.

### C. Potential energy curves

As the IHFSCC method allows the determination of multiple states available in a single calculation, it was quite easy to determine the equilibrium bond distances and vibrational symmetric stretch frequencies for a number of different states. These quantities are shown in Table V. An important difference between these results and those of previous calculations is the difference of the bond lengths for the

TABLE V. Bond lengths (in Å) and harmonic frequencies (in $cm^{-1}$) for the first 14 states of $NpO_2^+$ and $PuO_2^{2+}$, derived from potential energy surfaces for the symmetric stretch. These surfaces were obtained with the IHFSCC method using the "IH-u" model space.

| | $NpO_2^+$ | | | $PuO_2^{2+}$ | |
|---|---|---|---|---|---|
| State | $r_c$ (Å) | $\omega_e$ ($cm^{-1}$) | State | $r_c$ (Å) | $\omega_e$ ($cm^{-1}$) |
| $4_g$ | 1.700 | 1073 | $4_g$ | 1.643 | 1144 |
| $0_g$ | 1.701 | 1061 | $0_g$ | 1.654 | 1048 |
| $1_g$ | 1.699 | 1069 | $1_g$ | 1.643 | 1162 |
| $5_g$ | 1.699 | 1080 | $5_g$ | 1.637 | 1334 |
| $0_g$ | 1.699 | 1081 | $0_g$ | 1.637 | 1361 |
| $1_g$ | 1.698 | 1072 | $1_g$ | 1.637 | 1351 |
| $0_g$ | 1.695 | 1082 | $0_g$ | 1.634 | 1324 |
| $6_g$ | 1.701 | 1075 | $6_g$ | 1.642 | 1087 |
| $2_g$ | 1.697 | 1063 | $2_g$ | 1.636 | 1281 |
| $0_g$ | 1.691 | 1093 | $0_g$ | 1.637 | 1332 |
| $4_g$ | 1.698 | 1086 | $1_g$ | 1.630 | 1284 |
| $1_g$ | 1.701 | 1083 | $4_g$ | 1.636 | 1400 |
| $0_g$ | 1.701 | 1179 | $0_g$ | 1.637 | 1439 |
| $1_g$ | 1.724 | 1564 | $1_g$ | 1.640 | 1377 |

ground state of both molecules. For neptunyl, the bond length is about 0.02 Å shorter than the value given by Matsika and Pitzer,[7] whereas for plutonyl the bond length is about 0.05 Å shorter than that reported by Clavaguéra-Sarrio *et al.*[9] and 0.03 Å shorter than that given by Maron *et al.*.[43] These differences could be due to the inclusion of $6p$ orbitals in the correlated active space in our calculations, allowing the oxo ligands to move closer to the actinide, but we have not investigated this in detail.

The differences in the calculated excited energies shown here and those of previous works decrease to some extent if IHFSCC calculations are performed at the corresponding equilibrium geometries, thus indicating that part of these discrepancies are due to geometrical effects. We observed, however, that also in these situations the IHFSCC calculations generally show a better agreement with experiment.

The harmonic frequencies of the ground state of neptunyl and plutonyl differ by about 69 $cm^{-1}$ (1073 and 1144 $cm^{-1}$), which is of course mainly due to the difference in charge. It is interesting that the frequencies for the low-lying excited states of neptunyl are very similar to that of the ground state whereas for plutonyl variations of up to 200–300 $cm^{-1}$ are seen. Comparison to experimental data is difficult as it is well known that solvation and complexation lower the vibrational frequencies of actinyls considerably.[44] Madic *et al.*[45] gave Raman data for these ions in aqueous solution. The difference in values for the symmetric stretch of $NpO_2^+$ and $PuO_2^{2+}$ (767 and 833 $cm^{-1}$, respectively) of 66 $cm^{-1}$ is remarkably similar to our computed gas phase difference of 71 $cm^{-1}$.

## IV. CONCLUSIONS

In this work we have investigated the ground and excited states of the actinyl ions $NpO_2^+$ and $PuO_2^{2+}$. While the spectra of these ions had been studied before, there was still a good deal of uncertainty with respect to the ordering and spacing of different electronic states. In this work we were able to improve upon previous calculations on both aspects. First, we have established with greater certainty that the experimentally most intense peak found for both the actinyl ions has a $2_g$ symmetry. Second, the average errors we obtain compared to previous calculations are much smaller so that more definite assignments of these spectra could be made. This is particularly important for the higher excited states, where results from previous calculations varied considerably.

The use of the IHFSCC method allowed for the economical determination of several electronic states at once, while accurately describing both static and dynamic correlation energies. The IHFSCC method in its current form, however, is not without drawbacks. What is important is the limitation on the Fock-space sectors that are implemented. For instance, by using only sector $(0h,2p)$ only triplet $f^2$ states can be described, making charge transfer states of the neptunyl ion unaccessible. For quintet states, a mixed sector $(1h,3p)$ must be employed, but it is yet to be implemented in the DIRAC code. Another drawback is related to issues of convergence, which still demand experimentation with the $P$, $Q$ partitioning, and prevent the method to be used in a "black-box" manner that is desired when using the method for larger and more complex systems.

## ACKNOWLEDGMENTS

The authors wish to thank the Netherlands Organization for Scientific Research for financial support via the "Jonge Chemici" programme and like to acknowledge stimulating discussions with Dr. Ephraim Eliav and Professor Uzi Kaldor on the application of their IHFSCC methods. This research was performed in part using the Molecular Science Computing Facility (MSCF) in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at the Pacific Northwest National Laboratory, operated for the Department of Energy by Battelle.

[1] K. L. Nash, Solvent Extr. Ion Exch. **11**, 729 (1993).
[2] K. L. Nash, R. E. Barrans, R. Chiarizia, M. L. Dietz, M. Jensen, and P. G. Rickert, Solvent Extr. Ion Exch. **18**, 605 (2000).
[3] E. P. Horwitz, D. G. Kalina, H. Diamond, G. F. Vandegrift, and W. W. Schulz, Solvent Extr. Ion Exch. **3**, 75 (1985).
[4] N. C. Rasmussen, *Technologies for Separations and Transmutations* (National Academy, Washington, 1996).
[5] J. P. Bladeau, S. A. Zygmunt, L. A. Curtiss, D. T. Reed, and B. E. Bursten, Chem. Phys. Lett. **310**, 347 (1999).
[6] C. Clavaguera-Sarrio, V. Brenner, S. Hoyau, C. J. Marsden, P. Millie, and J. P. Dognon, J. Phys. Chem. B **107**, 3051 (2003).
[7] S. Matsika and R. M. Pitzer, J. Phys. Chem. A **104**, 4064 (2000).
[8] S. Matsika, Z. Zhang, S. R. Brozell, J. P. Bladeau, and R. M. Pitzer, J. Phys. Chem. A **105**, 3825 (2001).
[9] C. Clavaguera-Sarrio, V. Vallet, D. Maynau, and C. J. Marsden, J. Chem. Phys. **121**, 5312 (2004).
[10] I. Infante and L. Visscher, J. Chem. Phys. **121**, 5783 (2004).
[11] A. Landau, E. Eliav, and U. Kaldor, Chem. Phys. Lett. **313**, 399 (1999).
[12] A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **113**, 9905 (2000).
[13] A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **115**, 6862 (2001).
[14] A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **121**, 6634 (2004).
[15] M. Musial, L. Meissner, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **122**, 224110 (2005).

[16] D. Mukherjee and S. Pal, *Use of Cluster Expansion Methods in the Open-Shell Correlation Problem* (Academic, New York, 1989).
[17] U. Kaldor, E. Eliav, and A. Landau, in *Relativistic Electronic Structure Theory-Part 1: Fundamentals*, edited by P. Schwerdtfeger (Elsevier, Amsterdam, 2002).
[18] J. P. Malrieu, P. Durand, and J. P. Daudey, J. Phys. A **18**, 809 (1985).
[19] L. Visscher, E. Eliav, and U. Kaldor, J. Chem. Phys. **115**, 9720 (2001).
[20] H. J. Aa. Jensen, T. Saue, L. Visscher *et al.*, DIRAC04, Release 4.1, a relativistic *ab initio* electronic structure program, 2004.
[21] L. Visscher, Theor. Chem. Acc. **98**, 68 (1997).
[22] L. Visscher and K. G. Dyall, At. Data Nucl. Data Tables **67**, 207 (1997).
[23] I. Infante, E. Eliav, L. Visscher, and U. Kaldor (unpublished).
[24] K. G. Dyall (unpublished); basis sets are available from the DIRAC website http://dirac.chem.sdu.dk
[25] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[26] S. Matsika, R. M. Pitzer, and D. T. Reed, J. Phys. Chem. A **104**, 11983 (2000).
[27] L. Gagliardi, M. C. Heaven, J. W. Krogh, and B. O. Roos, J. Am. Chem. Soc. **127**, 86 (2005).
[28] L. Gagliardi, B. O. Roos, P. A. Malmqvist, and J. M. Dyke, J. Phys. Chem. A **105**, 10602 (2001).
[29] M. F. Zhou, L. Andrews, N. Ismail, and C. Marsden, J. Phys. Chem. A **104**, 5495 (2000).
[30] T. Fleig, H. J. A. Jensen, J. Olsen, and L. Visscher, J. Chem. Phys. **124**, 104106 (2006).
[31] D. Cohen and B. Taylor, J. Inorg. Nucl. Chem. **22**, 151 (1962).
[32] H. A. C. McKay, J. S. Nairn, and M. B. Waldron, J. Inorg. Nucl. Chem. **7**, 167 (1958).
[33] R. G. Denning, J. O. Norris, and W. Brown, Mol. Phys. **46**, 325 (1982).
[34] R. G. Denning, J. O. Norris, and W. Brown, Mol. Phys. **46**, 287 (1982).
[35] J. C. Eisenstein and M. H. L. Pryce, Proc. R. Soc. London, Ser. A **229**, 20 (1965).
[36] J. C. Eisenstein and M. H. L. Pryce, J. Res. Natl. Bur. Stand., Sect. A **69**, 217 (1965).
[37] J. C. Eisenstein and M. H. L. Pryce, J. Res. Natl. Bur. Stand., Sect. A **70**, 165 (1966).
[38] *Gmelin Handbooks of Inorganic Chemistry, Transuranium Elements A2*, 8th ed. (Springer-Verlag, New York, 1973).
[39] R. Sjoblom and J. C. Hindman, J. Am. Chem. Soc. **73**, 1744 (1951).
[40] W. C. Waggener, J. Phys. Chem. **62**, 382 (1958).
[41] T. W. Newton and F. B. Baker, J. Phys. Chem. **61**, 934 (1957).
[42] R. H. Bettsand B. G. Harvey, J. Chem. Phys. **16**, 1089 (1948).
[43] L. Maron, T. Leininger, B. Schimmelpfennig, V. Vallet, J. L. Heully, C. Teichteil, O. Gropen, and U. Wahlgren, Chem. Phys. **244**, 195 (1999).
[44] G. S. Groenewold, A. K. Gianotto, K. C. Cossel, M. J. van Stipdonk, D. T. Moore, N. Polfer, J. Oomens, W. de Jong, and L. Visscher, J. Am. Chem. Soc. **128**, 4802 (2006).
[45] C. Madic, G. M. Begun, D. E. Hobart, and R. L. Hahn, Inorg. Chem. **23**, 1914 (1984).

## C.2 Paper II

THE JOURNAL OF CHEMICAL PHYSICS **133**, 064305 (2010)

# The electronic structure of the triiodide ion from relativistic correlated calculations: A comparison of different methodologies

André Severo Pereira Gomes,[1,2,3] Lucas Visscher,[2,a)] Hélène Bolvin,[3] Trond Saue,[3] Stefan Knecht,[4,3] Timo Fleig,[4,5] and Ephraim Eliav[6]
[1]*Laboratoire PhLAM, CNRS UMR 8523, Université de Lille 1, F-59655 Villeneuve d'Ascq Cedex, France*
[2]*Department of Theoretical Chemistry, Faculty of Sciences, Amsterdam Center for Multiscale Modeling, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*
[3]*Laboratoire de Chimie Quantique, Institut de Chimie de Strasbourg, UMR 7177 CNRS/Université Louis Pasteur, 4 rue Blaise Pascal, F-67000 Strasbourg, France*
[4]*Institute of Theoretical and Computational Chemistry, Heinrich-Heine-University Duesseldorf, Universitaetsstr. 1 D-40225 Duesseldorf, Germany*
[5]*Laboratoire de Chimie et Physique Quantiques, I.R.S.A.M.C., Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse, France*
[6]*School of Chemistry, Tel Aviv University, 69978 Tel Aviv, Israel*

(Received 30 March 2010; accepted 12 July 2010; published online 12 August 2010)

The triiodide ion $I_3^-$ exhibits a complex photodissociation behavior, the dynamics of which are not yet fully understood. As a first step toward determining the full potential energy surfaces of this species for subsequent simulations of its dissociation processes, we investigate the performance of different electronic structure methods [time-dependent density functional theory, complete active space perturbation theory to second order (CASPT2), Fock-space coupled cluster and multireference configuration interaction] in describing the ground and excited states of the triiodide ion along the symmetrical dissociation path. All methods apart from CASPT2 include scalar relativity and spin-orbit coupling in the orbital optimization, providing useful benchmark data for the more common two-step approaches in which spin-orbit coupling is introduced in the configuration interaction. Time-dependent density functional theory with the statistical averaging of model orbital potential functional is off the mark for this system. Another choice of functional may improve performance with respect to vertical excitation energies and spectroscopic constants, but all functionals are likely to face instability problems away from the equilibrium region. The Fock-space coupled cluster method was shown to perform clearly best in regions not too far from equilibrium but is plagued by convergence problems toward the dissociation limit due to intruder states. CASPT2 shows good performance at significantly lower computational cost, but is quite sensitive to symmetry breaking. We furthermore observe spikes in the CASPT2 potential curves away from equilibrium, signaling intruder state problems that we were unable to curb through the use of level shifts. Multireference configuration interaction is, in principle, a viable option, but its computational cost in the present case prohibits use other than for benchmarking purposes. © *2010 American Institute of Physics*. [doi:10.1063/1.3474571]

## I. INTRODUCTION

Recent years have seen extraordinary advances in experimental techniques to probe chemical processes such as reaction dynamics in very short time frames. A wide range of techniques based on pump-probe schemes,[1] where the species under consideration are set in a nonstationary state by one light source (pump) and monitored by another (probe), providing information regarding the dynamical behavior of the system. A particularly interesting field, where such fast techniques are very helpful, is the study of the dynamics of stable negative ions.[2] By means of photodetachment or photodissociation these ions can be used to provide access to unstable neutral species that are difficult to study directly.

Some ionic species, however, are important in their own right, apart from being used as precursors to other systems. A very well-known example is the triiodide ion ($I_3^-$). From a chemist's point of view,[3–6] this relatively simple system is very interesting as it is (a) a structural analog of a transition state in $S_N2$ reactions and (b) an example of hyperconjugation, with a three-center four-electron bond. The widely accepted bonding picture of $I_3^-$ is that of a $(\sigma, \pi)$ system arising from the combinations of 5p orbitals on the three iodine atoms. In $\Lambda S$-coupling the ground state configuration is accordingly $\sigma_u^2 \pi_u^4 \pi_g^4 \pi_u^{*4} \sigma_g^2$, where the LUMO is the antibonding $\sigma_u^*$. This bonding picture is modified by spin-orbit coupling (SOC) (cf. Fig. 1), in particular for the MOs of *gerade* symmetry that are nonbonding and retain the atomic SO-mixing. In *ungerade* symmetry spin-orbit coupling is quenched due to bond formation and the orbitals are therefore essentially spin-pure. We have chosen to retain the $\Lambda S$ notation of all orbitals for easier reference to other work, but note that the MOs denoted $\sigma_{1/2g}$ and $\pi_{1/2g}$ in Fig. 1 corre-

[a)]Author to whom correspondence should be addressed. Electronic mail: visscher@chem.vu.nl.

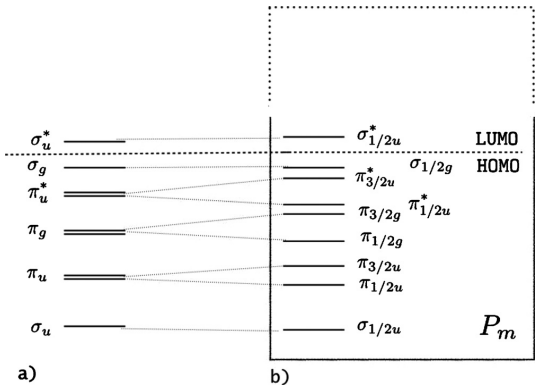



FIG. 1. Diagram of valence molecular orbitals of the triiodide ion, arising predominantly from the 5p manifold of the three iodine atoms: (a) spin-orbit free case and (b) spin-orbit case with the model $P_m$ space employed in the IHFSCC calculations indicated.

spond to linear combinations of $p_{1/2}$ and $p_{3/2}$ on the terminal atoms, respectively.

While early experimental information on the electronic structure of $I_3^-$ was obtained from UV spectroscopy,[7] more detailed information has been gained from photodissociation studies in more recent works. These studies have originally been performed in solution.[8–13] Questions arose as to whether it would be possible for the system to display a bent configuration at equilibrium, instead of the linear one proposed initially (the latter being the most accepted configuration). By contrast, experiments on the system in the gas-phase by Neumark and co-workers[14–16] and by Nagata and co-workers[17,18] strongly suggest that the ground state for this species is linear and centrosymmetric. Photofragment yield (PFY) spectra[15,17] reveal two photoabsorption bands peaking at 3.43–3.45 and 4.25–4.28 eV. They correspond to the C and D bands observed in solution,[7] and the peak values can be assigned as the vertical excitation energies from the ground to the excited $^3\Pi_u(0_u^+)$ and $^1\Sigma_u(0_u^+)$ states. Time of flight photofragment mass spectrometry by Nakanishi *et al.*[18] reveal a 40:60 ratio between three- and two-body dissociations both in the C and D bands. The three-body dissociation is dominated by charge-asymmetric dissociation that is with the negative charge on one of the terminal iodine atoms. The two-body dissociation in the C and D bands is dominated by the $(I^-+I_2)$ and $(I+I_2^-)$ channels, respectively. Although the same dissociation channels appear in the fast beam photofragment translation spectroscopy of Neumark and co-workers,[16] there are some discrepancies, notably in the calculated branching fractions for the two- and three-body dissociations as well as the $(I^-+I_2)/(I+I_2^-)$ ratio, which call for further investigation.

Compared to the wealth of experimental data on the triiodide species, relatively few theoretical studies have been performed. Kosloff and co-workers,[19] at about the time of the gas-phase experiments by Neumark,[14,15] presented an extensive study of the potential energy curves of the $I_3^-$, $I_2$, and $I_2^-$ species, which are thought to be involved in the photodissociation dynamics of $I_3^-$. Their results, obtained with a combination of multireference configuration interaction (MRCI) calculations and a diatomics-in-molecule (DIM) treatment where SOC effects were included, were the most accurate published at the time. However, the authors themselves considered the basis set used in the DIM treatment relatively small allowing for quantitative improvements on this early study. More recently, as a complement to their experimental work, Nakanishi *et al.*[18] performed comprehensive spin-orbit configuration interaction (SOCI) calculations which explored potential surfaces not only along the symmetric stretch of $I_3^-$ but also along the asymmetric stretch as well. Given the limited treatment of electron correlation and the still rather modest basis set size, also this theoretical investigation still leaves room for improvement.

Accordingly, the theoretical studies performed so far, although very helpful in understanding some aspects of the experimental data (such as the kind of states involved in the two experimentally observed absorption bands), are probably not sufficiently accurate for a direct comparison with experiment in terms of quantities associated with the dynamical behavior of the system, such as branching ratios. To enable this kind of analysis, it should be possible to calculate the whole of the potential energy surfaces accurately. This has motivated us to investigate the use of different methodologies for describing the electronic structure of the triiodide species to find the best candidate for calculation of a faithful representation of the potential energy surfaces that can be used in studies of the dynamics of the system.

Even though $I_3^-$ itself is a closed-shell species and can undergo two-body dissociation into two closed-shell species as well, it is important to account for SOC to achieve an accurate description of the full potential energy surfaces. Other possible channels involve dissociation into radicals, and furthermore SOC can have a significant impact on transition probabilities and intersystem crossings in the excitation/de-excitation processes. Apart from including SOC, theoretical methods should be able to describe the full potential energy surface (PES) and account for dynamical electron correlation at a reasonable cost, given the amount of calculations involved in determining a full PES.

As not all methods will fit this description we have chosen to compare expensive but accurate multireference coupled cluster (CC) and configuration interactions (CI) at selected points along the symmetric stretch of the molecule to methods that will allow full coverage of the ground and excited state potential energy surfaces. The wave function-based methods employed were (a) the intermediate Hamiltonian Fock-space coupled cluster method of Eliav and co-workers,[20–22] (b) the relativistic multireference CI method of Fleig and co-workers,[23,24] and (c) the spin-orbit complete active space (CAS) perturbation theory to second order (PT2) method (SO-CASPT2) of Roos and co-workers.[25,26] Apart from these, we decided to also explore the very economical time-dependent DFT (TDDFT) method[27] to provide a simple orbital picture for the excitation processes. These TDDFT calculations are of course not expected to yield better results than any of the methods above due to the well-known shortcomings of the current-day functionals (as, for

instance, in describing charge-transfer excitations[28,29] or that, within the adiabatic approximation, TDDFT can only describe single excitations[30–32]).

The paper is organized as follows: In Sec. II we present the details of the calculations performed with the different methodologies; in Sec. III we compare the performance of the different methods in calculating excitation energies of $I_3^-$, both at selected bond lengths and at the equilibrium geometries for the symmetric configuration. We also take the opportunity to discuss results for the triiodide radical ($I_3$), a species that was experimentally observed in photoionization studies involving $I_3^-$ in the gas-phase,[33,34] and that was also investigated theoretically by Kosloff and co-workers. Finally, in Sec. IV we assess the relative strengths and weaknesses of each method and provide concluding remarks.

## II. COMPUTATIONAL DETAILS

All calculations described were performed along the symmetric stretch coordinate, with I–I bond lengths $r_{I-I}$ in the range of [2.60:6.00] Å. In order to compare the different methodologies we have chosen two geometries in which to calculate the vertical excitation energies, apart from those obtained at the equilibrium structures for the different methods. These geometries are $r=2.84$ and 2.93 Å, corresponding to bond lengths in the vicinity of the equilibrium geometries for $I_3$ (from the MRCI calculations of Kosloff and co-workers[19]) and $I_3^-$ (equilibrium geometry reported from solid state studies[35]), respectively.

### A. Intermediate Hamiltonian Fock-space coupled cluster

Intermediate Hamiltonian Fock-space coupled cluster[20–22] (IHFSCC) calculations were performed with a development version of the DIRAC (Ref. 36) program. For describing the spectrum of $I_3^-$ with the IHFSCC method we have taken the anion as a starting point, and proceeded from the ground state through the $(1h,0p)$ and $(0h,1p)$ sectors in order to arrive at the $(1h,1p)$ sector and, therefore, at the excitation energies,

$$I_3^-(0h,0p) \rightarrow \{I_3(1h,0p), I_3^{2-}(0h,1p)\} \rightarrow I_3^{-(*)}(1h,1p). \tag{2.1}$$

For reasons of computational efficiency, the exact two-component Hamiltonian (X2C) scheme of Iliaš and Saue[37] was used. Two-electron spin same-orbit (SSO) and spin-other orbit (SOO) contributions were included via atomic mean-field integrals obtained with the AMFI[38,39] code.

The triple zeta basis sets of Dyall[40,41] were employed in all calculations. The starting large component 28s21p15d SCF set was kept uncontracted and augmented by 1s1p1d1f diffuse functions. It thus contains the 2s2p2d primitives recommended for polarization and valence correlation as well as the 2s2p2d primitives recommended for core-valence correlation. By further augmentation two different sets were defined: a valence correlation set ("aVTZ"), where one correlating f-type function was added to the augmented SCF set, yielding a (29s22p16d2f) basis set, and a core-valence correlating set ("aCVTZ"), which is a superset of aVTZ with additional 2f1g set of functions added.

In combination with these two sets different correlation spaces were employed: the first ($Q_1$) is used together with basis set aVTZ and includes the orbitals with orbital energies ($\varepsilon$) between $-1$ and 4 hartree. This means that in the occupied orbital space the $\sigma,\pi$ bonding system and three other $\sigma$-type orbitals arising from the 5s5p orbitals of iodine are present. The second correlation space ($Q_2$) is used together with basis set aCVTZ and includes the orbitals with orbital energies between $-3$ and 12 hartree. This corresponds to enlarging $Q_1$ by including the occupied 4d electrons of iodine apart from more virtual orbitals. The combinations aVTZ/$Q_1$ and aCVTZ/$Q_2$ will be referred to as IHFSCC(a) and IHFSCC(b), respectively.

A crucial ingredient of IHFSCC calculations to prevent convergence problems due to intruder states is the definition of the model ($P_m$) and intermediate ($P_I$) spaces that comprise the active space $P=P_m+P_I$.[21,22] After testing different spaces at the fixed geometries mentioned above, we found that convergence problems were generally avoided when the $P_m$ space contained 8 occupied ($1\sigma_g$, $1\pi_g$, $2\pi_u$, and $1\sigma_u$ orbitals) and 11 virtual orbitals ($2\sigma_g$, $1\pi_g$, $3\sigma_u$, and $2\pi_u$ orbitals) with a full $P$ space containing 11 occupied (5 of $g$ and 6 of $u$ parities) and 22 virtual (12 of $g$ and 10 of $u$ parities) orbitals (cf. Fig. 1). For bond lengths larger than 3.58 Å, however, it was not possible to obtain convergence for the $(1h,1p)$ sector for this partition.

### B. MRCI

MRCI calculations have been carried out with the relativistic double group CI program LUCIAREL,[23,24] which recently has been extended[42,43] for parallel computer applications and to allow for the computation of molecular properties.[44] In all of the calculations reported here the aCVTZ basis set and the X2C Hamiltonian including two-electron SSO and SOO corrections provided by the AMFI[38,39] code have been employed.

The molecular spinors for the CI calculations have been obtained by an average-of-configuration Hartree–Fock calculation, where the open shells were defined as the 8 occupied Kramers pairs as in the CC application above, and in addition the antibonding $\sigma_u$ orbital, corresponding to an active subspace with 16 electrons in 9 Kramers pairs. This type of Hartree–Fock wave function comprises a good starting point for relativistic MRCI studies since it provides a balanced description of ground and electronically excited states. The concept of general active spaces has been employed for constructing the CI expansion. In the present case, all Slater determinants with zero, one, and two particles in the external space (truncated at 3 hartrees) were included and all possible active space distributions were allowed for the remaining electrons. These active space distributions were defined by a CAS space corresponding to the above space used in the average-of-configuration Hartree–Fock calculation [16 electrons in 9 Kramers pairs, or 15 electrons in 9 Kramers pairs in the case of neutral $I_3$, "MRCI(a)"], and an additional space

including the iodine 5s orbitals where a restriction of up to two holes was imposed ["MRCI(b)"]. The resulting relativistic CI wave function describes the correlation of 16 (15 in the case of neutral $I_3$) or 22 electrons and consists of roughly $77 \times 10^6$ Slater determinants in the latter case.

### C. CASPT2

Calculations with the CASPT2 method were carried out with the MOLCAS code[45] (version 7.0), within the CASSCF/CASPT2/SO-RASSI approach. In this approach, scalar relativistic effects are included in the CASSCF (Ref. 25) and CASPT2 (Ref. 26) calculations via the Douglas–Kroll–Hess[46] Hamiltonian, and in a subsequent calculation the CASPT2 spin-free states are used by the RASSI program to set up a spin-orbit coupling Hamiltonian.[47] In this Hamiltonian the one- and two-electron spin-orbit integrals are calculated in a mean-field fashion via the AMFI[38,39] code. The basis set used in these calculations was the ANO-RCC (Ref. 48) set of TZP quality (7s6p4d2f1g). All calculations were carried out in $C_{2h}$ symmetry.

The CASSCF active space used consisted of 16 electrons in the nine orbitals arising from the 5p orbitals of the three iodine atoms. We have kept the core orbitals (i.e., up to and including the 3d orbitals) frozen in the CASSCF (and subsequent CASPT2) calculations. The CASSCF states that entered the multistate CASPT2 calculations have been obtained from state-averaged calculations over nine roots for $^1\Sigma$, six roots for $^1\Pi$, eight roots for $^3\Sigma$, and six roots for $^3\Pi$, respectively. In the CASPT2 calculations an IPEA shift of 0.25 a.u. was used.[49]

### D. TDDFT

TDDFT calculations[27] were carried out with the ADF (Ref. 50) code. Scalar relativistic and spin-orbit effects were included via the zeroth-order regular approximation.[51] In the TDDFT calculations noncollinear spin magnetization was used. We are also within the adiabatic local density approximation (LDA) approximation to TDDFT, therefore disregarding any time-dependence of the exchange-correlation kernel, while using only the derivatives of the LDA functional for computing the exchange-correlation contributions to the excitation energies.

We used the statistical averaging of model orbital potentials (SAOPs),[52] in combination with the TZ2P basis set.[53] It should be noted that in the ADF implementation SAOP is used to construct the Kohn–Sham potential during the SCF cycle, whereas the ground state energy is calculated with the PW91 (Ref. 54) energy expression. This makes the ground state properties close to those obtained with PW91, the differences arising from the modification of the potential. More details can be found in the ADF documentation (for instance, Ref. 55). Given the dependence of TDDFT on the functional used, we shall use the label TDDFT/SAOP.

## III. RESULTS AND DISCUSSION

In this section we discuss the results for the triiodide ion species obtained with the different methods. We begin by investigating the performance of the different methods in determining ground state spectroscopic constants, before addressing the electronic spectrum. Given the exploratory nature of this paper, we will restrict ourselves to selected structures along the symmetric stretch coordinate.

TABLE I. Ground state spectroscopic constants ($r_e$ in Å and $\omega_e$ in $cm^{-1}$) of $I_3^-$ calculated with the DFT, MRCI(a), CASPT2, and IHFSCC(a,b) methods.

| Method | $r_e$ | $\omega_e$ |
|---|---|---|
| DFT | 3.007 | 102 |
| MRCI(a) | 2.982 | 108 |
| CASPT2 | 2.888 | 119 |
| IHFSCC(a) | 2.971 | 112 |
| IHFSCC(b) | 2.946 | 114 |
| MRCI (Ref. 19, spin-free) | 2.930 | 114 |
| DIM+SO (Ref. 19) | 2.966 | 95 |
| Expt. (Ref. 35) | 2.93 | $112 \pm 1$ |

When comparing methodologies, it should be kept in mind that each method, with the exception of TDDFT/SAOP, can be tuned through the selection of active electrons and orbitals. As reference we take IHFSCC(b), which we consider the most accurate methodology employed in this work.

We will focus mainly on vertical excitations calculated at selected geometries, but will also address "adiabatic" excitations for this particular cut of the surface, as a way to gain insight on how the different methods represent the overall shape of the surfaces. Finally, we will take a more detailed look the $0_u$ states, in particular, the absorbing $0_u^+$ states, comparing the energetics and the excitation picture, in terms of the respective molecular orbitals for the different methods, before discussing the triiodide radical.

### A. Ground state spectroscopic constants for $I_3^-$

The ground state spectroscopic constants of $I_3^-$ obtained here are shown in Table I. We take the experimental bond length[35] in the solid state ($r_e$=2.93 Å) as a measure since to the best of our knowledge, there are no experimental bond length determinations for the triiodide ion in gas-phase.

With respect to this reference value CASPT2 shows a slight (about 0.04 Å) underestimation, whereas IHFSCC(a) shows a similar overestimation. Better agreement, with a deviation of less than 0.02 Å, is obtained with IHFSCC(b), where the 4d shell is included in the occupied space and the virtual space is truncated at a higher energy. The effect of electron correlation can be studied in a more systematic manner at the CASPT2 level. Our default scheme is a CAS consisting of the 5p manifold and in addition correlating 4s4p4d5s at the PT2 level, giving a bond length of 2.888 Å. Freezing 4s4p, which corresponds to the default CASPT2 correlation scheme in MOLCAS 7.2, increases the bond length to 2.914 Å. Freezing 4d gives an even more significant bond length extension to 2.966 Å, whereas in addition freezing 5s shortens the bond slightly to 2.958 Å. These results suggest that the inclusion of the 4s4p shells in the occupied space of the IHFSCC calculations will bring the bond length in even closer agreement with experiment and thus implies that environmental effects on the bond length are small. We were unable to optimize the $I_3^-$ bond distance at the MRCI(b) level

TABLE II. Comparison of vertical excitation energies $T_v$ (in eV) obtained with the different methods [TDDFT/SAOP, MRCI(a,b), CASPT2, and IHFSC-C(a,b)] for $I_3^-$ at $r_1=r_2=2.93$ Å. States of the same symmetry as those for the optically active excited states are shown in boldface. Statistical measures of the error compared to IHFSCC(b) are also shown (see text for discussion).

| | TDDFT/SAOP | | MRCI(a) | | MRCI(b) | | CASPT2 | | IHFSCC(a) | | IHFSCC(b) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| State | Ω | $T_v$ | Ω | $T_v$ | Ω | $T_v$ | Ω | $T_v$ | Ω | $T_v$ | Ω | $T_v$ |
| 1 | $2_g$ | 1.92 | $2_g$ | 2.32 | $2_g$ | 2.30 | $2_g$ | 2.24 | $2_g$ | 2.10 | $2_g$ | 2.05 |
| 2 | $1_g$ | 2.04 | $1_g$ | 2.44 | $0_u^-$ | 2.40 | $1_g$ | 2.32 | $1_g$ | 2.23 | $1_g$ | 2.18 |
| 3 | $0_u^-$ | 2.42 | $0_u^-$ | 2.47 | $1_u$ | 2.41 | $0_u^-$ | 2.47 | $0_u^-$ | 2.26 | $0_u^-$ | 2.20 |
| 4 | $1_u$ | 2.43 | $1_u$ | 2.48 | $1_g$ | 2.41 | $1_u$ | 2.47 | $1_u$ | 2.27 | $1_u$ | 2.20 |
| 5 | $0_g^-$ | 2.50 | $0_g^-$ | 2.94 | $0_g^-$ | 2.91 | $0_g^-$ | 2.76 | $0_g^-$ | 2.68 | $0_g^-$ | 2.64 |
| 6 | $0_g^+$ | 2.56 | $0_g^+$ | 2.98 | $0_g^+$ | 2.95 | $0_g^+$ | 2.82 | $0_g^+$ | 2.73 | $0_g^+$ | 2.69 |
| 7 | $1_g$ | 2.70 | $1_g$ | 3.13 | $1_g$ | 3.08 | $1_g$ | 2.85 | $1_g$ | 2.90 | $1_g$ | 2.86 |
| 8 | $2_u$ | 2.61 | $2_u$ | 3.25 | $2_u$ | 3.21 | $2_u$ | 3.10 | $2_u$ | 3.22 | $2_u$ | 3.17 |
| 9 | $1_u$ | 2.72 | $1_u$ | 3.30 | $1_u$ | 3.25 | $1_u$ | 3.11 | $1_u$ | 3.30 | $1_u$ | 3.24 |
| 10 | $\mathbf{0_u^+}$ | **3.14** | $\mathbf{0_u^+}$ | **3.71** | $\mathbf{0_u^+}$ | **3.66** | $\mathbf{0_u^+}$ | **3.52** | $\mathbf{0_u^+}$ | **3.52** | $\mathbf{0_u^+}$ | **3.51** |
| 11 | $2_g$ | 3.50 | $0_u^-$ | 4.04 | $0_u^-$ | 3.98 | $0_u^-$ | 3.79 | $0_u^-$ | 3.95 | $2_g$ | 3.88 |
| 12 | $0_u^-$ | 3.42 | $1_u$ | 4.09 | $1_u$ | 4.02 | $1_u$ | 3.80 | $2_g$ | 3.96 | $0_u^-$ | 3.91 |
| 13 | $1_g$ | 3.63 | $2_g$ | 4.19 | $2_g$ | 4.16 | $2_g$ | 3.98 | $1_u$ | 4.03 | $1_g$ | 4.00 |
| 14 | $1_u$ | 3.56 | $1_g$ | 4.29 | $1_g$ | 4.25 | $1_g$ | 4.06 | $1_g$ | 4.07 | $1_u$ | 4.00 |
| 15 | $0_g^-$ | 4.10 | $0_g^-$ | 4.81 | $\mathbf{0_u^+}$ | **4.75** | $\mathbf{0_u^+}$ | **4.51** | $\mathbf{0_u^+}$ | **4.33** | $\mathbf{0_u^+}$ | **4.33** |
| 16 | $0_g^+$ | 4.11 | $0_g^+$ | 4.82 | $0_g^-$ | 4.77 | $0_g^-$ | 4.51 | $0_g^-$ | 4.54 | $0_g^-$ | 4.48 |
| 17 | $1_g$ | 4.34 | $\mathbf{0_u^+}$ | **4.83** | $0_g^+$ | 4.78 | $0_g^+$ | 4.53 | $0_g^+$ | 4.54 | $0_g^+$ | 4.48 |
| 18 | $\mathbf{0_u^+}$ | **4.46** | $1_g$ | 4.96 | $1_g$ | 4.91 | $1_g$ | 4.60 | $2_u$ | 4.7 | $1_g$ | 4.68 |
| $\bar{\Delta}$ | | −0.24 | | 0.25 | | 0.20 | | 0.05 | | 0.05 | | |
| $\Delta_{\text{std}}$ | | 0.24 | | 0.11 | | 0.11 | | 0.15 | | 0.02 | | |
| $\bar{\Delta}_{\text{abs}}$ | | 0.31 | | 0.25 | | 0.21 | | 0.12 | | 0.05 | | |
| $\Delta_{\text{max}}$ | | 0.56 | | 0.50 | | 0.41 | | 0.27 | | 0.08 | | |

due to computational constraints. Based on the trends observed with CASPT2, one would expect that MRCI(b), which correlates the 5s5p, would *overestimate* the bond length further compared to MRCI(a). This is surprising since Vala *et al.*[19] with the same active space and basis set quality reproduced the experimental bond length in their spin-orbit free MRCI calculations. One possible source of error in the present MRCI calculations is the truncation of the virtual space, but from previous experience we deem it unlikely that this truncation affects the bond length by 0.05 Å. We instead believe that the results of Vala *et al.*[19] are fortuitous since our studies clearly show that correlation of the 4s4p4d manifold has a significant effect on the spectroscopic constants. Indeed, if we subtract the 0.07 Å gained by correlating 4s4p4d5s at the CASPT2 level from our MRCI(a) result we come quite close to the experimental value.

It is interesting to note that DFT, where in principle all orbitals are correlated, overestimates $r_e$ by about 0.07 Å. Clearly, the SAOP potential was not derived with the aim of providing accurate ground state structures. On the other hand, other generalized gradient approximation (GGA) functionals do not perform better, e.g., PW91 (2.996 Å), PBE (Ref. 56) (2.996 Å), M06-L (Ref. 57) (3.065 Å), and BP86 (3.007 Å). The latter value contrasts significantly with the bond length of 3.14 Å reported by Landrum *et al.*,[5] but they employed a TZP basis, thus not including the diffuse functions of the TZ2P basis as we do. Hybrid functionals do perform better, e.g., PBE0 (Ref. 58) (2.946 Å) and M06–2X (Ref. 59) (2.925 Å), whereas LDA overbinds, giving (2.913 Å). We note in passing that the effect of spin-orbit coupling on the bond lengths is on the order of 0.02 Å for the cited functionals, again an indication that the good agreement of the spin-orbit free MRCI calculation of Vala *et al.*[19] with the experimental bond length is likely to be fortuitous.

A similar picture is seen for the harmonic frequencies. Discrepancies between MRCI, CASPT2, IHFSCC, and both experiment[60] and the spin-free MRCI calculations of Kosloff are of the order of a few $cm^{-1}$, whereas DFT underestimates the frequency by 10 $cm^{-1}$. The spin-orbit numbers of Kosloff, on the other hand, show rather large discrepancies (17 $cm^{-1}$), which may be due to artifacts arising from their diatomics-in-molecule treatment.

### B. Benchmark calculations of the electronic spectra of $I_3^-$

In the comparison of calculated spectra, we consider all methods both at the experimental bond length ($r=2.93$ Å) and at their own optimal ground state equilibrium geometries (shown in Table I). The corresponding spectra can be found in Tables II and III, respectively.

As already mentioned, we discuss the performance of the other methods relative to IHFSCC(b). This method is accurate for the first 18 excited states, which go up to about 4.5–5.0 eV, but becomes less trustworthy for higher energies due to the increasing importance of double excitations in these states. Double excitations are readily captured by MRCI and CASPT2 but neither by TDDFT/SAOP nor by the IHFSCC calculations on the $(1h,1p)$ sector of Fock-space.

TABLE III. Vertical ($T_v$) excitation energies (in eV) calculated with the TDDFT/SAOP, MRCI(a), CASPT2, and IHFSCC(a,b) methods for $I_3^-$ calculated at the optimum bond length for each method. States of the same symmetry as those for the optically active excited states are shown in boldface. Statistical measures of the error compared to IHFSCC(b) are also shown (see text for discussion).

| | TDDFT/SAOP | | MRCI(a) | | CASPT2 | | IHFSCC(a) | | IHFSCC(b) | |
|---|---|---|---|---|---|---|---|---|---|---|
| State | $\Omega$ | $T_v$ | $\Omega$ | $T_v$ | $\Omega$ | $T_v$ | $\Omega$ | $T_v$ | $\Omega$ | $T_v$ |
| 1 | $2_g$ | 1.67 | $2_g$ | 2.15 | $2_g$ | 2.38 | $2_g$ | 1.98 | $2_g$ | 2.00 |
| 2 | $1_g$ | 1.78 | $1_g$ | 2.27 | $1_g$ | 2.45 | $1_g$ | 2.10 | $1_g$ | 2.13 |
| 3 | $0_u^-$ | 2.13 | $0_u^-$ | 2.28 | $1_u$ | 2.63 | $0_u^-$ | 2.11 | $0_u^-$ | 2.13 |
| 4 | $1_u$ | 2.14 | $1_u$ | 2.28 | $0_u^-$ | 2.63 | $1_u$ | 2.12 | $1_u$ | 2.14 |
| 5 | $0_g^-$ | 2.24 | $0_g^-$ | 2.77 | $0_g^-$ | 2.90 | $0_g^-$ | 2.55 | $0_g^-$ | 2.58 |
| 6 | $2_u$ | 2.26 | $0_g^+$ | 2.81 | $0_g^+$ | 2.95 | $0_g^+$ | 2.60 | $0_g^+$ | 2.64 |
| 7 | $0_g^+$ | 2.31 | $1_g$ | 2.94 | $1_g$ | 3.00 | $1_g$ | 2.77 | $1_g$ | 2.81 |
| 8 | $1_u$ | 2.38 | $2_u$ | 3.00 | $2_u$ | 3.31 | $2_u$ | 3.04 | $2_u$ | 3.10 |
| 9 | $1_g$ | 2.44 | $1_u$ | 3.05 | $1_u$ | 3.32 | $1_u$ | 3.11 | $1_u$ | 3.17 |
| 10 | $\mathbf{0_u^+}$ | **2.82** | $\mathbf{0_u^+}$ | **3.47** | $\mathbf{0_u^+}$ | **3.72** | $\mathbf{0_u^+}$ | **3.44** | $\mathbf{0_u^+}$ | **3.36** |
| 11 | $2_g$ | 3.07 | $0_u^-$ | 3.79 | $0_u^-$ | 3.99 | $2_g$ | 3.73 | $2_g$ | 3.79 |
| 12 | $0_u^-$ | 3.09 | $1_u$ | 3.84 | $1_u$ | 4.00 | $0_u^-$ | 3.77 | $0_u^-$ | 3.84 |
| 13 | $1_g$ | 3.20 | $2_g$ | 3.89 | $2_g$ | 4.23 | $1_g$ | 3.84 | $1_g$ | 3.90 |
| 14 | $1_u$ | 3.22 | $1_g$ | 3.98 | $1_g$ | 4.31 | $1_u$ | 3.85 | $1_u$ | 3.93 |
| 15 | $0_g^-$ | 3.66 | $0_g^-$ | 4.52 | $\mathbf{0_u^+}$ | **4.67** | $\mathbf{0_u^+}$ | **4.17** | $\mathbf{0_u^+}$ | **4.27** |
| 16 | $0_g^+$ | 3.68 | $0_g^+$ | 4.52 | $0_g^-$ | 4.75 | $0_g^-$ | 4.31 | $0_g^-$ | 4.38 |
| 17 | $1_g$ | 3.91 | $\mathbf{0_u^+}$ | **4.61** | $0_g^+$ | 4.78 | $0_g^+$ | 4.32 | $0_g^+$ | 4.39 |
| 18 | $\mathbf{0_u^+}$ | **4.29** | $1_g$ | 4.65 | $1_g$ | 4.85 | $1_g$ | 4.50 | $1_g$ | 4.59 |
| $\bar{\Delta}$ | | −0.49 | | 0.09 | | 0.32 | | −0.05 | | |
| $\Delta_{\text{std}}$ | | 0.29 | | 0.12 | | 0.12 | | 0.04 | | |
| $\bar{\Delta}_{\text{abs}}$ | | 0.49 | | 0.13 | | 0.32 | | 0.06 | | |
| $\Delta_{\text{max}}$ | | 0.84 | | 0.34 | | 0.50 | | 0.10 | | |

In the statistical analysis of the errors for each method we therefore include only the first 18 states that are dominated by single excitations. In this analysis that is found at the bottom of Tables II and III, we report the mean error $\bar{\Delta}$ and its standard deviation $\Delta_{\text{std}}$, as well as the mean absolute error $\bar{\Delta}_{\text{abs}}$ and the maximum absolute error $\Delta_{\text{max}}$ relative to the IHFSCC(b) results.

### 1. Vertical excitations

Inspecting Table II we immediately see some general trends: CASPT2 and MRCI tend to overestimate the IHFSCC excitation energies, whereas TDDFT/SAOP shows the lowest excitation energies among all methods considered. One can also see that the degree of electron correlation introduced by the different methods affects distinct regions of the spectrum differently. For the lowest ten excited states, the ordering of states is consistent with most methods whereas for higher states, where the density of states is higher, small variations in the correlation treatment result in significant reorderings. The IHFSCC(a) scheme yields only small deviations relative to the larger IHFSCC(b) calculation. This is quantified by the small mean errors (signed and absolute) and a small standard deviation (0.02 eV), suggesting that core-valence correlation does not play a prominent role in describing the transitions to low-lying excited states. Such effects can be studied with more ease at the CASPT2 level. We find that freezing the 4s4p shells at the PT2 level has no effect on the vertical excitation energies, whereas freezing the 4d shell brings about an upward shift of about 0.12 eV, somewhat larger than what is observed at the IHFSCC level. Freezing in addition the 5s shell leads to a mean upward shift of around 0.21 eV compared to our default CASPT2 scheme. With these observations in mind one can readily understand why CASPT2 reproduces better the IHFSCC(b) reference than the MRCI calculations since the latter have much more limited active spaces. TDDFT/SAOP underestimates ($\bar{\Delta}$=0.24 eV) the excitation energies, but perhaps more worrisome is that $\Delta_{\text{std}}$ is rather large (0.24 eV), indicating large nonsystematic errors.

While it is beyond the scope of this paper to present a detailed analysis of the difference between IHFSCC and other methods, it is nevertheless instructive to consider comparisons of excitation energies calculated by IHFSCC and the linear response coupled cluster (LRCC) method done by various authors.[61–65] From those studies it becomes clear that LRCC and IHFSCC may yield significantly different excitation energies. Some evidence from recent studies for different molecular systems performed by some of us[66] as well as by other authors[67–70] seems to point to a systematic upward shift in LRCC excitation energies compared to IHFSCC ones, similar to what is observed here when comparing MRCI and IHFSCC. This could be consistent with the different parametrization used in describing the excitations from the ground to the excited states (linear in MRCI and LRCC, and exponential for IHFSCC). It is less straightforward to rationalize the discrepancies between CASPT2 and IHFSCC as these methods describe the ground state with a rather different wave function. Based on previous experience[66,71–73] we expect CASPT2 results for low-lying

excited states to deviate up to a few tenths of an eV from IHFSCC, but is more difficult to predict either an increase or a decrease of the excitation energies.

In Table III we present vertical excitation energies at the ground state equilibrium geometry optimized for each method. With a spread of 0.12 Å between the optimized bond lengths, the discrepancies with respect to the reference IHFSCC(b) data increase. The effect is accentuated by the fact that the vertical excitation energies probe the repulsive wall of the excited states, as will become clear in Sec. III B 2. We first note that the vertical excitation energies obtained with the IHFSCC(b) scheme to the optically active states—3.36 and 4.27 eV—are close to the experimentally observed peaks in the PFY spectra [3.43 and 4.25 eV (Ref. 18)]. The agreement is also quite acceptable with the IHFSCC(a) scheme, which again indicates that core-valence correlation (correlating deeper than the 5s shell) can possibly be disregarded when constructing the full potential energy surfaces. The corresponding CASPT2 results—3.72 and 4.67 eV—are close to the SOCI values reported by Nakanishi *et al.*[18] (3.79 and 4.70 eV). MRCI(a) is reasonably close to the lower excitation energy, but severely overestimates the second one, thus providing a peak separation energy of 1.14 eV which is far from the experimentally observed 0.82 eV. In fact, it is only the IHFSCC method that is able to reproduce the experimental peak positions and separation; all other methods explored in this paper overestimate the splitting significantly with TDDFT/SAOP being simply off the mark for these excitations.

The statistical analysis at these selected geometries indicates that both MRCI and CASPT2 are able to provide a balanced treatment of the ground and excited state surfaces. A problem with MRCI(a), currently the only feasible CI approach to cover the whole surface, is the fact that the optically active states (indicated in boldface in both Tables II and III) that are well described with IHFSCC(b) exhibit relatively large errors when treated with MRCI(a). An important remark concerning CASPT2 is that our conclusions are based on calculations carried out in $C_{2h}$ symmetry with the rotation axis aligned with the molecular axis. Since we are presently only exploring the symmetric stretch one would rather be inclined to use the highest possible point group symmetry available for a centrosymmetric configuration, which is $D_{2h}$ with the MOLCAS code. However, the introduction of a vertical mirror plane places $\pi_x$ and $\pi_y$ orbitals in different irreducible representations and introduces a symmetry breaking to which CASPT2 is extremely sensitive. Vertical excitation energies at bond length 2.93 Å calculated in $D_{2h}$ and $C_{2v}$ symmetries give upper and lower bounds, respectively, for the corresponding $C_{2h}$ values within a span approaching 0.1 eV. The choice of symmetry can therefore seriously affect conclusions regarding the performance of the method. On the other hand, restricting symmetry to $C_{2h}$ means that the important distinction between $0^+$ and $0^-$ states is lost. The use of supersymmetry is hardly an option in the scan of the potential surfaces, so we have in the present work rather relied on scripts that analyze the CASSCF orbitals to extract this symmetry information. Having to work at reduced symmetry also implies that orbitals are optimized and thus averaged for a larger number of states.

We also note that the two other experimentally observed[7] bands—the A band (at 2.19 eV in $CH_2Cl_2$ and 2.16 eV in $Et_4NI_3$ crystal), assigned to a $^3\Pi_u$, and the B band (at 2.82 eV in $CH_2Cl_2$ and 2.71 eV in $Et_4NI_3$ crystal), assigned to a $^3\Sigma_u^+$ state—can indeed be correlated with the first and second $1_u$ states calculated for the vertical excitations with IHFSCC(b) (2.13 and 3.17 eV), IHFSCC(a) (2.11 and 3.11 eV), MRCI(a) (2.28 eV and 3.05 eV), and TDDFT/SAOP (2.14 and 2.38 eV), while agreement with SO-CASPT2 is less good (2.63 and 3.32 eV). While the agreement for the maxima of band A is rather good, for band B it is much less so. This could be due to the fact that this band is rather weak and only visible as a shoulder on the C band in the spectra reported by Gabes and Stufkens,[7] making the exact position dependent on the fitting procedure employed to establish the band maximum.

### 2. Explorations of the dissociative region

The discussion so far has dealt with geometries at and close to the ground state equilibrium structures. However, in order to be useful in modeling the dissociation process of $I_3^-$ the potential surfaces of the excited states far from the equilibrium region have to be properly described as well. In order to probe the relative accuracy of the methods, we have chosen to investigate displacements along the symmetric stretch coordinate, along which photodissociation via the C and D bands generally initiates. This represents a one-dimensional cut through the full three-dimensional surface enabling us to define local minima (and harmonic frequencies) for the different excited states in this restricted geometry. One should thereby keep in mind, however, that such extrema do not necessarily correspond to the true spectroscopic constants for the corresponding states as we did not investigate the curvature of the surface in directions orthogonal to the symmetric stretch coordinate.

A further restriction in the current study is the presence of intruder states that made it impossible to converge the IHFSCC calculations at internuclear distances larger than 3.6 Å. This means that we can only compare with the IHFSCC(b) reference at relatively short distances. Similarly, we could not obtain TDDFT/SAOP results beyond about 3.9 Å due to triplet instabilities. In the CASPT2 calculations we observe downward spikes in the potential curves, visible in Fig. 2, appear around 3.5 Å, signaling a weak intruder state problem. These features appear already in the spin-orbit free CASPT2 calculations, whereas the CASSCF calculations produce smooth potential curves. We were unable to curb these features, applying real[74] and imaginary[75] level shifts of up to 0.20 $E_h$. On the other hand, outside the spikes the potential curves appear smooth and so a pragmatic approach to fitting the potential surfaces would be to remove points for which the weight of the reference CASSCF state drops below a selected value since this characterizes the appearance of the spikes. A more satisfying solution would be to increase the active space, but this easily brings the computational cost out of practical reach.

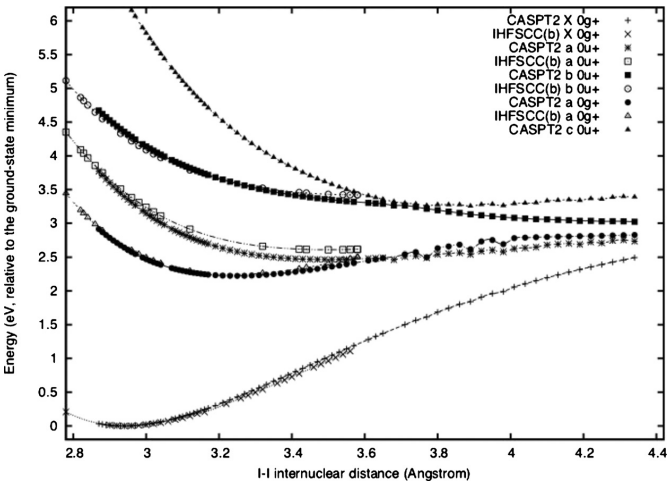



FIG. 2. Potential energy curves along the symmetric stretch coordinate for the ground state (GS, $0_g^+$) and first excited $0_g^+$ state, and lower $0_u^+$ states (the first two corresponding to the optically active states) from CASPT2 and IHFSCC(b) calculations. All potentials were scaled so that the ground state minimum corresponds to the origin; the CASPT2 excited states were further shifted (a) downward by 0.31 eV and (b) toward larger internuclear distances by 0.05 Å, to compensate the systematic errors compared to IHFSCC(b).

Potential curves along the symmetric stretch for the lower two $0_g^+$ and three $0_u^+$ states, calculated at the CASPT2 and IHFSCC(b) level, are given in Fig. 2. The crossing of the first excited $0_g^+$ and $0_u^+$ states around 3.7 Å translate into a conical intersection in a 2-dimensional (2D) plot involving also the antisymmetric stretch and plays an important role in the photodissociation via the C band, as discussed by Nakanishi *et al.*[18] These authors indicate a number of avoided crossings and conical intersections appearing in their SOCI calculations, such as an avoided crossing between the second and third $0_u^+$ state, clearly visible around 3.7 Å in Fig. 2, which renders the second $0_u^+$ state, associated with the D band, dissociative. Due to these features in the potential curves we have restricted the calculation of spectroscopic constants and adiabatic excitation energies to the ten lowest excited states of $I_3^-$, presented in Table IV.

The lower excited states are dominated by single excitations to the antibonding $\sigma^*_{1/2u}$ LUMO and bond lengths are therefore systematically considerably longer than observed for the ground state. This also implies, as already mentioned, that vertical excitations probe the repulsive wall of the excited states. We note, for instance, that the vertical and adiabatic excitation energies to the first excited $0_u^+$, calculated at the IHFSCC(b) level, are 3.36 and 2.61 eV, respectively, a difference of 0.75 eV, meaning that the photodissociation process starts off with a significant acceleration of the vibrational wave packet.

With respect to the spectroscopic constants given in Table IV, we observe excellent agreement between IHFSCC(a) and IHFSCC(b), again suggesting that correlation of the 4d shell is not crucial for a good description of the excited states potential surfaces. CASPT2 tends to consistently underestimate the bond lengths compared to IHFSCC(b) by about 0.07 Å, in line with the difference of 0.058 Å for the

TABLE IV. Spectroscopic constants ($r_e$ in Å and $\omega_e$ in $cm^{-1}$) and adiabatic ($T_e$) excitation energies (in eV) for the ten lowest excited states of $I_3^-$, calculated with the TDDFT/SAOP, MRCI(a), CASPT2, IHFSCC(a), and IHFSCC(b) methods. The lowest, optically active $0_u^+$ state is shown in boldface. Statistical measures of the error compared to IHFSCC(b) are also shown (see text for discussion). Values in parenthesis have reduced accuracy due to triplet instabilities (TDDFT) or lack of convergence (IHFSCC) and have not been included in the statistical analysis.

| | TDDFT/SAOP | | | | MRCI(a) | | | | CASPT2 | | | | IHFSCC(a) | | | | IHFSCC(b) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| State | Ω | $r_e$ | $\omega_e$ | $T_e$ | Ω | $r_e$ | $\omega_e$ | $T_e$ | Ω | $r_e$ | $\omega_e$ | $T_e$ | Ω | $r_e$ | $\omega_e$ | $T_e$ | Ω | $r_e$ | $\omega_e$ | $T_e$ |
| 1 | $2_g$ | 3.443 | 55 | 1.19 | $0_u^-$ | 3.501 | 56 | 1.53 | $2_g$ | 3.188 | 80 | 1.97 | $0_u^-$ | 3.439 | 60 | 1.45 | $0_u^-$ | 3.417 | 60 | 1.45 |
| 2 | $1_g$ | 3.462 | 52 | 1.29 | $1_u$ | 3.503 | 56 | 1.53 | $1_u$ | 3.318 | 71 | 1.98 | $1_u$ | 3.436 | 61 | 1.46 | $1_u$ | 3.414 | 61 | 1.46 |
| 3 | $0_u^-$ | (3.719) | (22) | (1.38) | $2_g$ | 3.383 | 61 | 1.64 | $0_u^-$ | 3.315 | 72 | 1.99 | $2_g$ | 3.270 | 81 | 1.59 | $2_g$ | 3.245 | 83 | 1.59 |
| 4 | $1_u$ | (3.696) | (29) | (1.38) | $1_g$ | 3.421 | 56 | 1.71 | $1_g$ | 3.208 | 76 | 2.02 | $1_g$ | 3.282 | 78 | 1.70 | $1_g$ | 3.255 | 80 | 1.71 |
| 5 | $2_u$ | 3.616 | 46 | 1.41 | $2_u$ | 3.644 | 47 | 1.89 | $1_u$ | 3.427 | 59 | 2.33 | $0_g^-$ | 3.383 | 52 | 2.05 | $0_g^-$ | 3.358 | 53 | 2.07 |
| 6 | $1_u$ | 3.669 | 39 | 1.52 | $1_u$ | 3.657 | 47 | 1.89 | $2_u$ | 3.417 | 60 | 2.35 | $1_g$ | ⋯ | ⋯ | ⋯ | $1_g$ | (3.514) | (29) | (2.20) |
| 7 | $0_g^+$ | 3.388 | 63 | 1.68 | $0_g^-$ | 3.478 | 48 | 2.16 | $0_g^-$ | 3.222 | 73 | 2.44 | $2_u$ | 3.475 | 63 | 2.18 | $2_u$ | 3.446 | 64 | 2.22 |
| 8 | $1_g$ | (3.877) | (54) | (1.81) | $1_g$ | 3.562 | 43 | 2.22 | $1_g$ | 3.260 | 68 | 2.47 | $0_g^+$ | 3.264 | 82 | 2.22 | $0_g^+$ | 3.240 | 84 | 2.24 |
| 9 | $0_g^-$ | 3.554 | 38 | 1.88 | $0_g^+$ | 3.388 | 62 | 2.28 | $0_g^+$ | 3.198 | 78 | 2.53 | $1_u$ | 3.492 | 62 | 2.21 | $1_u$ | 3.462 | 63 | 2.26 |
| 10 | $\mathbf{0_u^+}$ | **3.513** | **57** | **2.10** | $\mathbf{0_u^+}$ | **3.731** | **34** | **2.36** | $\mathbf{0_u^+}$ | **3.445** | **53** | **2.78** | $\mathbf{0_u^+}$ | ⋯ | ⋯ | ⋯ | $\mathbf{0_u^+}$ | **3.498** | **50** | **2.61** |
| $\bar{\Delta}$ | | 0.163 | −18 | −0.52 | | 0.152 | −15 | −0.07 | | −0.066 | 3 | 0.31 | | 0.025 | −1 | −0.02 | | | | |
| $\Delta_{\text{std}}$ | | 0.106 | 20 | 0.18 | | 0.051 | 8 | 0.19 | | 0.036 | 9 | 0.16 | | 0.003 | 1 | 0.02 | | | | |
| $\bar{\Delta}_{\text{abs}}$ | | 0.163 | 23 | 0.52 | | 0.152 | 15 | 0.14 | | 0.066 | 7 | 0.31 | | 0.025 | 1 | 0.02 | | | | |
| $\Delta_{\text{max}}$ | | 0.314 | −46 | −0.81 | | 0.233 | −24 | −0.37 | | −0.136 | 20 | 0.54 | | 0.030 | −2 | −0.05 | | | | |

TABLE V. Character of $0_u^+$ states of $I_3^-$ for TDDFT/SAOP, IHFSCC(b), CASPT2, and MRCI(a) at the respective ground state equilibrium geometries. These are given in terms of contributions from one-electron (for TDDFT, IHFSCC, and MRCI, where spin-orbit coupling is included at the SCF level) or many-electron states (for CASPT2). The oscillator strengths f associated to the transition from the ground state to these excited states are also shown (where available) along with the vertical excitation energy. The experimental absorption maxima are found at (a) 3.43 and 4.25 eV from gas-phase photofragment yield spectra (Ref. 18), (b) 3.41 and 4.22 eV in $CH_2Cl_2$ (solution) (Ref. 7), and (c) 3.38 and 4.28 eV in $Et_4NI_3$ (solid state) (Ref. 7).

| Method | State | Excited state composition | $T_v$ (eV) | f |
|---|---|---|---|---|
| TDDFT/SAOP | $0_u^+$ | $51\%\sigma_{1/2g}\rightarrow\sigma^*_{1/2u}+49\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}$ | 2.82 | 0.09 |
| | $0_u^+$ | $45\%\sigma_{1/2g}\rightarrow\sigma^*_{1/2u}+50\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}$ | 4.29 | 1.54 |
| MRCI | $0_u^+$ | $46\%_{1/2g}\rightarrow\sigma^*_{1/2u}+30\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}+7\%(\pi^1_{1/2g}\pi^1_{1/2u}\rightarrow\sigma^{*2}_{1/2u})$ | 3.47 | 0.37 |
| | $0_u^+$ | $29\%\sigma_{1/2g}\rightarrow\sigma^*_{1/2u}+47\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}+6\%(\sigma^1_{1/2g}\sigma^1_{1/2u}\rightarrow\sigma^{*2}_{1/2u})$ | 4.61 | 1.83 |
| IHFSCC | $0_u^+$ | $40\%\sigma_{1/2g}\rightarrow\sigma^*_{1/2u}$(LUMO)$+19\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}$(LUMO) $+22\%\sigma_{1/2g}\rightarrow\sigma^*_{1/2u}+15\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}$ | 3.36 | n/a[a] |
| | $0_u^+$ | $26\%\sigma_{1/2g}\rightarrow\sigma^*_{1/2u}$(LUMO)$+36\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}$(LUMO)$+8\%\sigma_{1/2g}$ $\rightarrow\sigma^*_{1/2u}+24\%\pi_{1/2g}\rightarrow\sigma^*_{1/2u}$ | 4.27 | n/a[a] |
| CASPT2 | $0_u^+$ | $14\%\ ^1\Sigma_u^+(98\%\sigma_g\rightarrow\sigma_u^*)+85\%\ ^3\Pi_u(91\%\pi_g\rightarrow\sigma_u^*)$ | 3.72 | 0.45 |
| | $0_u^+$ | $84\%\ ^1\Sigma_u^+(98\%\sigma_g\rightarrow\sigma_u^*)+15\%\ ^3\Pi_u(91\%\pi_g\rightarrow\sigma_u^*)$ | 4.67 | 2.37 |

[a]Not available.

ground state. On the other hand, MRCI(a) significantly overestimates the excited state bond lengths, approaching the errors of TDDFT/SAOP ($\bar{\Delta}\simeq 0.16$ Å), although with lower standard deviation. The same tendency is observed for harmonic frequencies, where CASPT2 on the other hand shows rather good performance.

It is interesting then to note that MRCI(a) outperforms CASPT2 for adiabatic excitation energies, whereas TDDFT/SAOP severely underestimates them. CASPT2 has, on the other hand, a reasonably small and consistent standard deviation of around 0.16 eV in excitation energies. This systematic nature of the errors can therefore perhaps be exploited to further bring the results close to IHFSCC(b) by applying global shifts to the CASPT2 potential energy surfaces. Such an approximation is demonstrated in Fig. 2, where we show, for the optically active $0_u^+$ states, the IHFSCC(b) cuts of the potential energy surface along corresponding cuts of the CASPT2 potential, after the CASPT2 bond lengths were uniformly shifted by 0.035 Å and a shift of 0.15 eV was applied to the excited states. After these two corrections, we observe a very good agreement between the two methods, with rather small discrepancies for the second $0_u^+$ state. Agreement for the first $0_u^+$ state, however, seems to be somewhat poorer than for the second.

#### 3. A closer look at the $0_u^+$ states

The two strong, allowed transitions for $I_3^-$ occur from the $\Omega=0_g^+$ ground state to two states with $\Omega=0_u^+$. Table V displays the dominant contributions to these transitions. If available we also show oscillator strengths for the corresponding transitions. Since the CASPT2 calculations were done in a two-step fashion we can also analyze the composition of these excited states in terms of spin-orbit free states. To a first approximation one can view these states as $^3\Pi_u$ and $^1\Sigma_u^+$, the former borrowing intensity from the latter through the spin-orbit coupling. This picture correlates very well with the DIM-SO results of Kosloff and co-workers.[19] In an orbital picture all methods give the same description, with different mixings of the two transitions from the occupied $\sigma_g$ and $\pi_g$ orbitals to the $\sigma_u$ LUMO. In the IHFSCC(b) calculations two additional excitations appear that can be interpreted as providing orbital relaxation of the LUMO. Since the DFT calculations also include spin-orbit coupling at the SCF stage we see a similar picture as for IHFSCC and MRCI, except that relaxation effects do not play an important role (since the virtual orbitals see the same potential as the occupied ones).

Nakanishi *et al.*[18] reported calculated squared transition moments ($\mu^2$) 7.31 and 13.51 $a_0^2$ for the two lower $0_u^+$ states, respectively, which translates into oscillator strengths of 0.68 and 1.56, thus giving somewhat more relative weight to the lower state than what we observe at the CASPT2 and IHFSCC(b) level and shown in Table V. Gabes and Stufkens[7] reported an intensity ratio of around 1.6:1.0 between the upper and lower $0_u^+$ states. Since CASPT2, MRCI, and TDDFT/SAOP all overestimate the peak separation between these states and thus underestimate the spin-orbit coupling between the underlying $^3\Pi_u$ and $^1\Sigma_u^+$ $\Lambda S$ states, these methods are also unable to reproduce the experimental intensity ratio. IHFSCC does get the peak separation right, but unfortunately we do not have access to oscillator strengths for this methodology.

### C. Benchmark calculations on $I_3$: Electronic spectra and electron affinities

Given that the IHFSCC results for the $(1h,0p)$ sector are generated as by-products of the excitation energies determination for $I_3^-$, we can also present spectroscopic constants and vertical and adiabatic excitation energies for this species, calculated with IHFSCC(b). We furthermore provide a comparison of this radical with IHFSCC, CASPT2, and MRCI at $r=2.84$ Å, which is the geometry employed by Kosloff and co-workers[19] for the same system. These results are also shown in Table VI.

TABLE VI. Comparison of excitation energies $T_v$ (in eV) for different methods [CASPT2, IHFSCC(a,b)], and MRCI(a) for neutral $I_3$ at $r_1=r_2=2.84$ Å]. Vertical ($T_v$) and adiabatic ($T_e$) excitation energies, as well as EAs are shown for IHFSCC(b).

| | MRCI[a] | MRCI(a) | CASPT2 | IHFSCC(a) | Expt. | | | IHFSCC(b) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| State ($\Omega$) | | $T_v(r_1=r_2=2.84$ Å) | | | $\omega_e$ | $T_e$ | EA[b] | $R_e$ | $\omega_e$ | $T_v$ [c] | $T_v$ | $T_e$ | EA[d] |
| X $(3/2_u)$ | 0.00 | 0.00 | 0.00 | 0.00 | $115\pm5$ | 0.00 | $4.15\pm0.12$ | 2.828 | 132 | 0.00 | 0.00 | 0.00 | 4.29 |
| A $(1/2_g)$ | 0.28 | 0.33 | 0.42 | 0.27 | | 0.28 | | 2.884 | 113 | 0.24 | 0.25 | 0.23 | |
| B $(1/2_u)$ | 0.61 | 0.65 | 0.55 | 0.64 | | 0.62 | | 2.837 | 129 | 0.65 | 0.65 | 0.65 | |
| C $(3/2_g)$ | 0.78 | 0.78 | 0.77 | 0.88 | | 0.68 | | 2.951 | 115 | 0.86 | 0.88 | 0.78 | |
| D $(1/2_g)$ | 1.62 | 1.63 | 1.55 | 1.68 | | | | 2.933 | 118 | 1.68 | 1.70 | 1.62 | |

[a]Results from Ref. 19 at $r_1=r_2=2.836$ Å.
[b]Reference 76; vertical detachment energy of $I_3^-$: 4.25 eV (Ref. 34).
[c]Calculated at $r_1=r_2=2.84$ Å.
[d]Value for the adiabatic EA. Vertical processes: IP($I_3^-$)=4.39 eV and EA($I_3$)=4.20 eV.

The excitation energies calculated at $r=2.84$ Å are in good agreement with those of Kosloff *et al.*, and, for the first two excited states, also with the experimental data. The $\Omega=3/2_g$ third excited state has a much longer bond length than the ground state which may explain the 0.1–0.2 eV overestimation found relative to the experimental value when calculated as vertical excitation. The adiabatic results for IHFSCC(b) indeed show a decrease of 0.10 eV from the vertical to the adiabatic excitation energy.

In line with the discussion above, we see that the results obtained with the different methods are fairly consistent. We find a ground state with $\Omega=3/2_u$, followed by states of $\Omega=1/2_g$, $1/2_u$, $3/2_g$, and $1/2_g$, respectively, in agreement with the experimental assignment.[33,34]

The CASPT2 excitation energies are lower than those of IHFSCC(b) by about 0.1–0.2 eV with the exception of the first excited state, which is higher for CASPT2. In this case, however, MRCI and IHFSCC are much more alike than for $I_3^-$, with discrepancies generally smaller than 0.1 eV. This strong similarity could be due to the fact that for the $(1h,0p)$ sector used here the exponential parametrization of IHFSCC for the excited states is truncated at the linear term and therefore is essentially the same as in MRCI.[62–65]

From the IHFSCC(b) calculations we furthermore obtain an adiabatic (vertical) electron affinity (EA) of 4.29 eV (4.20 eV) which compares well with the adiabatic EA of $4.15\pm0.12$ eV, including zero-point vibrational corrections, obtained experimentally from a thermodynamic cycle,[76] as well as the vertical detachment energy of $I_3^-$ of 4.25 eV reported by Choi *et al.*[34]

## IV. CONCLUSIONS

We have performed correlated electronic structure calculations including spin-orbit effects at high levels of theory on the triiodide ion ($I_3^-$) and the radical ($I_3$). The agreement between the different wave function-based methodologies employed is reasonable, as is their agreement with experimental results. Exploratory TDDFT calculations with the SAOP functional provide a qualitatively correct picture not too far from the equilibrium distance, but show unsystematic errors that prohibit use in quantitative description. The SAOP functional was chosen on the one hand because it has been constructed to correct the wrong asymptotic behavior of pure exchange-correlation potentials, and on the other hand because it has shown good performance for excitation energies of molecules containing heavy elements.[77–79] We believe that the principal reason for the relatively poor performance of TDDFT/SAOP is self-interaction errors.[80] This is indicated by the severe overestimation of the equilibrium bond lengths and the observation that the introduction of exact Hartree–Fock exchange through hybrid functionals gives spectroscopic constants in much better agreement with experiment. We have not undertaken a systematic study of the performance of other DFT functionals. Whereas there is a good reason to believe that some functionals will provide better vertical excitation energies at the equilibrium distance, we suspect that they will all encounter stability problems along the dissociation channels.

With respect to $I_3^-$, of the different methodologies evaluated, the intermediate Hamiltonian Fock-space coupled cluster is clearly the method that most accurately and consistently provides a picture which is both qualitatively and quantitatively correct for the excitation processes taking place in the initial steps in the photodissociation of the triiodide ion and the triiodide radical. Due to convergence problems this method is unfortunately not applicable to the complete potential energy surface, but we have been able to show that other wave function-based methods can reproduce the benchmark IHFSCC results rather well.

While it can be argued that for the triiodide species MRCI is slightly more accurate than SO-CASPT2, the latter has two interesting advantages: For one thing, it is computationally much more efficient than MRCI, and for another its errors seem to be very systematic for all electronic states considered. This systematic nature of errors is observed, in particular, for energies, allowing for global correction to be applied to the potential energy surfaces in order to bring them in agreement with IHFSCC. All is not well, however, with CASPT2. We observe a strong sensitivity of the method to symmetry breaking. A simulation of the photodissociation of $I_3^-$ will require as a minimum the generation of the 2D potential surfaces of linear $I_3^-$. However, in order to avoid symmetry breaking such a scan can at best be carried out in $C_2$ symmetry, for which the distinction between $0^+$ and $0^-$ states is lost and needs to be recovered by the use of scripts when analyzing the individual states. CASPT2 is furthermore susceptible to intruder states, and we have been unable

to remove the spikes observed in the potential curves toward dissociation by the use of level shifting techniques. One alternative is to explore a possibly more robust multireference perturbation theory such as n-electron valence second-order perturbation theory.[81]

MRCI clearly is another option, but presently too expensive to be able to generate a full set of potential energy surfaces for studies of the dynamics of the photodissociation process. The challenge of applying such methods to heavy-element systems is not only the mandatory treatment of spin-orbit coupling, but also the fact that heavy atoms are increasingly polarizable such that subvalence has to be correlated. In the present case we observe that the correlation of 4d is crucial for obtaining correct spectroscopic constants, less so for vertical excitation energies. Subvalence correlation can possibly be avoided by employing relativistic effective core potentials combined with core-polarization potentials.[82,83] Finally, we would like to suggest that a successful combination of MCSCF and DFT (Ref. 84) could be the ideal tool for generating potential surfaces for dynamics.

## ACKNOWLEDGMENTS

We dedicate this paper to a giant of quantum chemistry, Björn Roos. Bravely fighting illness, he succumbed in the end. In the preparation of this work he (as well as Per-Åke Malmqvist and Valera Veryazov) was most helpful in answering questions related to the SO-CASPT2 calculations. Once, only a few days after a mail with questions, he answered as usually to the point, but excusing himself for the late reply. His mail was written an early Sunday morning, at a time when most of his younger and healthier colleagues were probably still sound asleep.

We wish to thank Professor Ronnie Kosloff for calling our attention to the need for a more accurate description of the potential energy surfaces for the triiodide dissociation problem. Moreover, we thank the DEISA Consortium (cofunded by the EU, FP6 Project No. 508830/031513), for support within the DEISA Extreme Computing Initiative (www.deisa.org). In particular, we wish to acknowledge the work done by Walter Lioen (SARA/NL) in improving the efficiency of the four-index transformation code within the DIRAC program for the architectures used. We also acknowledge support by CINES ("Centre Informatique National de l'Enseignement Supérieur") under Grant No. phl2531. This work has been supported by the Netherlands Organization for Scientific Research (NWO) via a Vici grant for L.V.

[1] M. Fushitani Annu. Rep. Prog. Chem., Sect. C: Phys. Chem. **104**, 272 (2008).
[2] D. M. Neumark, J. Chem. Phys. **125**, 132303 (2006).
[3] R. J. Hach and R. E. Rundle, J. Am. Chem. Soc. **73**, 4321 (1951).
[4] G. C. Pimentel, J. Chem. Phys. **19**, 446 (1951).
[5] G. A. Landrum, N. Goldberg, and R. Hoffmann, J. Chem. Soc. Dalton Trans. **1997**, 3605.
[6] P. H. Svensson and L. Kloo, Chem. Rev. (Washington, D.C.) **103**, 1649 (2003).
[7] W. Gabes and D. J. Stufkens, Spectrochim. Acta, Part A **30**, 1835 (1974).
[8] M. Mizuno, J. Tanaka, and I. Harada, J. Phys. Chem. **85**, 1789 (1981).
[9] H. Isci and W. R. Mason, Inorg. Chem. **24**, 271 (1985).
[10] T. Kühne and P. Vöhringer, J. Chem. Phys. **105**, 10788 (1996).
[11] T. Kühne, R. Küster, and P. Vöhringer, Chem. Phys. **233**, 161 (1998).
[12] T. Kühne and P. Vöhringer, J. Phys. Chem. A **102**, 4177 (1998).
[13] A. Baratz and S. Ruhman, Chem. Phys. Lett. **461**, 211 (2008).
[14] M. T. Zanni, B. J. Greenblatt, A. V. Davis, and D. M. Neumark, J. Chem. Phys. **111**, 2991 (1999).
[15] H. Choi, R. T. Bise, A. A. Hoops, and D. M. Neumark, J. Chem. Phys. **113**, 2255 (2000).
[16] A. A. Hoops, J. R. Gascooke, A. E. Faulhaber, K. E. Kautzman, and D. M. Neumark, J. Chem. Phys. **120**, 7901 (2004).
[17] L. Zhu, K. Takahashi, M. Saeki, T. Tsukuda, and T. Nagata, Chem. Phys. Lett. **350**, 233 (2001).
[18] R. Nakanishi, N. Saitou, T. Ohno, S. Kowashi, S. Yabushita, and T. Nagata, J. Chem. Phys. **126**, 204311 (2007).
[19] J. Vala, R. Kosloff, and J. N. Harvey, J. Chem. Phys. **114**, 7413 (2001).
[20] L. Visscher, E. Eliav, and U. Kaldor, J. Chem. Phys. **115**, 9720 (2001).
[21] A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **113**, 9905 (2000).
[22] A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **115**, 6862 (2001).
[23] T. Fleig, J. Olsen, and L. Visscher, J. Chem. Phys. **119**, 2963 (2003).
[24] T. Fleig, H. J. Aa. Jensen, J. Olsen, and L. Visscher, J. Chem. Phys. **124**, 104106 (2006).
[25] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).
[26] K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).
[27] E. K. U. Gross and W. Kohn, Adv. Quantum Chem. **21**, 255 (1990).
[28] D. Tozer, J. Chem. Phys. **119**, 12697 (2003).
[29] M. E. Casida, F. Gutierrez, J. Guan, F.-X. Gadea, D. Salahub, and J.-P. Daudey, J. Chem. Phys. **113**, 7062 (2000).
[30] M. E. Casida, J. Chem. Phys. **122**, 054111 (2005).
[31] D. J. Tozer and N. C. Handy, Phys. Chem. Chem. Phys. **2**, 2117 (2000).
[32] N. T. Maitra, F. Zhang, R. J. Cave, and K. Burke, J. Chem. Phys. **120**, 5932 (2004).
[33] T. R. Taylor, K. R. Asmis, M. T. Zanni, and D. M. Neumark, J. Chem. Phys. **110**, 7607 (1999).
[34] H. Choi, T. R. Taylor, R. R. Bise, A. A. Hoops, and D. M. Neumark, J. Chem. Phys. **113**, 8608 (2000).
[35] T. Migchelsen and A. Vos, Acta Crystallogr. **23**, 796 (1967).
[36] DIRAC, a relativistic *ab initio* electronic structure program, release DIRAC04.0, written by H. J. Aa. Jensen, T. Saue, and L. Visscher with contributions from V. Bakken, E. Eliav, T. Enevoldsen, T. Fleig, O. Fossgaard, T. Helgaker, J. Laerdahl, C. V. Larsen, P. Norman, J. Olsen, M. Pernpointner, J. K. Pedersen, K. Ruud, P. Salek, J. N. P. van Stralen, J. Thyssen, O. Visser, and T. Winther, 2004.
[37] M. Iliaš and T. Saue, J. Chem. Phys. **126**, 064102 (2007).
[38] B. Schimmelpfennig, AMFI: an atomic mean-field code, Stockholm, Sweden, 1996.
[39] B. A. Heß, C. M. Marian, U. Wahlgren, and O. Gropen, Chem. Phys. Lett. **251**, 365 (1996).
[40] K. G. Dyall, Theor. Chem. Acc. **108**, 335 (2002).
[41] K. G. Dyall, Theor. Chem. Acc. **115**, 441 (2006).
[42] S. Knecht, H. J. Aa. Jensen, and T. Fleig, J. Chem. Phys. **128**, 014108 (2008).
[43] S. Knecht, H. J. Aa. Jensen, and T. Fleig, J. Chem. Phys. **132**, 014108 (2010).
[44] S. R. Knecht, "Parallel relativistic multiconfiguration methods: New powerful tools for heavy-element electronic-structure studies," Ph.D. dissertation, Heinrich-Heine-Universität Düsseldorf, 2009.
[45] G. Karlström, R. Lindh, P.-A. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, and L. Seijo, Comput. Mater. Sci. **28**, 222 (2003).
[46] B. A. Hess, Phys. Rev. A **33**, 3742 (1986).
[47] P.-A. Malmqvist, B. O. Roos, and B. Schimmelpfennig, Chem. Phys. Lett. **357**, 230 (2002).
[48] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark, J. Phys. Chem. A **108**, 2851 (2004).
[49] G. Ghigo, B. O. Roos, and P.-A. Malmqvist, Chem. Phys. Lett. **396**, 142 (2004).
[50] ADF, Amsterdam density functional program, Theoretical Chemistry, Vrije Universiteit Amsterdam, URL: http://www.scm.com, 2007.
[51] E. van Lenthe, J. G. Snijders, and E. J. Baerends, J. Chem. Phys. **105**, 6505 (1996).
[52] P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen, and E. J.

Baerends, J. Chem. Phys. **112**, 1344 (2000).
[53] E. van Lenthe and E. J. Baerends, J. Comput. Chem. **24**, 1142 (2003).
[54] J. Perdew, J. Chevary, S. Vosko, K. Jackson, M. Pederson, D. Sing, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[55] For further details on the use of the SAOP potential, see the ADF document http://www.scm.com/Doc/Doc2009.01/ADF/ADFUsersGuide/page95.html, retrieved July 14, 2010.
[56] K. B. J. P. Perdew and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[57] Y. Zhao and D. G. Truhlar, J. Chem. Phys. **125**, 194101 (2006).
[58] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).
[59] Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).
[60] M. T. Zanni, V. S. Batista, W. H. Miller, B. J. Greenblatt, A. V. Davis, and D. M. Neumark, J. Chem. Phys. **110**, 3748 (1999).
[61] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory* (Wiley, Chichester, 2000).
[62] R. J. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).
[63] L. Meissner and R. J. Bartlett, J. Chem. Phys. **94**, 6670 (1991).
[64] L. Meissner and R. J. Bartlett, J. Chem. Phys. **102**, 7490 (1995).
[65] D. Mukhopadhyay, S. Mukhopadhyay, R. Chaudhuri, and D. Mukherjee, Theor. Chim. Acta **80**, 441 (1991).
[66] F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet, and E. Eliav, J. Phys. Chem. A **113**, 12504 (2009).
[67] L. Meissner, J. Chem. Phys. **108**, 9227 (1998).
[68] M. Musial and R. J. Bartlett, J. Chem. Phys. **129**, 044101 (2008).
[69] M. Musial and R. J. Bartlett, J. Chem. Phys. **129**, 134105 (2008).
[70] M. Musial and R. J. Bartlett, Chem. Phys. Lett. **457**, 267 (2008).
[71] I. Infante, A. S. P. Gomes, and L. Visscher, J. Chem. Phys. **125**, 074301 (2006).
[72] G. La Macchia, I. Infante, R. Juraj, J. K. Gibson, and L. Gagliardi, Phys. Chem. Chem. Phys. **10**, 7278 (2008).
[73] F. Notter, S. Dubillard, and H. Bolvin, J. Chem. Phys. **128**, 164315 (2008).
[74] B. O. Roos and K. Andersson, Chem. Phys. Lett. **245**, 215 (1995).
[75] N. Forsberg and P.-Aa. Malmqvist, Chem. Phys. Lett. **274**, 196 (1997).
[76] K. Do, T. P. Klein, C. A. Pommerening, and L. S. Sunderlin, J. Am. Soc. Mass Spectrom. **8**, 688 (1997).
[77] J. Gao, W. Zou, W. Liu, Y. Xiao, D. Peng, B. Song, and C. Liu, J. Chem. Phys. **123**, 054102 (2005).
[78] K. Pierloot, E. van Besien, E. van Lenthe, and E. J. Baerends, J. Chem. Phys. **126**, 194311 (2007).
[79] R. Bast, H. J. A. Jensen, and T. Saue, Int. J. Quantum Chem. **109**, 2091 (2009).
[80] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[81] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J. P. Malrieu, J. Chem. Phys. **114**, 10252 (2001).
[82] P. Fuentealba, H. Preuss, H. Stoll, and L. V. Szentpály, Chem. Phys. Lett. **89**, 418 (1982).
[83] M. Dolg, in *Relativistic Electronic Structure Theory—Part 1. Fundamentals*, edited by P. Schwerdtfeger (Elsevier, Amsterdam, 2002), p. 523.
[84] E. Fromager, J. Toulouse, and H. J. A. Jensen, J. Chem. Phys. **126**, 074111 (2007).

## C.3 Paper III

# Benchmarking Electronic Structure Calculations on the Bare $UO_2^{2+}$ Ion: How Different are Single and Multireference Electron Correlation Methods?†

**Florent Réal,‡ André Severo Pereira Gomes,§ Lucas Visscher,§ Valérie Vallet,*,‡ and Ephraim Eliav||**

*Université Lille1 (Sciences et Technologies), Laboratoire PhLAM, CNRS UMR 8523, CERLA, CNRS FR 2416, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France, Amsterdam Center for Multiscale Modeling, Department of Theoretical Chemistry, Faculty of Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands, and School of Chemistry, Tel Aviv University, 69978 Tel Aviv, Israel*

*Received: April 23, 2009; Revised Manuscript Received: June 15, 2009*

In a recent investigation by some of us on the spectrum of the uranyl ($UO_2^{2+}$) ion [Réal, F.; Vallet, V.; Marian, C.; Wahlgren, U. *J. Chem. Phys.* **2007**, *126*, 214302], a sizable difference between CASPT2 and linear response coupled cluster (LRCC) was observed both with and without the perturbative inclusion of spin−orbit coupling. This poses a serious question as to which of the two would be more reliable for investigating molecules containing actinides. In this paper we address this question by comparing CASPT2 and LRCC to a method known to accurately describe the spectra of actinide-containing molecules: the four-component intermediate Hamiltonian Fock-space coupled cluster (IHFSCC) method, where electron correlation and spin−orbit coupling are treated on an equal footing. Our results indicate that for $UO_2^{2+}$ there is little difference between treatments of spin−orbit coupling, making electron correlation the main cause of discrepancies. We have found IHFSCC and LRCC to be the most alike in the overall description of excited states, even though individual LRCC energies are blue-shifted in comparison to IHFSCC due to the difference in the parametrization of the excited states' wave functions. CASPT2, on the other hand, shows good agreement with IHFSCC for individual frequencies but significantly less so for the spectrum as a whole, due to the difference in the degree of correlation recovered in both cases.

## 1. Introduction

The chemistry of actinide-containing molecules is a rich and fascinating subject, in particular because of their spectroscopic and luminescence properties. The luminescence of uranyl(VI) $UO_2^{2+}$ has been extensively studied experimentally both in aqueous solution and in crystals. However, the assignment of the energy levels responsible for the strong absorption in the UV range and the long-lived luminescent state that emits light in the visible (20 000−26 000 $cm^{-1}$) range is not trivial.[1−3] There have been several theoretical studies in the past decade on the electronic spectrum of uranyl(VI), either as a bare ion[4−7] or coordinated[5,6,8,9] to other species, with the 2-fold aim of comparing theoretical methods and reproducing the experimental data available in the condensed phase. Theoretical approaches face several challenges due to the large number of electrons, which should be treated explicitly, and to the accurate description of the strong interactions of the uranyl with its surroundings (ligands, host crystals, or solvent molecules). Due to these interactions, unambiguous comparison between theoretical and experimental data is often difficult, and it is of interest to obtain benchmark theoretical data for the bare uranyl ion for which no experimental data is available.

In order to reach benchmark accuracy different hierarchies of methods could be applied: one can start from a relativistic framework, either via four-component[10,11] or two-component[12,13] treatments, in which spin−orbit is included a priori, and use multireference coupled-cluster[14−17] or configuration interaction (CI) methods[18−21] to treat electron correlation effects. An alternative is to use a two-step approach. In the first step, the correlation effect on the states of interest is treated at the spin-free (SF) level with CI-based methods such as multireference MRCI[22] or complete active space with second-order perturbation theory CASPT2[23−25] approaches, or by applying coupled cluster theory in the framework of response theory.[26−29] In the second step of the calculation, spin−orbit interaction between the various spin-free states is accounted for by performing spin−orbit configuration interaction (SOCI) calculations.[25,30] The effect of electron correlation computed in the first step is taken into account by means of an effective Bloch Hamiltonian.[31] The diagonalization of the total Hamiltonian yields energies and eigenvectors which take into account both correlation and spin−orbit effects. All of these methods are rather demanding and are often replaced by methods that reduce the work in either the treatment of electron correlation, such as time-dependent density functional theory (TDDFT), or in the treatment of relativity, e.g., using pseudopotentials, or both. Most approximate correlation schemes lack however, the possibility of systematically improving the description by going to the next level in a well-defined hierarchy.

In recent theoretical studies on the uranyl(VI) spectrum[7,9,32] several theoretical methods have been compared. Within the TD-DFT scheme, most density functionals do not yield accurate excitation energies; however, geometries and relaxation energies of the excited states are in most cases reasonably well described. Réal and co-workers[7] also demonstrated that different wave function-based methods generally provide qualitatively similar

† Part of the "Russell M. Pitzer Festschrift".

* To whom correspondence should be addressed. E-mail: valerie.vallet@univ-lille1.fr.

‡ CNRS UMR 8523.

§ Vrije Universiteit Amsterdam.

|| Tel Aviv University.

10.1021/jp903758c CCC: $40.75 
Published on Web 07/16/2009

results. However, in quantitative terms significant differences appear between the two methods expected to yield the most accurate results: CASPT2 and linear response coupled cluster (LRCC). The LRCC spectrum was blue-shifted in comparison with that of CASPT2 by about 3000 $cm^{-1}$ and quite similar to spectra obtained from multireference CI (MRCI) or averaged-quadratic coupled-cluster (AQCC) calculations. The only experimental spectra available are for crystals such as $Cs_2UO2Cl_4$[3] or in solution.[33] Matsika and Pitzer[5] and Pierloot et al.[6,9] have shown that the environment (the equatorial ligands and the rest of the crystal) may modify the character of the excited states, apart from changing significantly the transition energies. This greatly reduces the usefulness of comparing calculated energy levels for the bare uranyl to experimental values.

The need for understanding the origin of the discrepancies among wave function based methods in the computed uranyl energy levels has motivated us to employ the intermediate Hamiltonian Fock space coupled cluster (IHFSCC) method, a true multireference coupled cluster method that, in its relativistic formulation,[14−16] allows us to consider spin−orbit coupling and electron correlation on the same footing. The accuracy of the relativistic IHFSCC approach has been demonstrated in several investigations on actinyl species.[34−37] Moreover, Fock-space coupled cluster provides an ideal measure for the relative accuracy and reliability of both LRCC and CASPT2 because it is fully size-extensive for both ground and excited states (as opposed to LRCC, which is formally so only for the ground state) and includes electron correlation to infinite order (as opposed to CASPT2, which does so to second order). By providing a comparison of LRCC and CASPT2 to the accurate IHFSCC, we complement the picture obtained from previous studies, since the former have been compared extensively to more approximate wave function-based methods (e.g., MRCI and AQCC) and TD-DFT.[7,9,32]

## 2. Computational Details

**2.1. Fock-Space Coupled Cluster.** The calculations of the excitation spectrum of $UO_2^{2+}$ were performed with a development version of the Dirac08 program,[38] using three different approaches to treat relativity. In the spin-free Dirac−Coulomb (SFDC)[11,39] calculations, we eliminate all spin−orbit coupling terms to allow for straightforward comparison with earlier works. This approach is an approximation to the regular 4-component Dirac−Coulomb (DC) calculation, in which only the usual approximation of (SS|SS) integrals by an a posteriori correction[40] is applied. The third approach concerns the eXact 2-Component (X2C) approach recently introduced by Iliaš and Saue[13] in which spin−orbit coupling is included from the start via atomic mean-field integrals calculated with the AMFI code.[41,42]

In most of the calculations the valence double (DZ) or triple-$\zeta$ (TZ) basis sets by Dyall[43] were used for the uranium atom, but we also considered the Fægri set,[44] that corresponds to triple-$\zeta$ (TZ) quality in the valence s functions, and quadruple- $\zeta$ (QZ) or higher quality for the higher angular momenta. For oxygen we employed the (aug)-cc-VDZ and aug-cc-pVTZ basis sets of Dunning and co-workers.[45] All of these basis sets were kept uncontracted in all calculations.

The potential energy curves for the symmetric stretch were sampled at 14 different uranium−oxygen bond lengths ($r_{UO}$) within the range $r_{UO} \in [1.58;1.92]$ (Å). This ground-state potential energy curve was described by the Dirac−Coulomb coupled cluster single and double without (CCSD) and with perturbative treatment of triples (CCSD(T)) method,[46−49] while

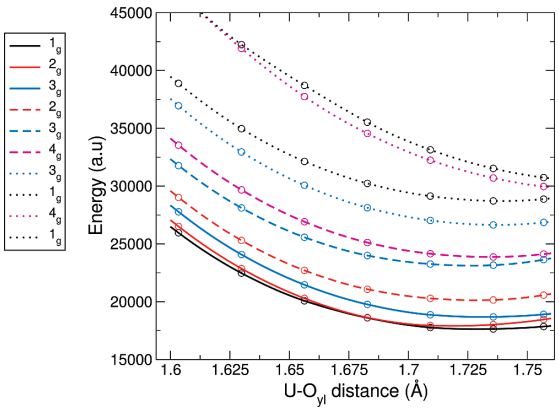



**Figure 1.** Potential curves of the uranyl (VI) bare ion along the symmetric stretching mode computed with the SO-IHFSCC method. The $1_g$ states are drawn in black, the $2_g$ in red, the $3_g$ in blue, and the $4_g$ in magenta.

the curves for the excited states were obtained within the IHFSCCSD scheme using the "one particle, one hole" sector (1*h*,1*p*) of Fock space. In the calculations, several active spaces were tested; only orbitals with orbital energies (in au) $\varepsilon \in [-6.00; 20.00]$ (34 electrons), $[-3.00; 20.00]$ (24 electrons; uranium 5d frozen), and $[-3.00; 40.00]$ (24 electrons; testing the effect of higher lying virtual orbitals) are included in the correlation treatment. These different active spaces will be referred to as AS1, AS2, and AS3, respectively.

In Fock-space coupled cluster calculations one should subdivide the space spanned by the active orbitals in two subspaces: the model or $P$ space, containing the active valence orbitals which are directly involved in the electronic excitations and the complementary $Q$ space that includes the remaining "correlation-active" orbitals from AS1, AS2 or AS3. In the present case this translates into including the highest occupied orbitals ($\sigma_{1/2u}$, $\sigma_{1/2g}$, $\pi_{1/2u}$, $\pi_{3/2u}$, $\pi_{1/2u}$ and $\pi_{3/2g}$) and the ten lowest unoccupied (the nonbonding uranium $f^{\delta}_{3/2u}$, $f^{\delta}_{5/2u}$ and $f^{\phi}_{5/2u}$, $f^{\phi}_{7/2u}$, as well as the antibonding $\sigma^*_{1/2u}$, $\sigma^*_{1/2g}$, $\pi^*_{1/2u}$, $\pi^*_{3/2u}$, $\pi^*_{1/2u}$ and $\pi^*_{3/2g}$) orbitals in the ($P$) model space. This space can be extended by considering in addition the deeper lying occupied orbitals (6s, 6p, and 5d on uranium and 2s on the oxygens) and/or higher virtual orbitals (see Figure 1 in ref 3 for a schematic spin-free picture on the composition of these orbitals).

The intermediate Hamiltonian facilitates such extensions of the active space by allowing for a further subdivision of the resulting model space into 2 subspaces, the main model ($P_m$) space and an intermediate model ($P_i$) space that is not dressed and serves as a buffer between the $P_m$ and $Q$ spaces, thus alleviating "intruder" state problems. The main model space $P_m$ (consisting here of 14 electrons and 34 spinors for AS1, and of 12 electrons and 24 spinors for AS2 and AS3, respectively) is built solely from excitations within the subset of main active valence orbitals, whereas every configuration from the $P_i$ subspace involves at least one intermediate active valence orbital.

One should realize that accurate solutions are only obtained for states dominated by $P_m$ components. The undressed $P_i$ space merely serves as a buffer to eliminate intruder state problems that make convergence in traditional Fock space approaches difficult. The scheme employed here is known in literature as IH2,[50,51] but for simplicity we refer to it simply as IHFSCC. This method forces the $P_i \rightarrow Q$ transition amplitudes of the wave

**TABLE 1: Calculated U−$O_{yl}$ Distances $R_e$ (Å) and Harmonic Frequencies $\omega_e$ ($cm^{-1}$) Computed for Various Correlation Methods Including All Relativistic Effects**

| | 4C | | | | 1C | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| method | this work | | de Jong et al.[a] | | Straka et al.[b] | | Réal et al.[c] | | Pierloot et al.[d] | |
| ref. | $R_e$ | $\omega_e$ | $R_e$ | $\omega_e$ | $R_e$ | $\omega_e$ | $R_e$ | $\omega_e$ | $R_e$ | $\omega_e$ |
| MP2 | 1.724 | 957 | 1.739 | 944 | 1.728 | 1053 | | | | |
| CCSD | 1.685 | 1103 | 1.697 | 1041 | | | 1.679 | 1038 | | |
| CCSD(T) | 1.703 | 1016 | 1.715 | 974 | 1.702 | 1111 | | | | |
| CASPT2 | | | | | | | | | 1.708 | 1103 |

[a] Reference 55a. [b] Spin-free values; ref 56. [c] Reference 7. [d] Reference 6.

function operator to be zero, which makes the scheme easy and efficient to use.[34,35]

**2.2. Linear Response CCSD.** Calculations with the linear response CCSD method[26−29] were made with the implementation available in the Dalton 2.0 package.[52] These were carried out to check the influence of uncontracting the basis sets in such calculations, in order to make the comparison of our current results and the previous spin-free LR-CCSD work (in which a contracted atomic natural orbital basis set was used) more straightforward. The Fægri basis set mentioned above was thereby used for uranium, along with the cc-pVTZ basis for oxygen. In this case molecular orbitals were obtained from a scalar relativistic Hartree−Fock calculation using the Douglas−Kroll−Hess Hamiltonian.[53,54] In the spin-free LR-CCSD calculation, the uranium atomic orbitals below the 6s were kept frozen as well as the 1s oxygen atomic orbitals, thus correlating 24 electrons. The highest virtual orbitals with energies $\varepsilon \geq 40$ au were discarded from the orbital correlation space.

**2.3. Computational Requirements.** Given this range of different methods employed in the current work it may be of interest to list some representative computational requirements for the calculations performed: the one-component correlated calculations can be performed on a typical Linux cluster with one processor and about 1 Gb memory, while the requirements of the four-component DC-IHFSCC go up to 4Gb memory (in solving the amplitude equations for the different sectors of the Fock space) and about 80 Gb of disk space (in the transformation from AO to MO basis, prior to the coupled cluster calculation) per processor, which are more easily met on a supercomputer.

## 3. Results and Discussion

**3.1. Ground-State Spectroscopic Constants.** The spectroscopic constants of the ground state of $UO_2^{2+}$ already provide information on possible differences between the combinations of basis sets and Hamiltonians that we wish to compare. The bond lengths and vibrational frequencies computed with the (4C−)MP2, (4C−)CCSD and (4C−)CCSD(T) methods are reported in Table 1 and may be compared to those available in the literature.[6,7,55,56]

Our calculations yield U−$O_{yl}$ bond lengths of 1.724, 1.685, and 1.703 Å for 4C-MP2, 4C−CCSD, and 4C−CCSD(T), respectively (from fitting 7 points near the respective minima with a 4th degree polynomial). It is clear that the inclusion of triple excitations in the coupled cluster allows for relaxation of the U−$O_{yl}$ bonding orbitals, resulting in a slight bond lengthening (0.02 Å), while the bond distance is overestimated by 4C-MP2. These results are quite similar to those obtained by de Jong et al.[55] in previous four-component calculations with the differences of 0.012 Å between the two sets of results likely arising from the different basis sets used in both studies - those used here are more flexible than those used by de Jong et al.[55] The 4C-MP2 results are also quite similar to those obtained by Straka et al.[56] with a scalar relativistic MP2 calculation, as was expected since spin−orbit effects are known to be of little importance for the properties of the uranyl closed-shell ground state.[57,58]

The SO-CASPT2 bond lengths[6] are shorter by about 0.02 Å than the 4C-MP2 ones, while they differ by less than 0.005 Å from the 4C−CCSD(T) values. This is surprising at first glance as one might expect MP2 and CASPT2 to yield similar results for the uranyl ground state for which the wave function is dominated by a single closed-shell determinant.[59] Even so, the ground-state CAS wave function (with an active space including the six bonding, six antibonding orbitals and four nonbonding orbitals) computed in the one component framework includes a significant contribution (about 13%) from double excited determinants. This leads to higher order excitations in CASPT2, relative to the MP2 and a better description of electron correlation, thus bringing the result closer to CCSD(T) accuracy.

Vibrational frequencies for the ground state, computed in the harmonic approximation, are also given in Table 1. As was the case for the bond lengths, we see in general good agreement with the results of de Jong et al.,[55] with discrepancies of about 40 $cm^{-1}$ for 4C−CCSD and 4C−CCSD(T) numbers, but only 13 $cm^{-1}$ for the 4C-MP2 ones. The frequency values vary in accordance with the trend observed for the equilibrium bond distances: the larger the bond length, the smaller the frequency. The SO-LR-CCSD results of Réal et al.[7] differ by 65 $cm^{-1}$ from the 4C−CCSD ones. Although the bond lengths computed by Straka et al.[56] agreed with the 4C-values within 0.001 Å, the stretching frequencies show discrepancies of about 100 $cm^{-1}$. Similarly, for SO-CASPT2 the equilibrium bond distance[9] are nearly identical to the 4C−CCSD(T) value, while the symmetric stretching frequency is larger by about 90 $cm^{-1}$. This could indicate that the methods differ somewhat in the description of the electron correlation along the internal coordinates, but we are uncertain as to what extent these differences are intrinsic to the methods or whether inaccuracies in the determination of the frequencies play a role. Our experience obtaining frequencies from polynomial fits near the minima for this system indicates that these are rather sensitive to the fitting procedure, and lack of information as to how these have been obtained in the previous works restricts our ability to test this dependence on the fit.

**3.2. Comparison of Vertical Transitions Computed in Scalar Relativistic Calculations.** We start our discussion of the excited state calculations by comparing results obtained without inclusion of spin−orbit coupling, to focus on the influence of the electron correlation treatment. Results from spin-free calculations are reported in Table 2. We first note that differences due to choice of basis set are small: the two sets of LR-CCSD transition energies differ by at most 1000 $cm^{-1}$. The differences due to the choice of correlation method are very significant, with LR-CCSD transition energies of both the *u* and

**TABLE 2: Spin-Free Vertical Transitions Energies $\Delta E^0$ (in $cm^{-1}$) of $UO_2^{2+}$ Computed at the IHFSCCSD, LR-CCSD, and CASPT2 levels[a]**

| | IHFSCCSD[b] | | | LR-CCSD[b] | | LR-CCSD[c] | | CASPT2[d] | |
|---|---|---|---|---|---|---|---|---|---|
| state | character | $\Delta E^0$ | $\Delta E^1$ | $\Delta E^0$ | $\Delta E^1$ | $\Delta E^0$ | $\Delta E^1$ | $\Delta E^0$ | $\Delta E^1$ |
| $a^3\Delta_g$ | $\sigma_u f_\delta$ | 20972 | | 24378 | | 24441 | | 22477 | |
| $a^3\Phi_g$ | $\sigma_u f_\phi$ | 23050 | (2078) | 26461 | (2083) | 26154 | (1713) | 23689 | (1212) |
| $a^1\Phi_g$ | $\sigma_u f_\phi$ | 27545 | (6573) | 30975 | (6598) | 30936 | (6495) | 27966 | (5489) |
| $a^1\Delta_g$ | $\sigma_u f_\delta$ | 30177 | (9205) | 33911 | (9534) | 33927 | (9486) | 31437 | (8960) |
| $a^1\Delta_u$ | $\sigma_g f_\delta$ | 37876 | (16904) | 41555 | (17178) | 42100 | (17659) | 36644 | (14167) |
| $a^3\Delta_u$ | $\sigma_g f_\delta$ | 38339 | (17367) | 41664 | (17287) | 42234 | (17793) | 37283 | (14806) |
| $a^1\Pi_g$ | $\pi_u f_\delta$ | 38776 | (17804) | *42137* | (17759) | *43098* | (18657) | *37733* | (15256) |
| $a^3\Pi_g$ | $\pi_u f_\delta$ | 39029 | (18057) | *41278* | (16900) | *42188* | (17747) | *37181* | (14704) |
| $a^3\Gamma_g$ | $\pi_u f_\phi$ | 39611 | (18639) | *42168* | (17790) | *42981* | (18540) | *37563* | (15086) |
| $a^3\Phi_u$ | $\sigma_g f_\phi$ | 40059 | (19087) | 43368 | (18991) | 43728 | (19287) | 38016 | (15539) |
| $a^1\Phi_u$ | $\sigma_g f_\phi$ | 40632 | (19660) | *44137* | (19759) | 44487 | (20046) | 38248 | (15771) |
| $a^1\Gamma_g$ | $\pi_u f_\phi$ | 41435 | (20463) | *43909* | (19532) | 44686 | (20245) | 39172 | (16695) |
| $a^1\Pi_u$ | $\pi_g f_\delta$ | 46126 | (25154) | *50479* | (26101) | *51417* | (26976) | *43192* | (20715) |
| $a^3\Pi_u$ | $\pi_g f_\delta$ | 46134 | (25162) | *48491* | (24113) | *49426* | (24985) | *42602* | (20125) |
| $b^1\Gamma_u$ | $\pi_g f_\phi$ | 48442 | (27470) | 51293 | (26915) | 52022 | (27581) | 44314 | (21837) |
| $b^3\Gamma_u$ | $\pi_g f_\phi$ | 49144 | (28172) | 51127 | (26749) | 52222 | (27781) | 44784 | (22307) |

[a] The energy differences $\Delta E^1$ with respect to the first excited state are reported in parentheses. Changes in the ordering of the states are marked in italics. [b] Uranium Fægri and oxygen cc-pVTZ basis sets; $R(U-O_{yl}) = 1.683$ Å. [c] Reference 7; uranium ANO-RCC-QZP and oxygen ANO-RCC-TZP basis sets; $R(U-O_{yl}) = 1.683$ Å. [d] Reference 6; uranium DK3 and oxygen ANO-L basis sets; $R(U-O_{yl}) = 1.708$ Å.

*g* states being about 3000−3500 $cm^{-1}$ higher than the corresponding IHFSCC numbers (calculated in the same basis and at the same distance). Since the ground state of the $UO_2^{2+}$ ion can be described by a single determinant, the correlation in the coupled cluster method is in principle equivalent in the spin-free 4C and 1C frameworks. Thus, the discrepancy must come from differences in the correlation treatment for the excited states; more precisely due to the differences between IHFSCC and LRCC. These differences, which arise from the parametrization of the excited states wave functions, have been addressed extensively by other authors in previous publications,[60−64] but usually for rather light molecules with relatively few valence electrons.

LRCC is based upon a linear parametrization in terms of the ground-state coupled-cluster wave function, corresponding to a wave operator $\Omega^{LRCC} = (1 + C_k)$ that, upon application to the ground-state coupled cluster reference wave function, yields the excited state wave functions

$$|\Psi_k^{LR}\rangle = \Omega^{LR}|CC\rangle = (1 + C_k)\exp(T)|\Phi_0\rangle \qquad (1)$$

The (IH)FSCC parametrization is based upon the exponential of the cluster operator *S*, with a wave operator that is written as

$$|\Psi_k^{FS}\rangle = \Omega^{FS}|\Phi_0\rangle = \exp(S)|\Phi_0\rangle = \exp(S')\exp(T)|\Phi_0\rangle \qquad (2)$$

where *S* has been subdivided into ground-state amplitudes *T* and separate cluster operators for each sector of the Fock space ($S' = S(1,0) + S(0,1) + S(1,1)$) under consideration. For ionization potentials and electron affinities, IHFSCC and LRCC methods are formally equivalent since the expansion of exp (*S*′) truncates on the linear term.[62,63]

For higher sectors such as the (1*h*,1*p*) used in this work, both methods yield different results, since Fock-space methods contain terms such as $S^{(0,1)}S^{(1,0)}$,[62,63] which cancel out the disconnected terms arising from the linear parametrization in $\Omega^{LRCC}$ to third order or higher.[60−63]

This feature makes Fock-space methods size-extensive for both ground and excited states,[63,64] and should be the main reason for the systematic difference of about 3000−3500 $cm^{-1}$ in excitation energies between the two methods (compare the first and second column of Table 2). Recent numerical comparisons between Fock-space and LRCC approaches by Musial and Bartlett,[65−67] are in qualitative agreement with our results, with LRCC yielding higher excitation energies than IHFSCC. In general the differences observed in the molecules considered then ($N_2$, $H_2O$, CO) are smaller than the ones computed here, which could be due to the larger number of electrons contributing to the correlation energy differences in $UO_2^{2+}$.

Comparison between the coupled cluster methods and the CASPT2 approach is less straightforward, as both methods already differ in the parametrization of the ground-state wave function. This is evident from the results discussed in the previous section and poses the question how best to compare vertical excitation energies. We have thereby chosen to present energies computed at a near-optimal bond length for each method.

Comparing the first and last column of Table 2, we see that the CASPT2 energies in most cases overestimate the excitation energies relative to IHFSCC, but in contrast to the systematic shift found with LR-CCSD, we also find some transitions computed up to a few thousand $cm^{-1}$ lower than with IHFSCC.

The individual CASPT2 energies, particularly for the lower excited states (up to about 38 000 $cm^{-1}$), are in rather good agreement with IHFSCC ones, but this agreement deteriorates as higher states are considered. In the IHFSCC calculations the excited states below 41 000 $cm^{-1}$ are dominated (>95%) by determinants within the $P_m$ space and excitation energies should therefore be reliable. Above 41 000 $cm^{-1}$, contributions from determinants in the $P_i$ space start to be more significant in particular for some Π states (such as the $a\ ^1\Pi_u$ at 46126 $cm^{-1}$), making the accuracy less certain and suggesting that discrepancies with CASPT2 could also arise from the inclusion of inaccurate "undressed" states in the IHFSCC calculation, even though our model space is larger than the largest CAS (12 electrons in 16 orbitals).

Apart from comparing the excitation energies directly, it is interesting here to analyze the energy differences between

**TABLE 3: SO-IHFSCC Vertical Excitation Energies (in $cm^{-1}$) of $UO_2^{2+}$ with Different Active Spaces, Hamiltonians and Basis Sets. The Energy Range of the Selected Orbitals in the CCSD Part Is between −3 au and 20 au, Except when the 5d Shell Is Correlated then the Lower Value Is about −7 au. Changes in the Ordering of the States Are Marked in Italics**

| basis U | Dyall DZ | | | | Dyall DZ | | | Dyall TZ | Fægri TZ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| basis O | aug-cc-pVDZ | | | | aug-cc-pVTZ | | | aug-cc-pVTZ | cc-pVTZ | | |
| 5d | | frozen | in $Q$ | in $P_i$ | frozen | | | frozen | frozen | | |
| state | Ω | X2C | | | Ω | X2C | 4C-DC | 4C-DC | Ω | 4C-DC | composition (wrt. $h$–$p$ determinants) |
| 1 | $2_g$ | 18777 | 18536 | 18571 | $1_g$ | 18789 | 18984 | 18506 | $1_g$ | 18610 | 81% $\sigma_{1/2u}f^{\delta}_{3/2u}$ + 15% $\pi_{1/2u}f^{\delta}_{3/2u}$ |
| 2 | $1_g$ | 18949 | 18555 | 18591 | $2_g$ | 18871 | 19065 | 18529 | $2_g$ | 18633 | 69% $\sigma_{1/2u}f^{\phi}_{5/2u}$ + 13% $\sigma_{1/2u}f^{\delta}_{3/2u}$ + 11% $\pi_{1/2u}f^{\phi}_{5/2u}$ |
| 3 | $3_g$ | 20128 | 19726 | 19760 | $3_g$ | 20042 | 20233 | 19662 | $3_g$ | 19765 | 75% $\sigma_{1/2u}f^{\phi}_{5/2u}$ + 13% $\pi_{1/2u}f^{\phi}_{5/2u}$ |
| 4 | $2_g$ | 21411 | 21093 | 21127 | $2_g$ | 21348 | 21536 | 20987 | $2_g$ | 21080 | 45% $\sigma_{1/2u}f^{\delta}_{3/2u}$ + 22% $\sigma_{1/2u}f^{\delta}_{5/2u}$ + 16% $\sigma_{1/2u}f^{\phi}_{5/2u}$ |
| 5 | $3_g$ | 24368 | 24202 | 24233 | $3_g$ | 24310 | 24496 | 23914 | $3_g$ | 23996 | 73% $\sigma_{1/2u}f^{\delta}_{5/2u}$ + 12% $\pi_{1/2u}f^{\delta}_{5/2u}$ + 7% $\sigma_{1/2u}f^{\phi}_{7/2u}$ |
| 6 | $4_g$ | 25668 | 25388 | 25417 | $4_g$ | 25509 | 25688 | 25049 | $4_g$ | 25117 | 83% $\sigma_{1/2u}f^{\phi}_{7/2u}$ + 15% $\pi_{1/2u}f^{\phi}_{7/2u}$ |
| 7 | $3_g$ | 28784 | 28531 | 28557 | $3_g$ | 28629 | 28795 | 28066 | $3_g$ | 28132 | 73% $\sigma_{1/2u}f^{\phi}_{7/2u}$ + 10% $\pi_{1/2u}f^{\phi}_{7/2u}$ + 8% $\sigma_{1/2u}f^{\phi}_{5/2u}$ |
| 8 | $2_g$ | 30594 | 30425 | 28557 | $2_g$ | 30694 | 30861 | 30177 | $2_g$ | 30230 | 56% $\sigma_{1/2u}f^{\delta}_{5/2u}$ + 24% $\sigma_{1/2u}f^{\delta}_{3/2u}$ + 10% $\pi_{1/2u}f^{\delta}_{5/2u}$ |
| 9 | $0_g^-$ | 32121 | 32132 | 32124 | $1_g$ | 35216 | 35423 | 34544 | $1_g$ | 34556 | 95% $\pi_{3/2u}f^{\phi}_{5/2u}$ |
| 10 | $0_g^+$ | 32384 | *32391* | *32382* | $4_g$ | 36214 | 36417 | 35495 | $4_g$ | 35528 | 97% $\pi_{3/2u}f^{\phi}_{5/2u}$ |
| 11 | $1_g$ | 33222 | *33221* | *33211* | $0_g^-$ | 36415 | 36602 | *35909* | $0_g^-$ | 35840 | 92% $\pi_{3/2u}f^{\delta}_{3/2u}$ |
| 12 | $1_g$ | 34665 | 34505 | 34555 | $3_g$ | 36552 | 36744 | *35903* | $3_g$ | 35930 | 97% $\pi_{3/2u}f^{\delta}_{3/2u}$ |
| 13 | $4_g$ | 35654 | 35515 | 35564 | $2_u$ | *36784* | *36995* | 36024 | $2_u$ | 36074 | 70% $\sigma_{1/2\ g}f^{\delta}_{3/2u}$ + 25% $\sigma_{1/2\ g}f^{\phi}_{5/2u}$ |
| 14 | $3_g$ | 35916 | 35939 | 35988 | $0_g^+$ | *36693* | *36883* | 36141 | $0_g^+$ | 36115 | 92% $\pi_{3/2u}f^{\delta}_{3/2u}$ |
| 15 | $2_g$ | 36153 | 36155 | 36147 | $1_u$ | 37140 | 37344 | 36393 | $1_u$ | 36434 | 93% $\sigma_{1/2\ g}f^{\delta}_{3/2u}$ |
| 16 | $0_g^-$ | 36160 | 36163 | 36210 | $2_u$ | 37302 | 37525 | 36472 | $2_u$ | 36528 | 69% $\sigma_{1/2\ g}f^{\phi}_{5/2u}$ + 28% $\sigma_{1/2g}f^{\delta}_{3/2u}$ |
| 17 | $0_g^+$ | 36409 | 36437 | 36483 | $3_u$ | 37461 | 37686 | 36603 | $3_u$ | 36669 | 97% $\sigma_{1/2g}f^{\phi}_{5/2u}$ |

excited states by taking the first excited state ($^3\Delta_g$) as the reference level, as shown under $\Delta E^1$ in Table 2. From those numbers, it is clear that the IHFSCC and LR-CCSD results are very much alike, both in terms of energies as in the states ordering with respect to symmetry. The difference between the methods appears to mainly arise from a stronger bias for the ground state in the LR-CCSD framework that leads to a systematic overestimation of the transition energies. Less systematic errors are found in comparison with CASPT2, which yields larger discrepancies in the spacing of the different excited states (reaching a few thousand wave numbers for higher states).

**3.3. Vertical Transition Energies Including Spin−Orbit Coupling.** While instructive for a comparison with other theoretical work, SF results cannot be compared directly to experiment. In order to do so we will now switch on spin−orbit interactions. In this section we will first discuss the tests done to choose the most suitable "active space" in the SO-IHFSCC calculations. These tests are all presented in Table 3, noting that we shall restrict ourselves to excited states lower than 38 000 $cm^{-1}$ which lie well within the accurate $P_m$ model space. Apart from the inherently more difficult theoretical description of higher excited states, we also note that these states form a very dense part of the spectrum, corresponding to the strong absorption broadband that is difficult to resolve experimentally. Thus, the effort in describing them accurately is of limited use for comparison with experimental work.

***3.3.1. Influence of Model $P_m$ and Intermediate $P_i$ Spaces.*** The effect of adding the 5d shell to either the occupied $Q$ or the $P_i$ spaces was explored with the X2C Hamiltonian and turned out to be relatively unimportant (columns 3 to 5 in Table 3). Extending the $Q$ virtual space by including all orbitals with energies $\varepsilon \leq 40.00$ au did also not bring any significant changes to the results relative to the default choice (orbitals up to 20 au). This limits the correlation-active space used in subsequent comparisons to the orbitals with energies between −3 au and 20 au (AS2).

The quality of SO-IHFSCC calculations depends on the partitioning of the correlation space into the main $P_m$ and intermediate $P_i$ spaces. With respect to the creation of valence holes in the occupied orbitals it is possible to extend the minimal $P_m$ space by including the deeper of the two valence occupied $\sigma_{1/2u}$ (see Figure 1 in ref 3) in $P_m$. Placing this $\sigma_{1/2u}$ in $P_m$ instead of in $P_i$, shifts the six lowest transition energies up by 1700 $cm^{-1}$ while leaving the other transitions basically unchanged. This significant change favors the inclusion of the deeper $\sigma_{1/2u}$ in the construction of the $P_m$ space. Increasing the $P$ space in the particle space by including all orbitals with energies up to 0.35 au (as opposed to the base value of 0.08 au) has very little effect on the excitation energies but was kept as this extension did not increase calculation times much. The final $P_m$ space is constructed from orbitals with energies between −1.50 and −0.30 au, while the buffering $P_i$ space adds determinants build from holes in the energy ranges [−3.0; −1.50 au] and/or particles in the energy range [−0.30; 0.35 au].

***3.3.2. Influence of Hamiltonian and Basis Set Quality.*** Our results indicate that the X2C method complemented with the atomic mean-field approximation for spin−orbit coupling indeed provides a very good approximation to the DC Hamiltonian, with differences below 200 $cm^{-1}$ upon changing Hamiltonian. This is important in terms of extending the applicability of relativistic methods, as this two-component scheme can yield significant savings in computational time both at the SCF and the 4-index transformation step prior to the correlated calculations. In our calculations, however, computational time was mostly spent in the coupled cluster stage of the calculation, making the choice for the DC approach appropriate. We thus continue at this level and next consider the influence of the basis set choice.

Changing the description of the oxygen atoms from double-ζ quality to triple-ζ hardly affects the eight lowest excitation energies but does have a significant effect on the energies above 31 000 $cm^{-1}$, with maximal changes of 4000 $cm^{-1}$. As a

**TABLE 4: Equilibrium Geometries ($R_e$, in Å), Vertical ($\Delta E^0$, in $cm^{-1}$), and Adiabatic ($\Delta T_e^0$, in $cm^{-1}$) Spectrum of the Lowest Fine Structure Excited States of $UO_2^{2+}$, Computed at the SO-IHFSCC, SO-LR-CCSD,[7] and SO-CASPT2[6a]**

| | SO-IHFSCC[b] | | | | | SO-LR-CCSD[c] | | | | | SO-CASPT2[d] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ω | $R_e$ | $\Delta E^0$ | $\Delta E^1$ | $\Delta T_e^0$ | $\Delta T_e^1$ | $R_e$ | $\Delta E^0$ | $\Delta E^1$ | $\Delta T_e^0$ | $\Delta T_e^1$ | $R_e$ | $\Delta E^0$ | $\Delta E^1$ | $\Delta T_e^0$ | $\Delta T_e^1$ |
| $0_g^+$ | 1.683 | 0 | −18610 | 0 | −17557 | 1.679 | 0 | −22967 | 0 | −21338 | 1.708 | 0 | −20104 | 0 | −18888 |
| $1_g$ | 1.724 | 18610 | 0 | 17557 | 0 | 1.732 | 22967 | 0 | 21338 | 0 | 1.765 | *20104* | 0 | *18888* | 0 |
| $2_g$ | 1.719 | 18633 | 23 | 17834 | 277 | 1.743 | 22789 | −178 | 21826 | 488 | 1.782 | *19195* | −909 | *17227* | −1661 |
| $3_g$ | 1.725 | 19765 | 1155 | 18627 | 1070 | 1.743 | 23897 | 930 | 22361 | 1023 | 1.783 | 20265 | 161 | 18293 | −595 |
| $2_g$ | 1.722 | 21080 | 2470 | 20082 | 2525 | 1.736 | 25237 | 2270 | 24027 | 2689 | 1.769 | 22320 | 2216 | 20911 | 2023 |
| $3_g$ | 1.720 | 23996 | 5386 | 23073 | 5516 | 1.735 | 27808 | 4841 | 26723 | 5385 | 1.769 | 25435 | 5331 | 24026 | 5138 |
| $4_g$ | 1.727 | 25117 | 6507 | 23857 | 6300 | 1.743 | 29054 | 6087 | 27923 | 6585 | 1.784 | 26312 | 6208 | 24190 | 5302 |
| $3_g$ | 1.730 | 28132 | 9522 | 26679 | 9122 | 1.750 | 32382 | 9415 | 30833 | 9495 | 1.796 | 29085 | 8981 | *26446* | 7558 |
| $2_g$ | 1.731 | 30230 | 11620 | 28757 | 11200 | 1.749 | 34706 | 11739 | *32912* | 11574 | 1.848 | 31314 | 11210 | *26500* | 7612 |
| $4_g$ | 1.772 | 35528 | 16918 | 29991 | 12434 | 1.798 | *39845* | 16878 | *32113* | 10775 | 1.848 | *33262* | 13158 | *26259* | 7371 |
| $1_g$ | 1.778 | 34556 | 15946 | 30680 | 13123 | 1.795 | *39059* | 16092 | *32815* | 11477 | 1.833 | *32921* | 12817 | 27923 | 9035 |

[a] Here $\Delta E^1$ and $\Delta T_e^1$ (in $cm^{-1}$) denote vertical and adiabatic excitations where the origin of the spectrum is taken to be the first excited state. The minima of the SO-IHFSCC calculations (this work) were obtained by extrapolating the symmetrical stretching mode by second-order polynomials. Changes in the ordering of the states are marked in italics. [b] Uranium Fægri and oxygen cc-pVTZ basis sets. [c] Reference 7; uranium ANO-RCC-QZP and oxygen ANO-RCC-TZP basis sets. [d] Reference 6; uranium DK3 and oxygen ANO-L basis sets.

consequence the ordering of states in this energy region changes. Increasing the basis set quality on uranium from double to triple-ζ (Dyall or Faegri types) is again relatively unimportant for the lowest excited states (about 300−500 $cm^{-1}$), but yield effects of about 1000 $cm^{-1}$ for the higher states. These tests indicate that the triple-ζ basis set size (as used in the spin-free calculations) is indeed a minimal requirement in order to obtain accurate spectroscopic data.

***3.3.3. Composition of the Excited States.*** Having defined a satisfactory computational model (AS2/uranium with a Faegri TZ basis set/oxygen with a cc-VTZ basis set) we may now analyze the influence of spin−orbit coupling on the computed spectrum. The decomposition of the electronic states obtained in the SO-IHFSCC calculations is given in Table 2 in terms of the most significant excited determinants with respect to the ground state. From those numbers it appears that the higher excited states can be described in terms of one main excitation while the eight lowest states have some multireference character. Another interpretation is a difference in spin−orbit induced mixing in these excited states which increases the π-character of the $\sigma_{1/2u}$ orbital relative to the ground state.

The first excited state is sometimes called the luminescent state due to its long lifetime in the 20 000−26 000 $cm^{-1}$ energy range observed in various environments, e.g. in the $Cs_2UO_2Cl_4$ crystal.[68] We calculate this state to be the $1_g$ state, corresponding to an excitation from the highest $\sigma_{1/2u}$ orbital to the nonbonding $f^{\delta}_{3/2u}$. This is in agreement with spin−orbit configuration interaction results.[4,69] Pierloot et al.[7] and Réal et al.,[7] however, found the $2_g$ state being the lowest. In all calculations the energy differences between the lowest excited states, and in particular the $1_g$ and $2_g$ states, are small (less than 1500 $cm^{-1}$) and obviously influenced by the quality of the basis sets and electron correlation treatment. All the excitation from the $\sigma_{1/2u}$ to the nonbonding manifolds are located below 31000 $cm^{-1}$. Above this threshold, the spectrum is dense and almost continuous. The states corresponding to an excitation from the bonding $\pi_{1/2u}$, $\pi_{3/2u}$ orbitals to the nonbonding $f^{\delta}_{3/2u}$, $f^{\delta}_{5/2u}$ and $f^{\phi}_{5/2u}$, $f^{\phi}_{7/2u}$ orbitals are close to the strongly absorbing states, arising from the $\sigma_{1/2g}$ to $f^{\delta}_{3/2u}$, $f^{\delta}_{5/2u}$ and $f^{\phi}_{5/2u}$, $f^{\phi}_{7/2u}$ excitations.

**3.4. Excited-State Structures and Comparison between Methods.** The potential energy curves along the symmetric stretching vibrational mode relative to the ground-state geometry calculated with the SO-IHFSCC method were investigated and reported in the Table 4. They are displayed in Figure 1. For larger bond distances (>1.75 Å), convergence problems appear in the SO-IHFSCC for the $(1h,0p)$ sector (which corresponds to calculating the wave functions for ground and excited states of $UO_2^{3+}$), that could not be remedied by further extending the $P_i$ space. For the two states with an equilibrium bond length above this distance we therefore had to rely on an extrapolation of the curve using a low order polynomial.

The minima of the excited states are reported in the Table 4. All of these U−$O_{yl}$ distances are slightly longer than the ground state distance. States arising from excitations from the bonding $\sigma_{1/2u}$ orbital to the nonbonding $f^{\delta}_{3/2u}$, $f^{\delta}_{3/2u}$ and $f^{\phi}_{5/2u}$, $f^{\phi}_{7/2u}$ have their optimum bond length at about 1.73 Å, while excitations from the $\pi_{1/2u}$, $\pi_{3/2u}$ bonding orbitals (last two states reported in Table 4) yield distances of about 1.78 Å, almost 0.1 Å longer than in the ground state. In general, the SO-IHFSCC bond lengths come close to the SO-LR-CCSD values,[7] about 0.02 Å shorter. This is significantly shorter than the SO-CASPT2 bond lengths, which differ up to 0.06 Å from the SO-IHFSCC values.

Table 4 also contains vertical and adiabatic excitation energies for the three methods in question. Starting with the vertical excitations, for which the comparison between the methods should be more reliable than for the adiabatic energies (again due to the limited precision in the SO-IHFSCC equilibrium geometries), we observe that the SO-IHFSCC transition energies are consistently lower by about 4000 $cm^{-1}$ than the ones computed with SO-LR-CCSD. Like in the scalar relativistic comparison, they come closer to the SO-CASPT2 absolute values, with a difference about 500 $cm^{-1}$ for the $2_g$ state and up to 2000 $cm^{-1}$ for higher excited states.

The adiabatic transitions computed with the various methods are in line with this picture, but yielding typically one to two thousand wave numbers smaller values for the lower excited states and up to five to six thousand wave numbers for the higher states. It should be noted that the largest differences between vertical and adiabatic energies occur for states with significant Π character ($\pi_{1/2u}$, $\pi_{3/2u} \rightarrow f^{\phi}_{5/2u}$), namely the second $4_g$ and $1_g$ states, for which one would indeed expect more significant variations upon geometry changes.

The trends observed here regarding the similarities in the calculated spectra for the three methods are also seen for our spin-free results. This provides a strong indication that the inclusion of spin−orbit coupling a posteriori, as done for LR-CCSD and CASPT2 within one-component frameworks, is quite accurate in this case. As a result, the bulk of the deviations is due to the differences in the correlation treatment, with some smaller contributions from differences in Hamiltonian and/or

basis sets. This is further supported upon inspecting the values of $\Delta E^1$ in Tables 2 and 4, since these clearly show very similar results for SO-IHFSCC (SO a priori) and SO-LR-CCSD (SO a posteriori), and between the spin-free and spin−orbit calculations.

## 4. Conclusion

In this work we have investigated the performance of three wave function based correlation treatments in the calculation of excitation energies for the uranyl cation. Differences between the two approaches (perturbation theory and coupled cluster), and between different coupled cluster approaches (in particular the linear response and intermediate Hamiltonian Fock-space coupled cluster) are subtle, with comparisons showing similar results from a qualitative perspective but with clear quantitative differences.

The most remarkable quantitative difference between LRCC and IHFSCC is the systematic upward shift of excitation energies in the former, compared to the latter. This shift probably stems from the different parametrization (linear and exponential, respectively) of the wave functions for the excited states in the two coupled cluster methods leading to a larger bias for the ground state in the LRCC calculation. We expect that these discrepancies between LRCC and (IH)FSCC calculations will become smaller with the inclusion of higher excitations within the coupled cluster framework similar to the observations made in work on lighter molecules.

The often-used CASPT2 approach gives a satisfactory agreement with the lowest IHFSCC excitation energies. For higher excitations and relative spacings between excited states the agreement between the two methods larger differences are observed.

The still significant discrepancies between theoretical methods reinforces the call for experimental gas-phase spectroscopic data on the bare uranyl ion to provide a rigorous testing ground for theoretical methods. Direct comparison with experimental data obtained for uranyl crystals or solvated uranyl complexes requires to consider larger chemical models. Work is currently in progress to compute the spectrum of $Cs_2UO_2Cl_4$ using embedding methods, as recently done for the spectrum of $NpO_2^{2+}$ in the $Cs_2UO_2Cl_4$ crystal.[36]

**Acknowledgment.** We dedicate this article to Prof. R. M. Pitzer, a pioneer in the accurate simulation of actinide spectroscopy. The authors thank Dr. Ivan Infante for providing us information on the relativistic basis sets. This study was supported by a joint project (JRP 06-11) of the EC-supported ACTINET Network of Excellence. F.R. acknowledges a fellowship from the ACTINET network of excellence. Funding was provided by the CNRS and French Ministère de l'Enseignement Supérieur et de la Recherche. Moreover, the authors wish to thank The Netherlands Organization for Scientific Research (NWO) for financial support via the Vici program and acknowledge computer time provided by the Dutch National Computing Facilities (NCF). Additional computational resources have also been provided by the Institut de Développement et de Ressources en Informatique Scientifique du Centre National de la Recherche Scientifique (IDRIS-CNRS) (Contract 71859) and by the Centre Informatique National de l'Enseignement Supérieur, CINES France (Project phl2531).

## References and Notes

(1) Denning, R. G. *Struct. Bonding (Berlin)* **1992**, *79*, 215–276.
(2) *The chemistry of the actinide and transactinide elements*; Katz, J. J., Morss, L. R., Fuger, J., Edelstein, N. M., Eds.; Springer-Verlag: Berlin, 2006.
(3) Denning, R. G. *J. Phys. Chem. A* **2007**, *111*, 4125–4143.
(4) Zhang, Z.; Pitzer, R. M. *J. Phys. Chem. A* **1999**, *103*, 6880–6886.
(5) Matsika, S.; Pitzer, R. M. *J. Phys. Chem. A* **2001**, *105*, 637–645.
(6) Pierloot, K.; van Besien, E. *J. Chem. Phys.* **2005**, *123*, 204309.
(7) Réal, F.; Vallet, V.; Marian, C.; Wahlgren, U. *J. Chem. Phys.* **2007**, *127*, 214302.
(8) Wang, Q.; Pitzer, R. M. *J. Phys. Chem. A* **2001**, *105*, 8370–8375.
(9) Pierloot, K.; van Besien, E.; van Lenthe, E.; Baerends, E. J. *J. Chem. Phys.* **2007**, *126*, 194311.
(10) Saue, T.; Fægri, K.; Helgaker, T.; Gropen, O. *Mol. Phys.* **1997**, *91*, 937–950.
(11) Visscher, L.; Saue, T. *J. Chem. Phys.* **2000**, *113*, 3996–4002.
(12) Iliaš, M.; Jensen, H. J. Aa. private communication 2009.
(13) Iliaš, M.; Saue, T. *J. Chem. Phys.* **2007**, *126*, 064102.
(14) Landau, A.; Eliav, E.; Ishikawa, Y.; Kaldor, U. *J. Chem. Phys.* **2000**, *113*, 9905–9910.
(15) Landau, A.; Eliav, E.; Ishikawa, Y.; Kaldor, U. *J. Chem. Phys.* **2001**, *115*, 6862–6865.
(16) Visscher, L.; Eliav, E.; Kaldor, U. *J. Chem. Phys.* **2001**, *115*, 9720–9726.
(17) Eliav, E.; Vilkas, M. J.; Ishikawa, Y.; Kaldor, U. *J. Chem. Phys.* **2005**, *122*, 224113.
(18) Fleig, T.; Olsen, J.; Visscher, L. *J. Chem. Phys.* **2003**, *119*, 2963–2971.
(19) Fleig, T.; Jensen, H. J. Aa.; Olsen, J.; Visscher, L. *J. Chem. Phys.* **2006**, *124*, 104106.
(20) Knecht, S.; Jensen, H. J. Aa.; Fleig, T. 2009, in preparation.
(21) Knecht, S.; Jensen, H. J. Aa.; Fleig, T. *J. Chem. Phys.* **2008**, *128*, 014108.
(22) Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481–488.
(23) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
(24) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
(25) Vallet, V.; Maron, L.; Teichteil, C.; Flament, J.-P. *J. Chem. Phys.* **2000**, *113*, 1391–1402.
(26) Christiansen, O.; Hättig, C.; Gauss, J. *J. Chem. Phys.* **1998**, *109*, 4745–4757.
(27) Christiansen, O.; Gauss, J.; Stanton, J. F. *Chem. Phys. Lett.* **1998**, *292*, 437–446.
(28) Christiansen, O.; Gauss, J.; Stanton, J. F. *Chem. Phys. Lett.* **1999**, *305*, 147–155.
(29) Kobayashi, R.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **2002**, *219*, 30–35.
(30) Malmqvist, P.-Å.; Roos, B. O.; Schimmelpfennig, B. *Chem. Phys. Lett.* **2002**, *357*, 230–240.
(31) Llusar, R.; Casarrubios, M.; Barandiarán, Z.; Seijo, L. *J. Chem. Phys.* **1996**, *105*, 5321–5330.
(32) Bast, R.; Jensen, H. J. Aa.; Saue, T. *Int. J. Quantum Chem.* **2009**, *109*, 2091–2112.
(33) Görller-Walrand, C.; De Houwer, S.; Fluyt, L.; Binnemans, K. *Phys. Chem. Chem. Phys.* **2004**, *6*, 3292–3298.
(34) Infante, I.; Gomes, A. S. P.; Visscher, L. *J. Chem. Phys.* **2006**, *125*, 074301.
(35) Infante, I.; Eliav, E.; Vilkas, M. J.; Ishikawa, Y.; Kaldor, U.; Visscher, L. *J. Chem. Phys.* **2007**, *127*, 124308.
(36) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353–5362.
(37) Ruipérez, F.; Danilo, C.; Réal, F.; Flament, J.-P.; Vallet, V.; Wahlgren, U. *J. Phys. Chem. A* **2009**, *113*, 1420–1428.
(38) DIRAC, a relativistic ab initio electronic structure program, Release DIRAC08 (2008), written by Visscher, L.; Jensen, H. J. Aa.; Saue, T. with new contributions from Bast, R.; Dubillard, S.; Dyall, K. G.; Ekström, U.; Eliav, E.; Fleig, T.; Gomes, A. S. P.; Helgaker, T. U.; Henriksson, J.; Iliaš, M.; Jacob, Ch. R.; Knecht, S.; Norman, P.; Olsen, J.; Pernpointner, M.; Ruud, K.; Sałek, P.; Sikkema J. (see http://dirac.chem.sdu.dk).
(39) Dyall, K. G. *J. Chem. Phys.* **1994**, *100*, 2118–2127.
(40) Visscher, L. *Theor. Chem. Acc.* **1997**, *98*, 68–70.
(41) *Program Amfi*, Schimmelpfennig, B., Stockholm University, Sweden.
(42) He$\beta$, B. A.; Marian, C. M.; Wahlgren, U.; Gropen, O. *Chem. Phys. Lett.* **1996**, *251*, 365–371.
(43) Dyall, K. *Theor. Chem. Acc.* **2007**, *117*, 491–500.
(44) Fægri, K. *Chem. Phys.* **2005**, *311*, 25–34.
(45) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
(46) Visscher, L.; Dyall, K.; Lee, T. J. *Int. J. Quantum Chem.: Quant. Chem. Symp.* **1995**, *29*, 411–419.
(47) Visscher, L.; Lee, T. J.; Dyall, K. G. *J. Chem. Phys.* **1996**, *105*, 8769–8776.
(48) Pernpointner, M.; Visscher, L.; de Jong, W. A.; Broer, R. *J. Comput. Chem.* **2000**, *21*, 1176–1186.

(49) Pernpointner, M.; Visscher, L. *J. Comput. Chem.* **2003**, *24*, 754–759.
(50) Landau, A.; Eliav, E.; Ishikawa, Y.; Kaldor, U. *J. Chem. Phys.* **2004**, *121*, 6634–6639.
(51) Kaldor, U.; Eliav, E.; Landau, A. Study of Heavy Elements by Relativistic Fock Space and Intermediate Hamiltonian Coupled Cluster Methods. *In Fundamental World of Quantum Chemistry*; Brandas, E. J., Kryachko, E. S., Eds.; Kluwer: Dordrecht, 2004; Vol. 3, pp 365−406.
(52) DALTON, a molecular electronic structure program, Release 2.0 (2005), see http://www.kjemi.uio.no/software/dalton/dalton.html.
(53) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89–155.
(54) He$\beta$, B. A. *Phys. Rev. A* **1986**, *33*, 3742–3748.
(55) (a) de Jong, W. A.; Visscher, L.; Nieuwpoort, W. C. *J. Mol. Struct. (THEOCHEM)* **1999**, *458*, 41–52. (b) de Jong, W. A.; Visscher, L.; Nieuwpoort, W. C. *J. Mol. Struct. (THEOCHEM)* **2002**, *581*, 259–259.
(56) Straka, M.; Dyall, K. G.; Pyykkö, P. *Theor. Chem. Acc.* **2001**, *106*, 393–403.
(57) Ismail, N.; Heully, J.-L.; Saue, T.; Daudey, J.-P.; Marsden, C. J. *Chem. Phys. Lett.* **1999**, *300*, 296–302.
(58) Garciá-Hernández, M.; Lauterbach, C.; Krüger, S.; Matveev, A.; Rösch, N. *J. Comput. Chem.* **2002**, *23*, 834–846.
(59) Fromager, E.; Réal, F.; Wåhlin, P.; Jenssen, H. J. A.; Wahlgren, U. *J. Chem. Phys.* **2009**, accepted.
(60) Meissner, L.; Bartlett, R. J. *J. Chem. Phys.* **1991**, *94*, 6670−6676.
(61) Meissner, L.; Bartlett, R. J. *J. Chem. Phys.* **1995**, *102*, 7490−7498.
(62) Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352.
(63) Mukhopadhyay, D.; Mukhopadhyay, S.; Chaudhuri, R.; Mukherjee, D. *Theor. Chim. Acta* **1991**, *80*, 441–467.
(64) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic Structure Theory*; John Wiley & Sons: Chichester, 2000.
(65) Musiał, M.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *129*, 044101.
(66) Musiał, M.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *129*, 134105.
(67) Musiał, M.; Bartlett, R. J. *Chem. Phys. Lett.* **2008**, *457*, 267−270.
(68) Denning, R. G.; Snellgrove, T. R.; Woodwark, D. R. *Mol. Phys.* **1976**, *32*, 419–442.
(69) Matsika, S.; Zhang, Z.; Brozell, S. R.; Blaudeau, J.-P.; Wang, Q.; Pitzer, R. M. *J. Phys. Chem. A* **2001**, *105*, 3825–3838.

JP903758C

## C.4 Paper IV

View Article Online / Journal Homepage / Table of Contents for this issue
Dynamic Article Links

PCCP

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 6249–6259

www.rsc.org/pccp

**PAPER**

# Electronic spectroscopy of $UO_2^{2+}$, $NUO^+$ and NUN: an evaluation of time-dependent density functional theory for actinides

**Paweł Tecmer,[a] André Severo Pereira Gomes,[b] Ulf Ekström[a] and Lucas Visscher*[a]**

***Received 15th November 2010, Accepted 1st February 2011***
**DOI: 10.1039/c0cp02534h**



The performance of the time-dependent density functional theory (TDDFT) approach has been evaluated for the electronic spectrum of the $UO_2^{2+}$, $NUO^+$ and NUN molecules. Different exchange–correlation functionals (LDA, PBE, BLYP, B3LYP, PBE0, M06, M06-L, M06-2X, CAM-B3LYP) and the SAOP model potential have been investigated, as has the relative importance of the adiabatic local density approximation (ALDA) to the exchange-correlation kernel. The vertical excitation energies have been compared with reference data obtained using accurate wave-function theory (WFT) methods.

## 1. Introduction

The importance of theoretical modeling in furthering the understanding of the chemistry of actinide-containing systems, compared to other branches of chemistry, is by now well established. This prominent role has to do with the experimental difficulties involved in actinide research: besides the acute radiotoxicity of most species, which places severe restrictions on laboratory manipulations, the wide range of oxidation states possible for early actinides (U–Am), together with a relative ease of changing their oxidation states often makes it difficult to isolate and characterize species.

One of the most active areas of chemical research on actinides is related to the reprocessing of nuclear waste, whose objective is to separate uranium and plutonium from the other (minor) actinides. Such separation has implications both to the recycling of irradiated nuclear fuels (by allowing the retrieval of important quantities of U and Pu from spent fuel and their subsequent reconversion back to usable fuel) and to the disposal of nuclear waste (as it decreases the volume of material to be stored). While separation methods based upon liquid–liquid extraction ion-exchange, such as the plutonium uranium extraction (PUREX)[1,2] or transuranic extraction (TRUEX)[3] processes are rather well established, the details of the interaction of the extraction ligand with the actinide species (such as the simple atomic ions $Ac^{n+}$ or the frequently found actinyl ($AcO_2^{n+}$) species) are still far from fully understood.

[a] *Amsterdam Center for Multiscale Modeling, Department of Theoretical Chemistry, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, Netherlands. E-mail: visscher@few.vu.nl*
[b] *Université de Lille 1 (Sciences et Technologies), Laboratoire PhLAM, CNRS UMR 8523, CNRS FR 2416, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France*

Bridging this gap in understanding would be particularly helpful in designing more efficient and selective complexing agents. Modeling this process is challenging, because it requires an accurate description of the actinide and the complexing species, as well as their interactions with the solvent environment. This is currently only possible with Density-Functional Theory (DFT), as its relatively modest computational cost makes the study of structure and energetics of relatively large model complexes possible, even in the condensed phase.[4–12]

Notwithstanding this progress, the success of DFT still depends on careful parametrization and benchmarking studies that establish the reliability of exchange-correlation functionals for a particular application. This is particularly serious for molecules containing heavy elements since such systems were usually not accounted for in the parametrization and validation stage of the currently available density functionals. One particular reason for concern is the strong (static and dynamic) electron correlation effects in actinides. The 5f, 7s, 6p and 6d orbitals should all be considered valence orbitals that can contribute to the chemical bonding. While energetically close, these orbitals have a rather different spatial character making the description of the exchange–correlation interaction by a density functional more difficult than for lighter elements. These difficulties have been investigated for uranium oxides,[13] showing that DFT, using proper functionals, is typically suitable for geometry optimization and thermochemistry of the electronic ground state for these systems. Based on a series of studies of solvated uranium fluorides and oxofluorides Schreckenbach and Shamov[14] conclude that GGA functionals yield accurate geometries and frequencies while hybrid density functional theory (DFT) functionals are found to be superior for the thermochemistry.

The question of suitability of DFT for studying actinide-containing molecules carries over to its time-dependent generalization (TDDFT), which is is used to calculate properties

depending upon the response of the density to an external perturbation. These calculations allow one to predict and analyze electronic spectra, polarizabilities or magnetizabilities, and vibrational Raman spectra, all of which are useful tools in studying the interactions of the actinide system with the complexing agents, or of that complex with the environment. While TDDFT has been shown to work rather well for some transition metal excited states,[15–17] currently there are still only relatively few studies[18–20] assessing the reliability of TDDFT for actinide-containing molecules.

The aim of this paper is, therefore, to explore different flavors of TDDFT for the calculation of electronic spectra for such systems—that is, evaluating (meta-)GGAs and (meta-)hybrid functionals within the adiabatic approximation. We will also study the further approximation of replacing the derivatives of the functionals in the exchange-correlation (XC) kernel by the simpler LDA approach (ALDA) that is often used to simplify the implementation of TDDFT.

Our initial focus will be on small molecules, namely $UO_2^{2+}$, $NUO^+$ and NUN, since: (a) they are closed-shell systems in their ground-states, thus avoiding problems with the validity of the reference state for TDDFT; (b) they are all isoelectronic, making it instructive to see how changes in, for instance, electronegativity of the ligands affect the spectra; (c) the –N=U=N– and O=U=N– groups appear in larger organometallic systems, so they can serve as stepping stones for description of their oxo (MO), imido (MNR) and phoshorane iminato ($MNPR_3$)[21–24] analogs, which are very important in nuclear waste separation, and (d) in contrast to $UO_2^{2+}$, which has received extensive attention from theoreticians,[18,20,25–30] the electronic spectra of the isoelectronic species ($NUO^+$ and NUN)[31–33] have not yet been investigated in detail.

The lack of experimental data requires that our assessment be done by comparison with accurate wavefunction-based (WFT) calculations, namely the complete active space second-order perturbation theory (CASPT2)[34,35] and intermediate Hamiltonian Fock-space coupled cluster (IHFSCC)[36–38] methods. Our aim in using both in tandem is to provide a cross-validation of WFT results, and to get access to the rich set of analysis tools available for CASPT2 (in our case population analysis of the excited states).

CASPT2 and IHFSCC are examples of multi-reference approaches to the electron correlation problem, and are known to perform well for actinide and other heavy element-containing molecules since such systems often present (nearly-)degenerate electronic states[19,39–41] (*e.g.* half-filled f shells, *etc.*). Unsurprisingly, they also show a very good performance in cases where the reference wavefunction is still dominated by a single determinant.[27,42]

The IHFSCC method is particularly interesting here because of its ability to yield a number of electronic states of the molecule (and those due to electron attachment or detachment) in a single calculation. Furthermore, as all states have a common Fermi vaccum, it has as advantage the elimination of a potential bias towards a given reference state. The Intermediate Hamiltonian ansatz thereby circumvents a well-known drawback of the Fock Space (FS) (and other multireference coupled cluster) approaches by largely eliminating the intruder states that could otherwise obstruct the convergence of the FSCC algorithm. As it is beyond the scope of this work to discuss in depth IHFSCC and other coupled-cluster approaches for the calculation of electronic spectra, we refer the reader to recent reviews.[43–45]

We note that spin–obit (SO) effects are small in the ground states of these molecules, with scalar relativistic and 4-component relativistic methods agreeing to within 1 pm on the bond distance of $UO_2^{2+}$.[27] The effects on the electronic spectra are more important but do not affect the comparison between different correlation methods that is the subject of the current paper. We will therefore focus exclusively on spin-free calculations to simplify the discussions, and refer to a previous paper by two of us for a more detailed discussion of the SO-CASP2 and SO-IHFSCC results for $UO_2^{2+}$.[27] We shall address spin–orbit effects on the spectra of NUN and $NUO^+$ in a subsequent publication.

## 2. Computational details

All calculations were performed with spin-free relativistic methods using ADF2008 (and ADF2009[46–48] for the (meta-)-hybrid functionals), as well as with a development version of the DIRAC08[49] program. To facilitate comparisons with the TDDFT calculations of van Besien and Pierloot[18] we used for $UO_2^{2+}$ ions the same geometry (a U–O bond distance of 1.708 Å). The geometries of NUN and $NUO^+$ were optimized with the PBE exchange-correlation functional, the ADF TZ2P basis set and the all-electron scalar relativistic ZORA (Zero Order Regular Approximation)[50] Hamiltonian. The U–N distance in the NUN molecule was determined to be 1.739 Å, whereas for $NUO^+$ it was found to be 1.698 Å. The U–O distance in $NUO^+$ is calculated to be 1.761 Å.

### 2.1 TDDFT

In the TDDFT calculations we applied the adiabatic approximation, where the frequency-dependent exchange-correlation kernel has been replaced by the local (in time) functional derivatives of the frequency-independent functional. In ADF the ALDA approximation is used for all XC functionals, whereas for DIRAC[20] we used the full derivatives of the functionals (obtained *via* the XCFun DFT library)[51] in the XC kernel in addition to the ALDA approach. We note that, in the case of hybrid functionals, in both ADF and DIRAC the fractional Hartree–Fock exchange is always included in the TDDFT kernel.

For both ADF and DIRAC, we have evaluated the following functionals: LDA,[52] PBE,[53] BLYP,[54–56] B3LYP,[57] PBE0[58] as well as the SAOP model potential.[59,60] Additionally, the functionals M06, M06-L, M06-2X[61,62] were evaluated in the ALDA approximation using ADF, whereas CAM-B3LYP[63] was evaluated in DIRAC for both the ALDA and full (non-ALDA) TDDFT kernels.

In all calculations with ADF TZ2P basis sets (U:15s13p9d5f; N,O:5s3p1d1f)[64] were used, whereas for DIRAC we used the triple zeta basis set of Dyall[65] for the uranium atom (33s29p20d13f5g2h), and the uncontracted aug-cc-pVTZ basis sets for oxygen and nitrogen (11s6p3d2f).[66,67] We have used the scalar ZORA[50] Hamiltonian in ADF, and the spin-free Dirac-Coulomb[68] (DC) Hamiltonian in DIRAC. In the DC

case the (*SS*|*SS*) integrals have been approximated by a point charge model.[69]

Our work focused on five low-lying vertical excitations determined by DFT calculations. For $UO_2^{2+}$ we studied transitions mainly consisting of excitations $3\sigma_u \rightarrow 1\phi_u$, $3\sigma_u \rightarrow 1\delta_u$, $2\pi_u \rightarrow 1\phi_u$, $3\sigma_g \rightarrow 1\phi_u$ and $3\sigma_g \rightarrow 1\delta_u$. For $NUO^+$ molecule the following dominant excitations were studied: $6\sigma \rightarrow 1\phi$, $6\sigma \rightarrow 1\delta$, $3\pi \rightarrow 1\phi$, $3\pi \rightarrow 1\delta$ and $5\sigma \rightarrow 1\phi$, while for NUN the dominant excitations are $3\sigma_u \rightarrow 1\phi_u$, $3\sigma_u \rightarrow 1\delta_u$, $3\sigma_g \rightarrow 1\phi_u$, $2\pi_u \rightarrow 1\phi_u$ and $3\sigma_u \rightarrow 4\sigma_g$.

### 2.2 CASPT2

CASPT2 calculations were carried out with the MOLCAS 7.4[70] program. For all CASPT2(MS-CASPT2)[34,35,71] calculations we utilized the scalar second-order Douglas–Kroll–Hess (DKH2) Hamiltonian,[72,73] together with the (contracted) ANO-RCC basis sets, optimized specifically for this Hamiltonian: (26s23p17d13f5g3h) → [10s9p7d5f3g2h] for uranium, (14s9p4d3f2g) → [5s4p3d2f] for oxygen and (14s9p4d3f2g) → [4s3p2d1f] for nitrogen.[74]

The most important point in the CASSCF calculation is the proper choice of active space. In the molecules investigated, the U–O and U–N bonds are formed out of the 6p, 7s, 5f, 6d orbitals of the uranium atom and the 2s and 2p orbitals of the oxygen and nitrogen atoms. While it would be ideal to take into account all molecular orbitals that are formed out of these atomic uranium, oxygen and nitrogen orbitals and correlate them in the CASPT2(MS-CASPT2) level, such an active space becomes too large to handle and we are forced to truncate the active space. It was possible to enlarge the active space for the $UO_2^{2+}$ compound in comparison to the previous work by Réal *et al.*,[27] and we took into active space 12 electrons and 16 orbitals—CAS(12,16): $3\sigma_g$, $1\pi_g$, $2\pi_u$, $3\sigma_u$, $1\delta_u$, $1\phi_u$, $3\pi_u$, $4\sigma_g$, $4\sigma_u$ and $2\pi_g$. For $NUO^+$ and NUN molecules we also correlated 12 electrons and 16 orbitals—CAS(12,16), that is $5\sigma$, $2\pi$, $3\pi$, $6\sigma$, $1\delta$, $1\phi$, $7\sigma$, $2\delta$, $4\pi$ and $8\sigma$ in the case of $NUO^+$ and $3\sigma_g$, $1\pi_g$, $2\pi_u$, $3\sigma_u$, $1\delta_u$, $1\phi_u$, $4\sigma_g$, $3\pi_u$, $1\delta_g$ and $4\sigma_u$ in the case of NUN. Due to technical problems encountered with MOLCAS 7.4 it was not possible to obtain the CAS(12,16) results for some of the higher-lying states of $NUO^+$, in order to obtain these energies we also employed a CAS(12,15) space in which the $7\sigma$ orbital was not taken in the active space.

In order to eliminate weakly intruding states in the second-order perturbation theory, we used the imaginary shift method[75] with a shift parameter of 0.25 Hartree when exploratory calculations indicated problems with intruder states.

### 2.3 IHFSCC

Intermediate Hamiltonian Fock space coupled cluster (IHFSCC)[36–38] calculations were performed with a development version of DIRAC08.[49] In those, the spin-free[68] DC Hamiltonian was used with the same uncontracted basis sets as described above for the TDDFT Dirac calculations.

In order to be consistent with our previous calculations on $UO_2^{2+}$, and due to the fact that in their ground state these molecules are well described by a single determinant, we have essentially followed the procedure outlined in ref. 27 (with the difference that we used 1.708 Å as the U–O bond length); that is, we have utilized the "one particle, one hole" sector (1h,1p) of Fock space to obtain the excitation energies, and have included in the correlated calculations orbitals with energies (in a.u.) $\varepsilon \in [-3.00;20.00]$, which correspond to 12 occupied and 253 (252) virtual orbitals for NUN ($NUO^+$).

In Fock space calculations it is necessary to subdivide the space spanned by the active orbitals in two subspaces: the model or *P* space, containing the active valence orbitals which are directly involved in the electronic excitations and the complementary *Q* space that includes the remaining "correlation-active" orbitals. As we are employing a formulation based on an intermediate Hamiltonian,[76] the *P* space is further divided into a main model ($P_m$) space and an intermediate model ($P_i$) space that is not dressed and serves as a buffer between the $P_m$ and *Q* spaces, in order to alleviate problems with the so-called intruder states (further details can be found in ref. 27 and references therein). One must keep in mind, however, that accurate solutions are only obtained for states dominated by $P_m$ components.

Thus, in the definition of the Fock space used here, *P* contains all the occupied plus the 63 lowest-lying virtuals (for NUN, 29 virtuals are contained in gerade irreducible representations, and 34 in ungerade ones), whereas the remaining virtuals have been assigned to the *Q* space. As for the partition into $P_m$ and $P_i$ orbital spaces we have, for the occupied orbitals, the five innermost orbitals in $P_i$—which correspond to $\{1\sigma_g,1\sigma_u,1\pi_u,2\sigma_g\}$ for NUN and $\{1\sigma,2\sigma,1\pi,3\sigma\}$ for $NUO^+$—and the remaining seven orbitals in $P_m$—namely $\{3\sigma_g,2\sigma_u,3\sigma_u,1\pi_g,2\pi_u\}$ for NUN and $\{4\sigma,5\sigma,6\sigma,2\pi,3\pi\}$ for $NUO^+$.

For the virtuals the $P_m$ active spaces correspond to 31 (22) orbitals for NUN ($NUO^+$), assuring that the resulting lowest-lying excited states are dominated by $P_m$ components. These orbitals correspond roughly to about two to three lowest-lying $\phi(=\phi_u)$ and $\delta(=\delta_u,\delta_g)$ orbitals of uranium, apart from the same number of $\pi$ and twice as many $\sigma$ orbitals. The difference between these sets is due to the difference in charge for both systems and the nature of the Hartree–Fock virtuals; contrary of what is obtained with DFT, the $\phi$,$\delta$ orbitals of uranium are not the lowest-lying ones—a number of $\pi$,$\sigma$ orbitals lie below and in between the former and had to be included in the model space as well.

## 3. Wavefunction benchmark calculations

As the main goal of this work is assessing the performance of TDDFT by comparison to benchmark WFT values we will begin with a brief discussion of the two WFT methods employed here, before having a closer look at the actual DFT and TDDFT results.

### 3.1 Electronic structure from IHFSCC

We refer the reader to the paper of Réal and coworkers[27] for a detailed discussion of the $UO_2^{2+}$ IHFSCC calculations. These calculations were done at a different bond length, leading to slightly different numbers in Table 1, compared to those previously reported (see Table 2 of the aforementioned paper), but analysis of the wavefunctions at both geometries yields essentially the same picture. For $UO_2^{2+}$ the lowest $\Phi_g$ and $\Delta_g$

**Table 1** Comparison of different XC functionals for $UO_2^{2+}$ (in eV). S–T indicates singlet–triplet splitting, MUE the mean absolute error and Max the maximum absolute error, respectively

| XC/Hamiltonian | $1^3\Phi_g$ | $1^1\Phi_g$ | S-T | $1^3\Delta_g$ | $1^1\Delta_g$ | S-T | $1^3\Gamma_g$ | $1^1\Gamma_g$ | S-T | $1^3\Phi_u$ | $1^1\Phi_u$ | S-T | $1^3\Delta_u$ | $1^1\Delta_u$ | S-T | MUE | Max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ZORA/LDA | 2.02 | 2.46 | 0.44 | 2.36 | 3.07 | 0.71 | 3.37 | 3.57 | 0.20 | 2.99 | 3.07 | 0.08 | 3.47 | 3.48 | 0.01 | 1.04 | 1.70 |
| DC/LDA | 2.04 | 2.54 | 0.50 | 2.37 | 3.22 | 0.85 | 3.38 | 3.66 | 0.28 | 3.00 | 3.08 | 0.08 | 3.47 | 3.50 | 0.03 | 1.00 | 1.69 |
| ZORA/PBE(ALDA) | 2.01 | 2.45 | 0.44 | 2.40 | 3.10 | 0.70 | 3.38 | 3.58 | 0.20 | 3.04 | 3.12 | 0.08 | 3.57 | 3.59 | 0.02 | 1.00 | 1.65 |
| DC/PBE(ALDA) | 2.03 | 2.53 | 0.50 | 2.41 | 3.25 | 0.84 | 3.38 | 3.65 | 0.27 | 3.04 | 3.12 | 0.08 | 3.55 | 3.58 | 0.03 | 0.97 | 1.65 |
| DC/PBE | 1.73 | 2.46 | 0.73 | 2.05 | 3.16 | 1.11 | 3.20 | 3.59 | 0.39 | 2.99 | 3.11 | 0.12 | 3.51 | 3.57 | 0.06 | 1.09 | 1.70 |
| ZORA/BLYP(ALDA) | 2.01 | 2.46 | 0.45 | 2.43 | 3.11 | 0.66 | 3.31 | 3.52 | 0.21 | 3.00 | 3.08 | 0.08 | 3.55 | 3.56 | 0.01 | 1.02 | 1.69 |
| DC/BLYP(ALDA) | 2.05 | 2.55 | 0.50 | 2.45 | 3.26 | 0.81 | 3.31 | 3.58 | 0.27 | 3.00 | 3.09 | 0.09 | 3.53 | 3.56 | 0.03 | 0.99 | 1.69 |
| DC/BLYP | 1.84 | 2.45 | 0.61 | 2.22 | 3.13 | 0.91 | 3.19 | 3.50 | 0.31 | 2.98 | 3.08 | 0.10 | 3.50 | 3.55 | 0.05 | 1.08 | 1.71 |
| ZORA/M06-L(ALDA) | 2.38 | 2.84 | 0.46 | 2.71 | 3.46 | 0.75 | 3.77 | 3.97 | 0.20 | 3.60 | 3.68 | 0.08 | 4.08 | 4.08 | 0.00 | 0.61 | 1.09 |
| ZORA/B3LYP(XALDA) | 2.30 | 2.80 | 0.51 | 2.51 | 3.39 | 0.88 | 4.01 | 4.23 | 0.22 | 4.22 | 4.31 | 0.09 | 4.57 | 4.58 | 0.01 | 0.34 | 0.56 |
| DC/B3LYP(XALDA) | 2.37 | 2.92 | 0.55 | 2.58 | 3.59 | 1.01 | 4.04 | 4.31 | 0.27 | 4.23 | 4.33 | 0.10 | 4.56 | 4.61 | 0.05 | 0.29 | 0.53 |
| DC/B3LYP | 2.21 | 2.84 | 0.63 | 2.40 | 3.49 | 1.08 | 3.95 | 4.26 | 0.31 | 4.20 | 4.32 | 0.12 | 4.53 | 4.59 | 0.06 | 0.35 | 0.62 |
| ZORA/PBE0(XALDA) | 2.35 | 2.88 | 0.53 | 2.50 | 3.43 | 0.94 | 4.25 | 4.48 | 0.23 | 4.56 | 4.66 | 0.10 | 4.84 | 4.85 | 0.01 | 0.19 | 0.36 |
| DC/PBE0(XALDA) | 2.42 | 2.99 | 0.57 | 2.58 | 3.63 | 1.05 | 4.29 | 4.57 | 0.28 | 4.57 | 4.68 | 0.11 | 4.84 | 4.89 | 0.05 | 0.16 | 0.28 |
| DC/PBE0 | 2.15 | 2.94 | 0.79 | 2.24 | 3.57 | 1.33 | 4.15 | 4.54 | 0.39 | 4.53 | 4.67 | 0.14 | 4.79 | 4.88 | 0.09 | 0.22 | 0.55 |
| ZORA/M06(XALDA) | 2.42 | 2.97 | 0.54 | 2.52 | 3.50 | 0.98 | 4.26 | 4.49 | 0.23 | 4.60 | 4.70 | 0.10 | 4.84 | 4.85 | 0.01 | 0.16 | 0.31 |
| ZORA/M06-2X(XALDA) | 2.34 | 2.94 | 0.60 | 2.23 | 3.35 | 1.12 | 4.94 | 5.21 | 0.27 | 5.63 | 5.74 | 0.11 | 5.57 | 5.62 | 0.05 | 0.56 | 1.00 |
| ZORA/SAOP(ALDA) | 3.07 | 3.51 | 0.44 | 3.27 | 4.00 | 0.73 | 4.29 | 4.48 | 0.19 | 4.21 | 4.29 | 0.08 | 4.55 | 4.56 | 0.01 | 0.37 | 0.79 |
| DC/SAOP(ALDA) | 3.01 | 3.50 | 0.49 | 3.24 | 4.10 | 0.86 | 4.33 | 4.57 | 0.24 | 4.22 | 4.31 | 0.09 | 4.59 | 4.63 | 0.04 | 0.35 | 0.76 |
| DC/CAM-B3LYP(XALDA) | 2.56 | 3.13 | 0.57 | 2.71 | 3.77 | 1.06 | 4.41 | 4.69 | 0.26 | 4.70 | 4.81 | 0.11 | 4.93 | 5.00 | 0.07 | 0.15 | 0.29 |
| DC/CAM-B3LYP | 2.43 | 3.07 | 0.64 | 2.56 | 3.69 | 1.13 | 4.36 | 4.65 | 0.29 | 4.69 | 4.81 | 0.12 | 4.91 | 4.99 | 0.08 | 0.15 | 0.28 |
| ZORA/TDHF | 3.01 | 3.78 | 0.77 | 2.40 | 4.00 | 1.60 | 7.12 | 7.17 | 0.05 | 8.82 | 9.00 | 0.18 | 8.15 | 8.54 | 0.39 | 2.19 | 4.26 |
| CASPT2 | 2.91 | 3.40 | 0.49 | 2.77 | 3.88 | 1.11 | 4.61 | 4.83 | 0.22 | 4.82 | 4.85 | 0.03 | 4.72 | 4.64 | −0.08 | 0.16 | 0.31 |
| CASPT2[a] | 2.94 | 3.47 | 0.57 | 2.79 | 3.90 | 1.11 | 4.66 | 4.86 | 0.20 | 4.71 | 4.74 | 0.03 | 4.63 | 4.55 | −0.08 | 0.16 | 0.33 |
| IHFSCC | 2.70 | 3.24 | 0.54 | 2.48 | 3.57 | 1.09 | 4.57 | 4.78 | 0.21 | 4.69 | 4.74 | 0.05 | 4.76 | 4.71 | −0.05 | | |

[a] Ref. 25.

**Table 2** Comparison of different XC functionals for $NUO^+$ (in eV). S–T indicates singlet–triplet splitting, MUE the mean absolute error and Max the maximum absolute error, respectively

| XC/Hamiltonian | $1^3\Phi$ | $1^1\Phi$ | S–T | $1^3\Delta$ | $1^1\Delta$ | S–T | $1^3\Gamma$ | $1^1\Gamma$ | S–T | $1^3\Pi$ | $1^1\Pi$ | S–T | $2^3\Phi$ | $2^1\Phi$ | S–T | MUE | Max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ZORA/LDA | 0.77 | 1.09 | 0.32 | 1.34 | 1.73 | 0.39 | 2.00 | 2.14 | 0.14 | 2.61 | 2.98 | 0.37 | 2.61 | 2.65 | 0.04 | 0.80 | 1.31 |
| DC/LDA | 0.86 | 1.20 | 0.34 | 1.40 | 1.88 | 0.48 | 2.08 | 2.29 | 0.21 | 2.67 | 3.13 | 0.46 | 2.64 | 2.84 | 0.20 | 0.69 | 1.23 |
| ZORA/PBE(ALDA) | 0.85 | 1.17 | 0.32 | 1.45 | 1.86 | 0.41 | 2.13 | 2.27 | 0.14 | 2.79 | 3.16 | 0.37 | 2.79 | 2.82 | 0.03 | 0.66 | 1.18 |
| DC/PBE(ALDA) | 0.84 | 1.20 | 0.36 | 1.41 | 1.92 | 0.51 | 2.07 | 2.29 | 0.22 | 2.71 | 3.10 | 0.39 | 2.71 | 2.88 | 0.17 | 0.68 | 1.24 |
| DC/PBE | 0.52 | 1.15 | 0.63 | 1.11 | 1.87 | 0.76 | 1.91 | 2.24 | 0.33 | 2.59 | 3.07 | 0.48 | 2.54 | 2.82 | 0.28 | 0.81 | 1.40 |
| ZORA/BLYP(ALDA) | 0.87 | 1.19 | 0.32 | 1.50 | 1.90 | 0.40 | 2.06 | 2.21 | 0.15 | 2.75 | 3.12 | 0.37 | 2.75 | 2.78 | 0.03 | 0.68 | 1.25 |
| DC/BLYP(ALDA) | 0.87 | 1.23 | 0.36 | 1.47 | 1.96 | 0.49 | 2.01 | 2.23 | 0.22 | 2.66 | 3.05 | 0.39 | 2.66 | 2.84 | 0.18 | 0.69 | 1.30 |
| DC/BLYP | 0.67 | 1.17 | 0.50 | 1.28 | 1.88 | 0.60 | 1.89 | 2.15 | 0.26 | 2.59 | 3.01 | 0.42 | 2.54 | 2.74 | 0.20 | 0.80 | 1.42 |
| ZORA/M06-L(ALDA) | 1.13 | 1.48 | 0.35 | 1.68 | 2.15 | 0.47 | 2.55 | 2.70 | 0.15 | 3.18 | 3.54 | 0.36 | 3.18 | 3.21 | 0.03 | 0.34 | 0.76 |
| ZORA/B3LYP(XALDA) | 1.25 | 1.68 | 0.43 | 1.67 | 2.29 | 0.62 | 2.66 | 2.82 | 0.16 | 3.17 | 3.54 | 0.37 | 3.17 | 3.22 | 0.05 | 0.28 | 0.65 |
| DC/B3LYP(XALDA) | 1.29 | 1.75 | 0.46 | 1.68 | 2.40 | 0.72 | 2.62 | 2.85 | 0.23 | 3.11 | 3.51 | 0.40 | 3.12 | 3.29 | 0.17 | 0.26 | 0.69 |
| DC/B3LYP | 1.13 | 1.68 | 0.55 | 1.52 | 2.33 | 0.81 | 2.56 | 2.80 | 0.24 | 3.07 | 3.48 | 0.41 | 3.03 | 3.23 | 0.20 | 0.31 | 0.75 |
| ZORA/PBE0(XALDA) | 1.31 | 1.76 | 0.45 | 1.65 | 2.34 | 0.69 | 2.87 | 3.05 | 0.18 | 3.32 | 3.70 | 0.38 | 3.32 | 3.37 | 0.05 | 0.18 | 0.44 |
| DC/PBE0(XALDA) | 1.34 | 1.83 | 0.49 | 1.67 | 2.46 | 0.79 | 2.84 | 3.08 | 0.24 | 3.28 | 3.68 | 0.40 | 3.28 | 3.46 | 0.18 | 0.19 | 0.47 |
| DC/PBE0 | 1.06 | 1.78 | 0.68 | 1.35 | 2.41 | 1.06 | 2.70 | 3.05 | 0.35 | 3.18 | 3.66 | 0.48 | 3.14 | 3.42 | 0.28 | 0.27 | 0.61 |
| ZORA/M06(XALDA) | 1.33 | 1.79 | 0.46 | 1.64 | 2.36 | 0.72 | 2.92 | 3.09 | 0.17 | 3.35 | 3.72 | 0.37 | 3.35 | 3.40 | 0.05 | 0.16 | 0.39 |
| ZORA/M06-2X(XALDA) | 1.32 | 1.86 | 0.54 | 1.34 | 2.28 | 0.94 | 3.39 | 3.60 | 0.21 | 3.54 | 3.91 | 0.37 | 3.57 | 3.63 | 0.06 | 0.21 | 0.41 |
| ZORA/SAOP(ALDA) | 1.93 | 2.26 | 0.33 | 2.34 | 2.78 | 0.44 | 3.02 | 3.16 | 0.14 | 3.50 | 3.84 | 0.34 | 3.50 | 3.54 | 0.04 | 0.31 | 0.78 |
| DC/SAOP(ALDA) | 1.84 | 2.21 | 0.37 | 2.28 | 2.82 | 0.54 | 2.96 | 3.16 | 0.20 | 3.47 | 3.83 | 0.36 | 3.47 | 3.63 | 0.16 | 0.30 | 0.72 |
| DC/CAM-B3LYP(XALDA) | 1.50 | 1.99 | 0.49 | 1.83 | 2.63 | 0.80 | 2.91 | 3.16 | 0.25 | 3.35 | 3.76 | 0.41 | 3.35 | 3.53 | 0.18 | 0.18 | 0.40 |
| DC/CAM-B3LYP | 1.37 | 1.94 | 0.57 | 1.69 | 2.56 | 0.87 | 2.85 | 3.12 | 0.27 | 3.31 | 3.73 | 0.42 | 3.28 | 3.48 | 0.20 | 0.19 | 0.46 |
| ZORA/TDHF | 2.05 | 2.79 | 0.74 | 4.73 | 5.15 | 0.42 | 5.51 | 5.76 | 0.25 | 6.09 | 6.33 | 0.24 | 5.93 | 6.15 | 0.22 | 2.26 | 3.17 |
| CASPT2 | 1.88 | 2.23 | 0.35 | 1.84 | 2.42 | 0.58 | | 3.19 | | | 3.35 | | | | | 0.06 | 0.29 |
| CASPT2(12,15) | 1.89 | 2.32 | 0.43 | 1.90 | 2.55 | 0.65 | 3.18 | 3.26 | 0.08 | 3.20 | 3.39 | 0.19 | | | | 0.06 | 0.34 |
| IHFSCC | 1.59 | 2.06 | 0.47 | 1.56 | 2.37 | 0.81 | 3.31 | 3.45 | 0.14 | 3.34 | 3.50 | 0.16 | 3.36 | 3.38 | 0.02 | | |

states are dominated by excitations of the $\sigma_u \rightarrow \phi_u$ and $\sigma_u \rightarrow \delta_u$ kind, respectively, while the $\Gamma_g$ states arises from predominantly $\pi_u \rightarrow \phi_u$ excitation. The $\Phi_u$ and $\Delta_u$ states correspond to exciations from $\sigma_g \rightarrow \phi_u$ and $\sigma_g \rightarrow \delta_u$, respectively.

For NUN, the $\Phi_g$ and $\Phi_u$ states in Table 3 differ essentially in the occupied $\sigma$ orbital ($\sigma_u$ for the first and $\sigma_g$ for the second). These two $\sigma$ orbitals, in turn, differ only in the degree of contributions from the orbitals centered on N and U, with $\sigma_u$ having more U character and the $\sigma_g$ having more N character. And, since the $\phi_u$ orbital is essentially a pure f from U, we can argue that the two transitions have different degrees of charge-transfer character. The other higher-lying excitations are dominated by $\pi_u \rightarrow \phi_u$ (for $\Gamma_g$, where $\pi_u$ has

**Table 3** Comparison of different XC functionals for NUN (in eV). S–T indicates singlet–triplet splitting, MUE the mean absolute error and Max the maximum absolute error, respectively

| XC/Hamiltonian | $1^3\Phi_g$ | $1^1\Phi_g$ | S–T | $1^3\Delta_g$ | $1^1\Delta_g$ | S–T | $1^3\Phi_u$ | $1^1\Phi_u$ | S–T | $1^3\Gamma_g$ | $1^1\Gamma_g$ | S–T | $1^3\Sigma_u$ | $1^1\Sigma_u$ | S–T | MUE | Max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ZORA/LDA | 1.07 | 1.51 | 0.44 | 1.64 | 2.24 | 0.58 | 1.88 | 1.97 | 0.09 | 2.26 | 2.47 | 0.21 | 2.34 | 2.60 | 0.26 | 0.78 | 1.55 |
| DC/LDA | 1.13 | 1.63 | 0.50 | 1.67 | 2.41 | 0.74 | 1.95 | 2.06 | 0.11 | 2.32 | 2.60 | 0.28 | 2.21 | 2.48 | 0.27 | 0.74 | 1.46 |
| ZORA/PBE(ALDA) | 1.05 | 1.49 | 0.44 | 1.68 | 2.26 | 0.58 | 1.90 | 2.00 | 0.10 | 2.28 | 2.49 | 0.21 | 2.35 | 2.60 | 0.25 | 0.77 | 1.52 |
| DC/PBE(ALDA) | 1.12 | 1.62 | 0.50 | 1.70 | 2.43 | 0.73 | 1.96 | 2.07 | 0.11 | 2.32 | 2.60 | 0.28 | 2.23 | 2.49 | 0.26 | 0.74 | 1.45 |
| DC/PBE | 0.72 | 1.54 | 0.82 | 1.31 | 2.34 | 1.03 | 1.89 | 2.05 | 0.16 | 2.13 | 2.54 | 0.41 | 2.14 | 2.47 | 0.33 | 0.83 | 1.49 |
| ZORA/BLYP(ALDA) | 1.07 | 1.51 | 0.44 | 1.73 | 2.28 | 0.65 | 1.92 | 2.01 | 0.09 | 2.22 | 2.43 | 0.21 | 2.24 | 2.51 | 0.27 | 0.80 | 1.51 |
| DC/BLYP(ALDA) | 1.16 | 1.66 | 0.50 | 1.76 | 2.44 | 0.68 | 1.98 | 2.09 | 0.11 | 2.27 | 2.54 | 0.27 | 2.14 | 2.41 | 0.27 | 0.76 | 1.43 |
| DC/BLYP | 0.90 | 1.55 | 0.65 | 1.53 | 2.32 | 0.79 | 1.94 | 2.08 | 0.14 | 2.13 | 2.46 | 0.33 | 2.10 | 2.40 | 0.30 | 0.82 | 1.44 |
| ZORA/M06-L(ALDA) | 1.32 | 1.77 | 0.45 | 1.91 | 2.55 | 0.64 | 2.20 | 2.30 | 0.10 | 2.63 | 2.83 | 0.20 | 3.13 | 3.31 | 0.18 | 0.52 | 1.22 |
| ZORA/B3LYP(XALDA) | 1.29 | 1.79 | 0.50 | 1.74 | 2.53 | 0.79 | 2.92 | 3.04 | 0.12 | 2.77 | 3.00 | 0.23 | 2.62 | 2.93 | 0.31 | 0.34 | 0.60 |
| DC/B3LYP(XALDA) | 1.40 | 1.95 | 0.55 | 1.81 | 2.74 | 0.93 | 2.98 | 3.11 | 0.13 | 2.84 | 3.13 | 0.29 | 2.56 | 2.87 | 0.31 | 0.31 | 0.53 |
| DC/B3LYP | 1.21 | 1.87 | 0.66 | 1.62 | 2.64 | 1.02 | 2.95 | 3.10 | 0.15 | 2.74 | 3.07 | 0.33 | 2.53 | 2.86 | 0.33 | 0.33 | 0.63 |
| ZORA/PBE0(XALDA) | 1.32 | 1.84 | 0.52 | 1.69 | 2.55 | 0.86 | 3.18 | 3.30 | 0.12 | 2.97 | 3.21 | 0.24 | 2.87 | 3.17 | 0.30 | 0.19 | 0.40 |
| DC/PBE0(XALDA) | 1.42 | 1.99 | 0.57 | 1.77 | 2.76 | 0.99 | 3.23 | 3.37 | 0.14 | 3.05 | 3.35 | 0.30 | 2.79 | 3.08 | 0.29 | 0.16 | 0.32 |
| DC/PBE0 | 1.09 | 1.94 | 0.85 | 1.38 | 2.70 | 1.32 | 3.16 | 3.36 | 0.20 | 2.92 | 3.31 | 0.39 | 2.69 | 3.06 | 0.37 | 0.21 | 0.45 |
| ZORA/M06(XALDA) | 1.34 | 1.88 | 0.54 | 1.70 | 2.59 | 0.89 | 3.16 | 3.29 | 0.13 | 2.98 | 3.21 | 0.23 | 2.97 | 3.26 | 0.29 | 0.20 | 0.39 |
| ZORA/M06-2X(XALDA) | 1.26 | 1.86 | 0.60 | 1.30 | 2.40 | 1.10 | 3.97 | 4.12 | 0.15 | 3.49 | 3.76 | 0.27 | 2.87 | 3.32 | 0.45 | 0.26 | 0.60 |
| ZORA/SAOP(ALDA) | 2.03 | 2.47 | 0.44 | 2.50 | 3.06 | 0.56 | 2.89 | 2.99 | 0.10 | 3.04 | 3.23 | 0.19 | 2.61 | 2.86 | 0.25 | 0.47 | 1.02 |
| DC/SAOP(ALDA) | 1.98 | 2.47 | 0.49 | 2.46 | 3.18 | 0.72 | 2.93 | 3.04 | 0.11 | 3.09 | 3.35 | 0.26 | 2.45 | 2.71 | 0.26 | 0.48 | 0.98 |
| DC/CAM-B3LYP(XALDA) | 1.75 | 2.14 | 0.39 | 1.91 | 2.91 | 1.00 | 3.43 | 3.57 | 0.14 | 3.21 | 3.48 | 0.27 | 2.82 | 3.16 | 0.34 | 0.19 | 0.43 |
| DC/CAM-B3LYP | 1.65 | 2.08 | 0.43 | 1.75 | 2.83 | 1.08 | 3.41 | 3.57 | 0.16 | 3.09 | 3.42 | 0.33 | 2.83 | 3.18 | 0.35 | 0.13 | 0.28 |
| ZORA/TDHF | 1.88 | 2.66 | 0.78 | 1.37 | 2.98 | 1.61 | 6.79 | 6.98 | 0.19 | 5.53 | 5.84 | 0.31 | 3.87 | 4.54 | 0.67 | 1.52 | 3.46 |
| CASPT2 | 1.80 | 2.30 | 0.50 | 1.77 | 2.66 | 0.89 | 3.23 | 3.29 | 0.06 | 3.04 | 3.51 | 0.47 | 2.62 | 2.85 | 0.23 | 0.22 | 0.33 |
| IHFSCC | 1.51 | 2.06 | 0.55 | 1.48 | 2.56 | 0.88 | 3.38 | 3.52 | 0.14 | 3.37 | 3.58 | 0.21 | 2.87 | 3.14 | 0.27 | | |

dominant contributions from N, but still some U character), and $\sigma_u \rightarrow \sigma_g$ (for $\Sigma_u$, where both σs have N and U character).

For $NUO^+$, shown in Table 2, the picture is similar, but having as significant differences that the lowest Φ and Δ states are made up of excitations to the uranium ϕ,δ from σ orbitals with either N–U character or U–O character, whereas the Γ and Φ arise from excitations to the same uranium ϕ,δ from a π orbital with N–U character.

The Hartree–Fock virtual orbitals correspond to a system with one electron added relative to the reference determinant. While this is optimal for calculating electron affinities, it is not so for excitation energies calculated within the (1h,1p) sector of Fock space. In order to compare to the TDDFT results, the model spaces ($P$ and $P_m$) should include the $\delta_u$,$\phi_u$ orbitals. These are the lowest virtuals for $UO_2^{2+}$, both for Hartree–Fock and DFT, but the decreasing charge on the metal in $NUO^+$ and particularly in NUN places other orbitals at lower energy. For instance, for $NUO^+$, three σ and two π virtual orbitals have lower energies than the relevant δ,ϕ orbitals, while for NUN several (*e.g.* two $\sigma_u$, three $\pi_u$, three $\sigma_g$ and one $\pi_g$) orbitals are in between the HOMO and the $\delta_u$,$\phi_u$. These differences illustrate the need for increased model spaces in this work, compared to $UO_2^{2+}$, for which the $\delta_u$ and $\phi_u$ are the lowest-lying virtuals.

A related difference to the $UO_2^{2+}$ case is the extent to which the participation of a second set of δ,ϕ virtuals is important for $NUO^+$ and NUN. This is due to the energy separation of the first and second δ and ϕ virtuals, which for $UO_2^{2+}$ is of about 8 eV, but decreases to within 1.5–2 eV for $NUO^+$ and NUN, making these virtuals more important for orbital relaxation in the latter case.

Lastly, we note that in all three cases the $T_1$ diagnostic[77] for the (0h,0p) sector (which here is equivalent to a conventional CCSD calculation), namely 0.045 for $UO_2^{2+}$, 0.048 for $NUO^+$ and 0.049 for NUN, is rather similar and a bit higher than what is usually considered an indicator (<0.02) of single-reference character in light systems. This is typical for heavy elements, and should not be taken as an indication of multi-reference character.

### 3.2 CASPT2 electronic structure

The CASSCF wavefunction analysis points to ground-states of essentially single reference character, with weights for the HF determinant of about 0.86, 0.92 and 0.91 for $UO_2^{2+}$, $NUO^+$ and NUN, respectively. Also of importance is the fact that the determinants which contribute to the remaining 0.10 are made up of double excitations, and therefore do not point to important orbital relaxation effects.[78]

CASPT2 provides in a fairly straightforward manner information about the bond orders in the ground state (2.935(U–O) for $UO_2^{2+}$, 2.952(U–N) for NUN, and 1.957(U–O) and 2.977 (U–N) for $NUO^+$). Changes in the electron density upon excitation can be studied by means of an analysis of the Mulliken charges for each excited state. From these charges, summarized in Table 4, we can see that there is a general trend of displacing a small amount of density towards the uranium atom in comparison to the ground state. This effect is more systematic (in the sense of all states showing a transfer of charge from the ligands to the uranium) in the case of $UO_2^{2+}$ than for NUN, for which only the $\Sigma_u$ states have a rather pronounced ligand to uranium charge transfer relative to the ground state. For $NUO^+$ the most important effect is the migration of charge from one end of the molecule (the O atom) to the other (the N atom) rather than a net movement of charge towards the central atom.

### 3.3 Comparison of WFT excitation energies

In a previous investigation of the performance of IHFSCC and CASPT2 in calculating the electronic spectrum of

**Table 4** CASPT2 Mulliken charges for the ground and excited states of $UO_2^{2+}$, $NUO^+$ and NUN

| State | $UO_2^{2+}$ U | O | $NUO^+$ State | U | N | O | NUN State | U | N |
|---|---|---|---|---|---|---|---|---|---|
| $X^1\Sigma_g^+$ | 2.41 | −0.21 | $X^1\Sigma$ | 1.72 | −0.39 | −0.33 | $X^1\Sigma_g^+$ | 0.77 | −0.39 |
| $^1\Phi_g$ | 2.36 | −0.18 | $^1\Phi$ | 1.70 | −0.57 | −0.12 | $^1\Phi_g$ | 0.78 | −0.39 |
| $^3\Phi_g$ | 2.36 | −0.18 | $^3\Phi$ | 1.71 | −0.57 | −0.14 | $^3\Phi_g$ | 0.76 | −0.38 |
| $^1\Delta_g$ | 2.35 | −0.17 | $^1\Delta$ | 1.73 | −0.55 | −0.19 | $^1\Delta_g$ | 0.76 | −0.38 |
| $^3\Delta_g$ | 2.37 | −0.18 | $^3\Delta$ | 1.70 | −0.57 | −0.13 | $^3\Delta_g$ | 0.79 | −0.39 |
| $^1\Gamma_g$ | 2.25 | −0.13 | $^1\Gamma$ | 1.70 | −0.57 | −0.12 | $^1\Phi_u$ | 0.74 | −0.37 |
| $^3\Gamma_g$ | 2.26 | −0.13 | $^3\Gamma$ | — | — | — | $^3\Phi_u$ | 0.72 | −0.36 |
| $^1\Phi_u$ | 2.25 | −0.13 | $^1\Pi$ | 1.74 | −0.57 | −0.17 | $^1\Gamma_g$ | 0.78 | −0.39 |
| $^3\Phi_u$ | 2.26 | −0.13 | $^3\Pi$ | — | — | — | $^3\Gamma_g$ | 0.79 | −0.40 |
| $^1\Delta_u$ | 2.26 | −0.13 | $2^1\Phi$ | — | — | — | $^1\Sigma_u$ | 0.64 | −0.33 |
| $^3\Delta_u$ | 2.26 | −0.13 | $2^3\Phi$ | — | — | — | $^3\Sigma_u$ | 0.64 | −0.32 |



$UO_2^{2+}$,[27] two of us observed that CASPT2 typically shows discrepancies with respect to IHFSCC for individual excitation energies within a range of 0.1–0.4 eV. Furthermore, for the lowest $\Phi_g$ and $\Delta_g$, both singlet and triplet, CASPT2 overestimated the excitation energies, whereas for higher-lying states the opposite is true. The singlet–triplet splittings, however, were quite similar for both methods. The same general trends are seen here for $NUO^+$ and NUN in Tables 2 and 3, respectively. The differences between individual excitation energies for both methods are typically in the 0.1–0.3 eV range. We may conclude that the two WFT methods give the same semi-quantitative result and will take IHFSCC as the reference method, given its more systematic treatment of the excited states and inclusion of dynamic correlation effects beyond second-order perturbation theory.

## 4. The performance of DFT and TDDFT

Our TDDFT tests were focused on a subset of XC functionals, covering the following basic classes: LDA, GGAs (PBE, BLYP) and meta-GGAs (M06-L), hybrids (B3LYP and PBE0, with 20 and 25% of HF exchange, respectively) and meta-hybrids (M06 and M06-2X, with 27 and 54% of HF exchange, respectively), model potentials (SAOP) and range-separated hybrids (CAM-B3LYP).

We have chosen to represent all systems by restricted (closed-shell) Kohn–Sham calculations, given the evidence both from our wavefunction calculations and from previous studies (*e.g.* that of Kaltsoyannis,[10] Pierloot and coworkers,[18,25] those of Réal and coworkers[26] and that of Fromager[78] and coworkers) that this yields a proper description for their ground states. The suitability of this approach for uranyl was further demonstrated in the recent Kramers-restricted TDDFT calculations by Bast *et al.*[20]

### 4.1 Ground-state electronic structure

Before discussing the performance of (TD)DFT for the different electronically excited states, it is instructive to discuss the molecular orbitals (MOs) and the chemical bonding. The essentials of bonding in the actinyls are nicely summarized in a review by Denning;[79] in particular, for $UO_2^{2+}$ the currently accepted picture, in terms of the highest-lying occupied MOs, is that of a system of σ,π orbitals arising from the combination of oxygen 2p orbitals and the 5f,6p orbitals of uranium, ordered as $\pi_g > \pi_u > \sigma_g > \sigma_u$. In $UO_2^{2+}$ the contribution from the uranium 6p shows up in the relatively large gap between the σ orbitals, due to the repulsion between $\sigma_u$ and 6p,[80] a so-called "pushing from below" interaction.

To our knowledge only the aforementioned work of Kaltsoyannis has paid attention to the valence MO picture of NUN and $NUO^+$. Using a GGA functional (PB86), he found: (a) the same orbital ordering for NUN and $UO_2^{2+}$, but with a smaller (larger) energy gap between the σ (π) orbitals; and (b) an ordering of type $\pi > \sigma > \pi > \sigma$ for $NUO^+$, with the (energetically) lower π,σ pair mostly centered over the U–O bond, whereas the HOMO and HOMO-1 are mostly centered over the U–N bond. This picture is qualitatively consistent already with the Hartree–Fock results, but in order to obtain reliable information about the orbital ordering electron correlation should also be included. To remain within an orbital picture, we shall do so by comparing the vertical ionization potentials (IPs) obtained from the (1h,0p) sector in the IHFSCC calculations to those obtained from DFT.

The DFT IPs are here taken as approximations to the negative of the respective orbital energies, and while this is strictly valid only for the HOMO,[81] our results (shown in Table 5) indicate that this is a good approximation, in line with the findings of Chong and coworkers.[82]

**4.1.1 Describing the occupied space.** From the numbers in Table 5, we can confirm that the orbital scheme outlined above is maintained for all three molecules, with one exception for M06-2X (which for NUN places the $\sigma_u$ orbital below the $\sigma_g$). On the other hand, we observe significant quantitative differences as far as the differences in energy between orbitals (for a given molecule) are concerned. These are very much dependent on the type of functional in use (LDA/GGA, hybrids, metaGGAs/hybrids, *etc.*); for instance, the energy difference between the HOMO and HOMO-1 for NUN or $UO_2^{2+}$ can be halved just by going from GGAs to hybrids. Also striking is the fact that there is very little difference (of the order of 0.1 eV) in energy between $\sigma_g$ and $\sigma_u$ for IHFSCC in the case of $UO_2^{2+}$, whereas the DFT calculations yield differences between 0.3 and 1 eV.

These numbers are rather insensitive to the Hamiltonian; in general, we observe discrepancies no larger than 0.1 eV (but generally lower than 0.05 eV) between the ZORA and DC Hamiltonians (keeping in mind that one also might have effects due to the different basis sets used by ADF and DIRAC). This is in line with experience that approximate

**Table 5** Comparison of DFT and IHFSCC for the first three ionization potentials (IPs) for $UO_2^{2+}$, $NUO^+$ and NUN (in eV). As these ionized states in the IHFSCC are dominated by contributions from a single orbital and the DFT values are approximated by the negative of the orbital energies, we identify the IPs with the respective orbitals (which range from HOMO-2 to HOMO for DFT). ZORA and DC are Zero Order Regular Approximated and Dirac-Coulomb Hamiltonians, respectively

| | | NUN | | | $NUO^+$ | | | $UO_2^{2+}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\pi_u$ | $\sigma_g$ | $\sigma_u$ | σ(U–O) | π | σ(U–N) | $\pi_u$ | $\sigma_g$ | $\sigma_u$ |
| LDA | ZORA | 6.58 | 6.08 | 5.50 | 14.57 | 13.78 | 12.59 | 23.45 | 22.95 | 22.22 |
| | DC | 6.53 | 6.05 | 5.46 | 14.51 | 13.63 | 12.47 | 23.39 | 22.87 | 22.16 |
| PBE | ZORA | 6.32 | 5.84 | 5.22 | 14.33 | 13.52 | 12.32 | 23.19 | 22.74 | 21.94 |
| | DC | 6.29 | 5.82 | 5.21 | 14.28 | 13.38 | 12.23 | 23.12 | 22.66 | 21.90 |
| BLYP | ZORA | 6.18 | 5.76 | 5.15 | 14.22 | 13.36 | 12.24 | 23.02 | 22.59 | 21.84 |
| | DC | 6.16 | 5.77 | 5.17 | 14.18 | 13.22 | 12.16 | 22.95 | 22.52 | 21.81 |
| M06-L | ZORA | 6.25 | 5.72 | 5.07 | 14.40 | 13.53 | 12.12 | 23.19 | 22.91 | 21.93 |
| B3LYP | ZORA | 7.26 | 6.85 | 6.50 | 15.63 | 14.45 | 13.59 | 24.36 | 24.03 | 23.39 |
| | DC | 7.24 | 6.83 | 6.50 | 15.61 | 14.40 | 13.50 | 24.32 | 23.98 | 23.38 |
| PBE0 | ZORA | 7.57 | 6.85 | 6.50 | 16.01 | 14.90 | 13.92 | 24.76 | 24.45 | 23.78 |
| | DC | 7.54 | 7.08 | 6.79 | 15.98 | 14.76 | 13.82 | 24.72 | 24.39 | 23.76 |
| M06 | ZORA | 7.56 | 7.10 | 6.97 | 16.07 | 14.92 | 13.99 | 24.73 | 24.46 | 23.97 |
| M06-2X | ZORA | 8.86 | 8.30 | 8.58 | 17.66 | 16.24 | 15.45 | 26.27 | 25.90 | 25.64 |
| SAOP | ZORA | 10.16 | 9.92 | 9.28 | 18.73 | 17.69 | 16.69 | 27.82 | 27.64 | 26.72 |
| | DC | 10.08 | 9.83 | 9.10 | 18.62 | 17.52 | 16.49 | 27.76 | 27.55 | 26.57 |
| CAM-B3LYP | DC | 9.02 | 8.48 | 8.44 | 17.49 | 16.22 | 15.37 | 26.23 | 25.78 | 25.40 |
| IHFSCC | DC | 10.15 | 9.45 | 9.43 | 18.66 | 17.76 | 16.74 | 27.76 | 27.15 | 27.08 |

(but computationally efficient) two-component relativistic schemes such as ZORA yield accurate valence energies.

Looking at the performance of the various DFT approaches, we find the best agreement with IHFSCC for the SAOP potential. Comparing the DC values it shows errors in the range between −0.5 and 0.4 eV. It is followed by CAM-B3LYP and M06-2X, which systematically underestimate the IPs by about 1 eV and 1–1.5 eV, respectively. Next come two blocks, one grouping the remaining (meta)hybrids, and another encompassing LDA and the (meta)GGAs, which underestimate the IPs by about 3 and 4 eV, respectively.

While other factors, like the correlation part of the functionals, also play a role these trends may be rationalized by considering the degree to which self-interaction errors (SIE) are eliminated from each functional. The reader is referred to ref. 83–86 for discussion of the self-interaction problem. Important to note in the present context is the inability of approximate functionals to, in a one-electron picture (1-SIE), correctly cancel out the Coulomb interaction of the electron with itself through the exchange interaction (which in Hartree–Fock theory is exactly cancelled out) or, in a many-electron case (N-SIE), to properly describe the discontinuity of the derivative of the total energy with respect to (fractional) changes in particle number (the so-called integer discontinuity). The inability of GGAs to properly represent such discontinuities in the energy and exchange-correlation (xc) potential (the latter denoted here by $\Delta_{xc}$) have been shown to be behind the failures of TDDFT in describing charge-transfer excitations,[87] or to reproduce the correct asymptotic behavior of the exchange-correlation potential.[88,89] The latter defect is remedied by model potentials with the correct asymptotic (long-range) behavior, like LB94[60] and SAOP.[59]

Based on the analysis of Teale and coworkers,[90] valid for the case of pure functionals, we can estimate the value of $\Delta_{xc}$ from the relation

$$\Delta_{xc} = 2(I^0 + \varepsilon_{HOMO}) \quad (1)$$

where $I^0$ is a reference ionization potential (taken to be the IHFSCC value here) and $\varepsilon_{HOMO}$ is the Kohn–Sham orbital energy for the HOMO of the molecules (with the factor two arising from the assumption that, instead of a discontinuity, (meta)GGAs will exhibit an averaged potential over such discontinuity). From the differences between the LDA or GGAs and IHFSCC in Table 5 one sees that $\Delta_{xc}$ may have values of up to 8 eV, an indication that the effect of the integer discontinuity on the spectra of these systems is potentially large. We furthermore note that the asymptotically correct SAOP potential indeed provides a good agreement with the IHFSCC values.

An alternative to imposing the proper asymptotic behavior with a model potential is to introduce non-locality and reduce SIE *via* the inclusion of Hartree–Fock exchange (as in hybrid functionals), *via* an explicit dependence of the functional on the kinetic energy (metaGGAs) or a combination of both (metahybrids). This is also done in range-separated hybrids such as CAM-B3LYP or others[91,92] that offer a more detailed control over the incorporation of exact exchange than is possible in conventional hybrids.[84,93,94] While the analysis of Teale is not applicable to (meta)hybrids, their better performance (in particular when compared to the analogous GGA, as one can then suppose similar errors due to electron correlation or other factors) does indeed suggest improvements relative to the non-hybrid functionals.

**4.1.2 Remarks on the virtual space.** A detailed discussion concerning the representation of the virtual orbital space will not be made here. The main reason for that lies at the very different meaning of the virtual orbital energies[95] when pure (*e.g.* LDA or (meta)GGAs) or hybrid functionals are employed. It is well-known that for pure functionals the virtual orbital energies are good approximations to the ionization potentials of excited states, whereas in Hartree–Fock they represent approximations to electron affinities. For hybrids they are thus somewhere in between these two values making it

difficult to compare these values with the IHFSCC values (that strictly represent electron affinities).

A consequence of the difference between Hartree–Fock and pure DFT is that one finds, for the GGA functionals employed here, the low-lying virtuals to be uranium-centered $f_\phi$ and $f_\delta$ orbitals (to which we will observe the transitions from the occupied σ,π orbitals discussed above), whereas for the hybrids these are often found higher in energy than other orbitals such as the σ and π antibonding orbitals.

### 4.2 The performance of different functionals for the excited states

All investigated XC functionals were subsequently compared to the wave function methods in Table 1 ($UO_2^{2+}$), Table 2 ($NUO^+$) and Table 3 (NUN), which are known to perform very well for molecules containing heavy elements.[19,27,39,41,42] Since previous works indicate that IHFSCC energies are generally in better agreement with experiment than those from CASPT2, we chose to employ the former as our reference.

The tables contain, apart from the individual excitation energies and singlet–triplet splittings, the mean unsigned (MUE) and largest absolute (Max) errors with respect to IHFSCC for each molecule. We also provide a global picture in Fig. 1, where the (signed) errors are depicted for each individual excitation.

**4.2.1 General trends.** The statistical measures help to identify trends and we start discussing the MUE. For that, LDA and the GGAs considered show essentially the same results for all molecules, namely large underestimations with GGAs showing no clear improvement over LDA. The meta-GGA M06-L on the other hand does show improvement over both LDA and GGAs, almost halving the error. For the (meta)hybrids the errors are smaller still, about four times smaller than those of GGAs or LDA. We thereby note that whereas for NUN and $UO_2^{2+}$ the excitation energies are generally underestimated, for $NUO^+$ some functionals also show slight overestimations.

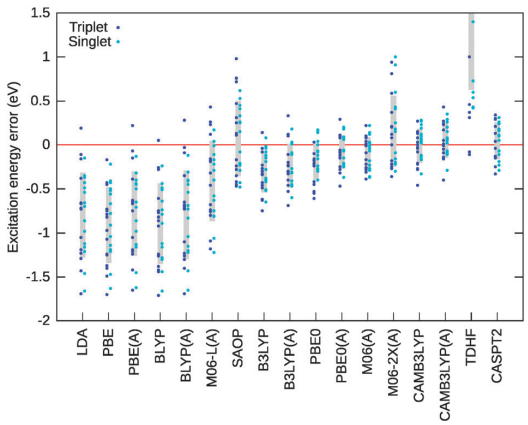



**Fig. 1** Errors with respect to IHFSCC for all excitations and all molecules. The gray boxes enclose a range of one sample standard deviation above and below the average error. Dots show individual errors for each excitation energy. The two highest (DFT) states for $NUO^+$ have been left out of the analysis. (ALDA)—evaluated using the ALDA approximation.

The B3LYP functional shows somewhat larger MUEs than PBE0 or M06(XALDA). The latter two show nearly identical results which is not so surprising as M06 bears a number of similarities to PBE0 (about the same amount of Hartree–Fock exchange, and exchange and correlation functionals based on those of PBE). It is nevertheless unfortunate that the higher flexibility in the functional form available for M06 does not translate into better accuracy than observed for PBE0. This may be due to the ALDA approximation employed in the current ADF implementation of this functional, however.

The M06-2X(XALDA) functional, in spite of having the same functional form as M06 (except for the amount of Hartree–Fock exchange), shows a worse performance than B3LYP, again indicating the important role played by the exchange energy. The model potential (SAOP) performs comparably to M06-2X(XALDA), while the range-separated (CAM-B3LYP) functional tends to show yet an improvement over PBE0 or M06. Furthermore, CAM-B3LYP generally matches the performance of CASPT2, even slightly outperforming it for NUN and $UO_2^{2+}$.

Nearly the same ranking of functional performance is seen for the largest absolute errors (Max), which are generally two to three times larger than the corresponding MUEs. The superior performance of (meta)hybrids and CAM-B3LYP compared to LDA or GGAs is evident for all three molecules, and it is interesting to see that, while the (meta)hybrids show a sightly better agreement with IHFSCC for NUN and $UO_2^{2+}$ in comparison to $NUO^+$, for (meta)GGAs the opposite is true.

We believe that, as stated above for the ionization potentials, the large errors seen for LDA and GGAs have to do with a poor description of exchange energies, that are quite different for the ground and excited states. The xc kernel plays a significant role in determining the accuracy here, and in it, the amount of HF exchange is important as the differences between M06(XALDA) and M06-2X(XALDA), and the similarities between PBE0(XALDA) and M06(XALDA) can attest. However, it remains to be seen whether or not the same will hold for other actinide compounds, especially in connection to typical charge-transfer or Rydberg-type excitations (and where one would expect CAM-B3LYP to clearly outperform the other hybrids).

Providing an understanding of the differences in standard deviation between the different groups of functionals, on the other hand, seems to be a much more difficult task. We can at this time only speculate that, at least to some extent, N-SIE effects that affect the various exited states differently will be important. In that case, one could expect that calculations with functionals that show large N-SIE, as is the case with GGAs[84,93,94] would then exhibit larger standard deviations. In this respect, the SAOP potential is perhaps an interesting example. We have already discussed that SAOP reduces the SIE by correcting the long-range part of the potential. However, since SAOP was conceptualized to be used with the ALDA approximation, the errors inherent to the LDA functional find themselves back into the response calculation (as indicated by the large standard deviations for the excitation energies).

These errors are incidentally of about the same order of any other calculations with (meta)GGAs and, more generally, of calculations with the ALDA kernel.

Also worth noting here is that, in spite of its better performance on the mean error compared to the investigated (local) GGA functionals, M06-L(ALDA) ultimately remains more in line with them than with the (meta)hybrids. At this point we cannot exclude that improvements could be seen if we departed from the ALDA kernel given which, as discussed below, can have significant effects on individual excitations.

**4.2.2 ALDA or Exact derivatives in the kernel?.** In order to judge whether the ALDA approximation changes significantly the electronic spectrum of investigated molecules we consider calculations with the PBE, PBE0, BLYP, B3LYP and CAM-B3LYP xc functionals with and without ALDA. The effect of the ALDA approximation is depicted in Fig. 2 for the lowest two excitations of each of the molecules considered.

From Fig. 2 and the values in Tables 1 through 3 we conclude that for the singlet states the effect of ALDA is small, but a surprisingly large discrepancy occurs for the triplet states, especially in the lowest electronic transitions (differences up to 0.4 eV). For the higher transitions the effect is smaller (difference up to 0.04) in all investigated molecules and functionals. PBE0 and B3LYP suffer somewhat less than PBE and BLYP from the ALDA approximation for the triplet energies, due to the fractional Hartree–Fock exchange still present in the kernel. We may thus expect that the ALDA approximation also has a large effect on triplet energies for the M06-L functional (for which an implementation of the full kernel is not yet available) but less so for M06 and in particular M06-2X which contain a large portion of Hartree–Fock exchange. Finally we note that by its construction as a model potential, it is difficult to remove the ALDA ansatz for SAOP, and here the large errors (see Fig. 1) in the triplet energies cannot be easily remedied.

**4.2.3 Comparison to previous calculations and benchmarks.** The results discussed above are in line with those from recent benchmark calculations on different molecular databases that do not include molecules containing heavy centers.[96–102] In particular, the recent comparison of the M06 family to other functionals by Jacquemin and coworkers[99] points to the same general trends seen here: M06 does show the best overall performance in the family and is close to PBE0, while M06-2X shows a slightly worse performance. They also show that M06-L outperform different GGAs, but still cannot match the accuracy of conventional hybrids functionals such as B3LYP.

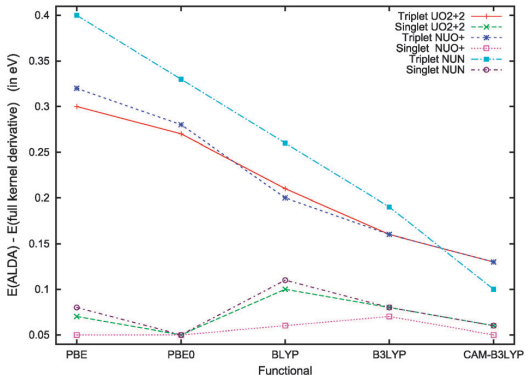



**Fig. 2** Errors for the first singlet and triplet Φ states due to the ALDA approximation for the PBE, PBE0, BLYP, B3LYP and CAM-B3LYP functionals.

It is difficult to directly compare our MUE values and those of Jacquemin and coworkers,[99] or those of Silva-Junior and coworkers,[101,102] due to the different methodologies used to obtain the reference values (and the extent to which basis set effects can influence the WFT[102] or TDDFT[103] results). We can nevertheless observe, for hybrids and metahybrids such as B3LYP and M06, a rather good agreement between our MUE and those of the literature (and similarly for the unsigned errors shown in Fig. 1). For GGAs, on the other hand, the values in the literature seem to be much smaller than ours. We are not able at this time to say whether this is definitely a degradation of performance for the GGAs for actinides or whether this is an artifact due to the limited size of our benchmark set.

Considering now calculations on molecules with heavy elements we confirm, for the uranyl spectrum, the observations of Bast *et al.*[20] who included spin–orbit coupling and compared the performance of functionals relative to the LR-CCSD results of Réal and coworkers.[26] While the latter have not considered the M06 family of functionals, they also reported a lowest MUE for CAM-B3LYP with LDA and GGAs severely underestimating the excitation energies.

Comparing our results to the benchmark calculations of Zhao and Truhlar (Table 17 in ref. 62) where a broad range of excitation energies calculated with different functionals are compared to reference values, we do not see the same drastic improvement going from hybrid functionals (B3LYP, PBE0) to metahybrids (M06, M06-2X). In our application, M06-2X brings the excitation energies too close to the Hartree–Fock values and introduces significant errors. Most likely this discrepancy between our particular molecules and excitations and the data presented by Zhao and Truhlar is due to the fact that we do not include Rydberg or extreme charge transfer states in our benchmark. In these cases it is crucial to use functionals with a correct treatment of the nonlocality of the change in the electron density. Based on our results we cannot recommend M06-2X for the systems and excitations studied in this work, despite its good performance in other benchmarks.

## 5. Conclusions

We investigated the performance of different classes of approximate exchange–correlation functionals in describing ten low-lying valence excitations for the uranyl ion ($UO_2^{2+}$) and two isoelectronic analogs, $NUO^+$ and NUN, by comparing them to wavefunction calculations (CASPT2 and Fock-space coupled cluster). A marked characteristic of such systems, all of which are closed-shell species in the ground state, is that the low-lying excited states under consideration correspond to excitation from the σ,π bonding orbitals to unoccupied orbitals which are essentially uranium f orbitals.

We can identify the following trends regarding the functional's performance: (a) LDA and (meta)GGAs show somewhat larger mean errors than (meta)hybrids or model potential such as SAOP; however, the standard deviation for those is significantly larger than for the (meta)hybrids; (b) one hardly observes an improvement for metaGGAs or meta hybrids in comparison to GGAs or hybrids, with perhaps the exception of the improvement of the mean error for M06-L over the GGAs considered; and (c) the performance of M06, PBE0 and CAM-B3LYP approaches that of CASPT2, both in terms of relatively small MAEs and standard deviations for the excitations. Of course, with only three molecules studied, one cannot rule out that the present agreement is fortuitous, but based on this benchmark M06, PBE0 and especially CAM-B3LYP appear appropriate for quantitative studies of actinide spectroscopy. Other hybrid functionals such as M06-2X and B3LYP are suited for (semi)quantitative or qualitative work, but we would strongly argue against employing non-hybrid (meta)GGAs even for qualitative investigations of excited states of actinyls.

In view of those trends, we believe that, while the correlation functional does play an important role in the accuracy of results—as seen in the differences between different functionals of same kind (GGAs, hybrids, *etc.*), what appears to be a critical factor governing the accuracy of the functionals tested is the degree of non-locality introduced through inclusion of HF exchange in hybrids or meta-hybrids. We could thereby rationalize why: (i) hybrids outperform their pure GGA counterparts; (ii) M06-L(ALDA) shows some improvement over the GGAs regarding the mean error but not in the standard deviation; (iii) the SAOP model yields excellent ionization potentials and mean errors for the excitation energies but has standard deviations similar to GGAs.

It is also clear that one must go beyond the ALDA approximation, given the rather large differences observed between the low-lying triplet states. Equally (or perhaps more) important, however, is that non-local effects should also be incorporated to the exchange-correlation kernel, as done for all (meta)hybrids and CAM-B3LYP, if one wishes to approach the accuracy of methods such as CASPT2.

It is, finally, interesting to note that for excitation energies the choice of relativistic (spin-free) Hamiltonian is almost irrelevant, so one can safely investigate the spectra of actinide-containing molecules with the more approximate two-component methods (such as ZORA), instead of using four-component approaches.

## Acknowledgements

The authors thank Dr Florent Réal, Dr Radovan Bast, and Dr Erik van Lenthe for helpful discussions on the practical aspects of carrying out the MOLCAS CASPT2 and the DIRAC and ADF TDDFT calculations, respectively. The authors acknowledge financial support from The Netherlands Organization for Scientific Research (NWO) *via* the Vici program, as well as computer time provided by the Dutch National Computing Facilities (NCF) at the Huygens and LISA facilities at SARA.

ASPG also acknowledges the use of computational resources from CINES ("Centre Informatique National de l'Enseignement Supérieur") and CCRT ("Centre de Calcul Recherche et Technologie"), under grant "phl2531".

## References

1 K. Nash, *Solvent Extr. Ion Exch.*, 1993, **11**, 729.
2 K. L. Nash, R. E. Barrans, R. Chiarizia, M. L. Dietz, M. Jensen and P. Rickert, *Solvent Extr. Ion Exch.*, 2000, **18**, 605.
3 E. P. Horwitz, D. G. Kalina, H. Diamond, G. F. Vandegrift and W. W. Schulz, *Solvent Extr. Ion Exch.*, 1985, **3**, 75.
4 G. Schreckenbach and G. A. Shamov, *Acc. Chem. Res.*, 2010, **43**, 19.
5 M. Pepper and B. E. Bursten, *Chem. Rev.*, 1991, **91**, 719–741.
6 Q.-J. Pan, G. A. Shamov and G. Schreckenbach, *Chem.–Eur. J.*, 2010, **16**, 2282.
7 M. del C. Michelini, J. Marçalo, N. Russo and J. H. Gibson, *Inorg. Chem.*, 2010, **49**, 3836.
8 N. Iché-Tarrat and C. J. Marsden, *J. Phys. Chem. A*, 2008, **112**, 7632.
9 G. S. Groenewold, A. K. Gianotto, M. E. McIlwain, M. J. van Stipdonk, M. Kullman, D. T. Moore, N. Polfer, J. Oomens, I. Infante, L. Visscher, B. Siboulet and W. A. de Jong, *J. Phys. Chem. A*, 2008, **112**, 508.
10 N. Kaltsoyannis, *Inorg. Chem.*, 2000, **39**, 6009.
11 N. Ismail, J.-L. Heully, T. Saue, J.-P. Daudey and C. J. Marsden, *Chem. Phys. Lett.*, 1999, **300**, 296.
12 T. Privalov, P. Macak, B. Schimmelpfennig, E. Fromager, I. Grenthe and U. Wahlgren, *J. Am. Chem. Soc.*, 2004, **126**, 9801.
13 G. A. Shamov, G. Schreckenbach and T. N. Vo, *Chem.–Eur. J.*, 2007, **13**, 4932.
14 G. Schreckenbach and G. A. Shamov, *Acc. Chem. Res.*, 2010, **43**, 19.
15 A. Rosa, E. J. Baerends, S. J. A. van Gisbergen, E. van Lenthe, J. A. Groeneveld and J. G. S. Snijders, *J. Am. Chem. Soc.*, 1999, **121**, 10356.
16 S. J. A. van Gisbergen, J. A. Groeneveld, A. Rosa, J. G. Snijders and E. J. Baerends, *J. Phys. Chem. A*, 1999, **103**, 6835.
17 S. J. A. van Gisbergen, A. Rosa, G. Ricciardi and E. J. Baerends, *J. Chem. Phys.*, 1999, **111**, 2499.
18 K. Pierloot, E. van Besien, E. van Lenthe and E. J. Baerends, *J. Chem. Phys.*, 2007, **126**, 194311.
19 F.-P. Notter, S. Dubillard and H. Bolvin, *J. Chem. Phys.*, 2008, **128**, 164315.
20 R. Bast, H. J. A. Jensen and T. Saue, *Int. J. Quantum Chem.*, 2009, **109**, 2091.
21 S. Fortier, G. Wu and T. W. Hayton, *J. Am. Chem. Soc.*, 2010, **132**, 6888.
22 D. R. Brown and R. G. Denning, *Inorg. Chem.*, 1996, **35**, 6158.
23 L. P. Spencer, P. Yang, B. L. Scott, E. R. Batista and J. M. Boncella, *Inorg. Chem.*, 2009, **48**, 2693.
24 S. Fortier and T. W. Hayton, *Coord. Chem. Rev.*, 2010, **254**, 197.
25 K. Pierloot and E. van Besien, *J. Chem. Phys.*, 2005, **123**, 204309.
26 F. Réal, V. Vallet, C. Marian and U. Wahlgren, *J. Chem. Phys.*, 2007, **127**, 214302.
27 F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet and E. Eliav, *J. Phys. Chem. A*, 2009, **113**, 12504.
28 F. Ruipérez, C. Danilo, F. Réal, J.-P. Flament, V. Vallet and U. Wahlgren, *J. Phys. Chem. A*, 2009, **113**, 1420.
29 R. Bast, *Ph.D. thesis*, l'Université Luis Pasteur, Strasburg, 2007.
30 S. Matsika, Z. Zhang, S. R. Brozell, J.-P. Blaudeau, Q. Wang and R. M. Pitzer, *J. Phys. Chem. A*, 2001, **105**, 3825.
31 M. Zhou, N. Ismail, C. Marsden and L. Andrews, *J. Phys. Chem. A*, 2000, **104**, 5495.
32 L. Gagliardi and B. O. Roos, *Chem. Soc. Rev.*, 2007, **36**, 893.
33 R. D. Hunt, J. T. Yustein and L. Andrews, *J. Chem. Phys.*, 1993, **98**, 6070.
34 K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, *J. Phys. Chem.*, 1990, **94**, 5483.
35 K. Andersson, P.-Å. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218.

36 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2000, **113**, 9905.
37 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 6862.
38 L. Visscher, E. Eliav and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 9720.
39 I. Infante, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2006, **125**, 074301.
40 I. Infante, E. Eliav, M. J. Vilkas, Y. Ishikawa, U. Kaldor and L. Visscher, *J. Chem. Phys.*, 2007, **127**, 124308.
41 G. L. Macchia, I. Infante, J. Raab, J. K. Gibson and L. Gagliardi, *Phys. Chem. Chem. Phys.*, 2008, **10**, 7278.
42 A. S. P. Gomes, L. Visscher, H. Bolvin, T. Saue, S. Knecht, T. Fleig and E. Eliav, *J. Chem. Phys.*, 2010, **133**, 064305.
43 R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.*, 2007, **79**, 291–351.
44 L. Meissner and M. Musiał, *Recent Progress in Coupled Cluster Methods*, Springer Sience + Business Media B.V., 2010, p. 395.
45 V. V. Ivanov, D. I. Lyakh and L. Adamowicz, *Phys. Chem. Chem. Phys.*, 2009, **11**, 2355–2370.
46 G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. F. Guerra, E. J. Baerends, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931.
47 C. F. Guerra, J. G. Snijders, G. te Velde and E. J. Baerends, *Theor. Chem. Acc.*, 1998, **99**, 391.
48 ADF2009.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.
49 DIRAC, a relativistic *ab initio* electronic structure program, Release DIRAC08 (2008), written by L. Visscher, H. J. Aa. Jensen, and T. Saue, with new contributions from R. Bast, S. Dubillard, K. G. Dyall, U. Ekström, E. Eliav, T. Fleig, A. S. P. Gomes, T. U. Helgaker, J. Henriksson, M. Iliaš, Ch. R. Jacob, S. Knecht, P. Norman, J. Olsen, M. Pernpointner, K. Ruud, P. Sałek, and J. Sikkema (see http://dirac.chem.sdu.dk).
50 E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1993, **99**, 4597.
51 U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen and K. Ruud, *J. Chem. Theory Comput.*, 2010, **6**, 1971.
52 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200.
53 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
54 A. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
55 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
56 B. G. Johnson, P. M. W. Gill and J. A. Pople, *J. Chem. Phys.*, 1993, **98**, 5612.
57 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.
58 M. Ernzerhof and G. Scuseria, *J. Chem. Phys.*, 1999, **110**, 5029.
59 P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344.
60 R. van Leeuwen and E. J. Baerends, *Phys. Rev. A*, 1994, **49**, 2421.
61 Y. Zhao and D. G. Truhlar, *J. Chem. Phys.*, 2006, **125**, 13126.
62 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215.
63 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51.
64 E. van Lenthe and E. J. Baerends, *J. Am. Chem. Soc.*, 2003, **24**, 1142.
65 K. G. Dyall, *Theor. Chem. Acc.*, 2007, **117**, 483.
66 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.
67 R. A. Kendall, T. H. Dunning and R. Harrison, *J. Chem. Phys.*, 1989, **96**, 6796.
68 L. Visscher and T. Saue, *J. Chem. Phys.*, 2000, **113**, 3996.
69 L. Visscher, *Theor. Chem. Acc.*, 1997, **98**, 68.
70 G. Karlström, R. Lindh, P.-A. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady and L. Seijo, *Comput. Mater. Sci.*, 2003, **28**, 222.
71 J. Finleya, P.-A. Malmqvist, B. O. Roos and L. Serrano-Andrés, *Chem. Phys. Lett.*, 1998, **288**, 299.
72 N. Douglas and N. M. Kroll, *Ann. Phys. (Leipzig)*, 1974, **82**, 89.
73 B. A. Hess, *Phys. Rev. A*, 1986, **33**, 3742.
74 P.-O. Widmark, P.-A. Malmqvist and B. O. Roos, *Theor. Chim. Acta*, 1990, **77**, 291.
75 N. Forsberg and P.-A. Malmqvist, *Chem. Phys. Lett.*, 1997, **274**, 196.
76 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2004, **121**, 6634–6639.
77 T. J. Lee and P. R. Taylor, *Int. J. Quantum Chem. Symp.*, 1989, **23**, 199.
78 E. Fromager, F. Real, P. Wahlin, U. Wahlgren and H. J. A. Jensen, *J. Chem. Phys.*, 2009, **131**, 054107.
79 R. G. Denning, *J. Phys. Chem.*, 2007, **111**, 4125.
80 K. Tatsumi and R. Hoffmann, *Inorg. Chem.*, 1980, **19**, 2656–2658.
81 J. F. Janak, *Phys. Rev. B*, 1978, **18**, 7165.
82 D. P. Chong, O. V. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2002, **117**, 1760.
83 A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 194112.
84 P. Mori-Sánchez, A. J. Cohen and W. T. Yang, *J. Chem. Phys.*, 2006, **125**, 201102.
85 S. Kümmel and L. Kronin, *Rev. Mod. Phys.*, 2008, **80**, 30.
86 T. Körzdorfer, S. Kümmel and M. Mundt, *J. Chem. Phys.*, 2008, **129**, 014110.
87 D. J. Tozer, *J. Chem. Phys.*, 2003, **119**, 12697.
88 D. J. Tozer and N. C. Handy, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2117.
89 J. B. Krieger, Y. Li and J. Iafrate, *Phys. Rev. A*, 1992, **45**, 101.
90 A. M. Teale, F. D. Proft and D. J. Tozer, *J. Chem. Phys.*, 2008, **129**, 044110.
91 A. J. Cohen, P. Mori-Sanchez and W. Yang, *J. Chem. Phys.*, 2007, **126**, 191109.
92 B. G. Janesko, T. M. Henderson and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2009, **11**, 443–454.
93 T. M. Henderson, A. F. Izmaylov, G. Scalmani and G. E. Scuseria, *J. Chem. Phys.*, 2009, **131**, 044108.
94 J.-W. Song, M. A. Watson, A. Nakata and K. Hirao, *J. Chem. Phys.*, 2008, **129**, 184113.
95 K. Capelle, *Braz. J. Phys.*, 2006, **36**, 1318.
96 D. Jacquemin, E. A. Perpéte, I. Cioni and C. Adamo, *J. Chem. Theory Comput.*, 2010, **6**, 1532.
97 J. Preat, D. Jacquemin, J.-M. A. V. Wathelet and E. A. Perpéte, *J. Phys. Chem. A*, 2006, **110**, 8144.
98 D. Jacquemin, V. Wathelet, E. A. Perpéte, I. Cioni and C. Adamo, *J. Chem. Theory Comput.*, 2009, **5**, 2420.
99 D. Jacquemin, E. A. Perpete, I. Ciofini, C. Adamo, R. Valero, Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2010, **6**, 2071.
100 M. A. Rohrdanz, K. M. Martins and J. M. Herbert, *J. Chem. Phys.*, 2009, **130**, 054112.
101 M. Silva-Junior, M. Schreiber, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.*, 2008, **129**, 104103.
102 M. Silva-Junior, M. Schreiber, S. Sauer and W. Thiel, *J. Chem. Phys.*, 2010, **133**, 174318.
103 I. Ciofini and C. Adamo, *J. Phys. Chem. A*, 2007, **111**, 5549–5556.

## C.5 Paper V

THE JOURNAL OF CHEMICAL PHYSICS **141**, 041107 (2014)

# Communication: Relativistic Fock-space coupled cluster study of small building blocks of larger uranium complexes

Paweł Tecmer,[1,a)] André Severo Pereira Gomes,[2] Stefan Knecht,[3] and Lucas Visscher[1]
[1]*Amsterdam Center for Multiscale Modeling (ACMM), VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*
[2]*Laboratoire PhLAM, CNRS, UMR 8523, Université de Lille 1, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France*
[3]*ETH Zürich, Laboratory of Physical Chemistry, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland*

(Received 3 July 2014; accepted 20 July 2014; published online 31 July 2014)

We present a study of the electronic structure of the $[UO_2]^+$, $[UO_2]^{2+}$, $[UO_2]^{3+}$, NUO, $[NUO]^+$, $[NUO]^{2+}$, $[NUN]^-$, NUN, and $[NUN]^+$ molecules with the intermediate Hamiltonian Fock-space coupled cluster method. The accuracy of mean-field approaches based on the eXact-2-Component Hamiltonian to incorporate spin–orbit coupling and Gaunt interactions are compared to results obtained with the Dirac–Coulomb Hamiltonian. Furthermore, we assess the reliability of calculations employing approximate density functionals in describing electronic spectra and quantities useful in rationalizing Uranium (VI) species reactivity (hardness, electronegativity, and electrophilicity).
 [http://dx.doi.org/10.1063/1.4891801]

Research on uranium chemistry focuses on the interaction between the atom and its nearest ligands as understanding this interaction is crucial in technological challenges such as assessing the feasibility of separating, storing, and disposing nuclear waste and in its use in catalysis.[1,2] Key properties are ionization potentials (IPs) and electron affinities (EAs), hardness and electronegativity of small building blocks of uranium species, because these can serve as indicators of chemical reactivity of larger uranium complexes. Particularly useful are accurate optical spectroscopy studies of prototypical uranium compounds, as this constitutes a very sensitive probe of the uranium environment.[3] Because the interpretation of experimental spectra for actinde complexes is often not straightforward,[3] theoretical investigations provide a means to gain further insight. This is, however, not an easy task due to the required balanced treatment of both electron correlation and relativistic effects.[4–7] Several publications[8–11] underlined the need of highly sophisticated (and computationally expensive) wave function approaches to obtain reliable results.[12,13] Among these, the relativistic intermediate Hamiltonian Fock-space coupled cluster singles and doubles (IH-FSCCSD) method stood out as the most reliable theoretical model.[14–18,20,21]

It is a well-established fact that spin–orbit interactions are essential to reach spectroscopic accuracy.[4–6] Unfortunately, an accurate four-component Dirac–Coulomb (DC) Hamiltonian has the downside of increased computational expense, notably for the AO–MO transformation step for correlated calculations. This high computational cost is a limiting factor to investigate larger molecular systems with accurate correlation methods. An alternative is found in two-component approaches, in which the large and small components will be decoupled and solutions can be expressed in terms of the large-component functions only.[4] A particularly interesting two-component method is based on the so-called "eXact 2-Component"[22] (X2C) which can provide an exact decoupling for the matrix representation of the Hamiltonian. However, in order to construct the decoupling, one needs solutions for a coupled system. This drawback has motivated the development of two main computational schemes: one based on the decoupling in an atomic mean-field[23] (AMF) fashion,[24–26] and another in which the decoupling is based on the molecular four-component solutions,[27] with two-electron interactions being approximated in a molecular mean-field (MMF) fashion. The AMF approach is computationally more advantageous than the MMF, since only two-electron integrals over large component basis functions are required at the Self-Consistent Field (SCF) step, but at the cost of discarding multi-center contributions to the spin–orbit interactions. While the AMF approach has seen a wider range of applications than the MMF one, not much is known of their accuracy in general and for actinides in particular.

Another relativistic effect worth considering for actinides is the Gaunt (G) interaction,[28] which describes the magnetic interaction between the spin current of one electron and the orbital current of another one.[4]

In this communication, we highlight the importance of both spin–orbit coupling and the (approximate) Gaunt interactions on the electronic structure of the $[UO_2]^{2+}$, $[NUO]^+$, and NUN isoelectronic series (and their electron-attached and ionized counterparts).

*Vertical excitation energies.* The spin–orbit coupling electronic spectra obtained from the IH-FSCCSD approach using different relativistic Hamiltonians (and the ground-state geometries from Ref. 18) are compiled in Table I ($UO_2^{2+}$), Table II ($NUO^+$), and Table III (NUN). (We refer the reader to the supplementary material[19] for computational details). For

[a)]Present address: Department of Chemistry and Chemical Biology, McMaster University, Hamilton, 1280 Main Street West L8S 4M1, Canada. Electronic mail: tecmer@mcmaster.ca

TABLE I. Fifteen lowest-lying IH-FSCCSD vertical excitation energies of the $UO_2^{2+}$ molecule ($r_{U\text{-}O}$ = 1.708 Å). Excitation energies are given in eV.

| Ω | Character (from DC) | DC | DC(G) | X2C/AMF | X2C/MMF | X2C(G)/MMF |
|---|---|---|---|---|---|---|
| $2_g$ | 65% $\sigma_{1/2u}\phi_{5/2u} + 17\%\ \pi_{1/2u}\phi_{5/2u}$ | 1.764 | 1.727 | 1.741 | 1.765 | 1.728 |
| $1_g$ | 70% $\sigma_{1/2u}\delta_{3/2u} + 20\%\ \pi_{1/2u}\delta_{3/2u}$ | 1.843 | 1.801 | 1.819 | 1.843 | 1.801 |
| $3_g$ | 66% $\sigma_{1/2u}\phi_{5/2u} + 18\%\ \pi_{1/2u}\phi_{5/2u}$ | 1.919 | 1.880 | 1.896 | 1.919 | 1.881 |
| $2_g$ | 47% $\sigma_{1/2u}\delta_{3/2u} + 20\%\ \sigma_{1/2u}\delta_{5/2u}$ | 2.131 | 2.085 | 2.108 | 2.131 | 2.085 |
| $3_g$ | 65% $\sigma_{1/2u}\delta_{5/2u} + 22\%\ \pi_{1/2u}\delta_{5/2u}$ | 2.514 | 2.453 | 2.491 | 2.513 | 2.452 |
| $4_g$ | 72% $\sigma_{1/2u}\phi_{7/2u} + 20\%\ \pi_{1/2u}\phi_{7/2u}$ | 2.592 | 2.532 | 2.571 | 2.592 | 2.532 |
| $3_g$ | 65% $\sigma_{1/2u}\phi_{7/2u}$ | 2.992 | 2.938 | 2.972 | 2.992 | 2.938 |
| $2_g$ | 50% $\sigma_{1/2u}\delta_{5/2u} + 22\%\ \sigma_{1/2u}\delta_{3/2u}$ | 3.394 | 3.346 | 3.374 | 3.394 | 3.346 |
| $1_g$ | 96% $\pi_{3/2u}\phi_{5/2u}$ | 3.853 | 3.817 | 3.829 | 3.853 | 3.817 |
| $4_g$ | 97% $\pi_{3/2u}\phi_{5/2u}$ | 3.968 | 3.932 | 3.945 | 3.968 | 3.933 |
| $3_g$ | 97% $\pi_{3/2u}\delta_{3/2u}$ | 4.079 | 4.042 | 4.056 | 4.079 | 4.042 |
| $2_u$ | 81% $\sigma_{1/2g}\phi_{5/2u} + 16\%\ \sigma_{1/2g}\delta_{3/2u}$ | 4.084 | 4.038 | 4.057 | 4.084 | 4.039 |
| $0_g^-$ | 96% $\pi_{3/2u}\delta_{3/2u}$ | 4.086 | 4.049 | 4.064 | 4.087 | 4.049 |
| $0_g^+$ | 96% $\pi_{3/2u}\delta_{3/2u}$ | 4.122 | 4.082 | 4.098 | 4.122 | 4.082 |
| $3_u$ | 98% $\sigma_{1/2g}\phi_{5/2u}$ | 4.125 | 4.079 | 4.099 | 4.125 | 4.080 |

each excited state, the main character of molecular spinors involved in the electronic transition is provided next to the Ω quantum number. The low-lying electronic excitations in the uranyl cation ($UO_2^{2+}$) are dominated by transitions from the occupied $\sigma^{(O)}_{1/2u}$- and $\pi^{(O)}_{1/2u}$-spinors to the nonbonding $\phi^{(U)}_{5/2u}$-, $\phi^{(U)}_{7/2u}$-, $\delta^{(U)}_{3/2u}$-, and $\delta^{(U)}_{5/2u}$-spinors (cf. Table I). In the upper part of the spectrum, electrons are excited from the lower occupied $\sigma^{(O)}_{1/2g}$- and $\pi^{(O)}_{3/2u}$-spinors. This picture is in line with previous four-component studies on this molecule,[29] numerical differences being due to the elongated U–O distance (1.683 vs. 1.708 Å) and larger active space utilized in this work, which led to some quantitative changes in the upper part of the spectrum. The new active spaces, corresponding to a significant enlargement of the number of active virtual spinors and the incorporation of the U 5*d*-shell in the Q-space,[29] were motivated by the observation that for $NUO^+$ and NUN including the 5*d* spinors greatly improved the IH-FSCCSD convergence of sector (0*h*, 1*p*), and that determinants in which higher virtuals are occupied become increasingly important for the $P_m$ space in sector (1*h*, 1*p*).

Similar to $UO_2^{2+}$, in the $NUO^+$ molecule low-lying excitations occur mostly from the light element (the nitrogen atom) to the uranium atom (see Table II). Both electronic spectra are well described by electronic transitions from the occupied $\sigma^{(N)}_{1/2}$- and $\sigma'^{(N)}_{1/2}$-spinors in the lower part, and $\pi^{(N)}_{3/2}$- and $\pi^{(N)}_{1/2}$-spinors in the upper part to the nonbonding $\phi^{(U)}_{5/2}$-, $\phi^{(U)}_{7/2}$-, $\delta^{(U)}_{3/2}$-, and $\delta^{(U)}_{5/2}$-spinors (compare Tables I and II). A significant difference is, however, that the low-lying excited states in the $NUO^+$ molecule lie much lower in energy than in $UO_2^{2+}$ (ca. 0.7 eV).

The electronic states for NUN are quite distinct from those for $UO_2^{2+}$ and $NUO^+$, in that they possess very pronounced multi-reference character (cf. Table III) and show a

TABLE II. Fifteen lowest-lying IH-FSCCSD vertical excitation energies of the $NUO^+$ molecule ($r_{U\text{-}O}$ = 1.761 Å, $r_{U\text{-}N}$ = 1.698 Å). Excitation energies are given in eV.

| Ω | Character (from DC) | DC | DC(G) | X2C/AMF | X2C/MMF | X2C(G)/MMF |
|---|---|---|---|---|---|---|
| 2 | 71% $\sigma_{1/2}\phi_{5/2} + 15\%\ \sigma'_{1/2}\phi_{5/2}$ | 1.018 | 0.987 | 1.001 | 1.020 | 0.992 |
| 3 | 69% $\sigma_{1/2}\phi_{5/2}$ | 1.147 | 1.114 | 1.130 | 1.149 | 1.120 |
| 1 | 68% $\sigma_{1/2}\delta_{3/2} + 15\%\ \sigma'_{1/2}\delta_{3/2}$ | 1.215 | 1.179 | 1.197 | 1.216 | 1.183 |
| 2 | 54% $\sigma_{1/2}\delta_{3/2}$ | 1.440 | 1.401 | 1.423 | 1.441 | 1.405 |
| 4 | 72% $\sigma_{1/2}\phi_{7/2} + 15\%\ \sigma'_{1/2}\phi_{7/2}$ | 1.778 | 1.725 | 1.762 | 1.779 | 1.728 |
| 3 | 69% $\sigma_{1/2}\delta_{5/2}$ | 1.811 | 1.758 | 1.794 | 1.811 | 1.760 |
| 3 | 62% $\sigma_{1/2}\phi_{7/2}$ | 2.101 | 2.052 | 2.086 | 2.102 | 2.056 |
| 2 | 57% $\sigma_{1/2}\delta_{5/2}$ | 2.361 | 2.316 | 2.345 | 2.362 | 2.319 |
| 1 | 96% $\pi_{3/2}\phi_{5/2}$ | 2.656 | 2.621 | 2.640 | 2.662 | 2.630 |
| 4 | 97% $\pi_{3/2}\phi_{5/2}$ | 2.743 | 2.708 | 2.726 | 2.748 | 2.716 |
| 3 | 83% $\pi_{1/2}\phi_{5/2}$ | 2.932 | 2.894 | 2.918 | 2.937 | 2.903 |
| $0^-$ | 93% $\pi_{3/2}\delta_{3/2}$ | 2.947 | 2.911 | 2.932 | 2.953 | 2.919 |
| 3 | 88% $\pi_{3/2}\delta_{3/2}$ | 2.987 | 2.950 | 2.971 | 2.992 | 2.958 |
| $0^+$ | 93% $\pi_{3/2}\delta_{3/2}$ | 2.988 | 2.949 | 2.973 | 2.993 | 2.957 |
| 2 | 82% $\pi_{1/2}\phi_{5/2}$ | 3.016 | 2.966 | 2.991 | 3.010 | 2.974 |

TABLE III. Fifteen lowest-lying IH-FSCCSD vertical excitation energies of the NUN molecule ($r_{U\text{-}N}$ = 1.739 Å). Excitation energies are given in eV.

| Ω | Character (from DC) | DC | DC(G) | X2C/AMF | X2C/MMF | X2C(G)/MMF |
|---|---|---|---|---|---|---|
| $2_g$ | $52\%\ \sigma_{1/2u}\phi_{5/2u} + 26\%\ \pi_{1/2u}\phi_{5/2u}$ | 0.956 | 0.923 | 0.936 | 0.957 | 0.927 |
| $3_g$ | $50\%\ \sigma_{1/2u}\phi_{5/2u} + 24\%\ \pi_{1/2u}\phi_{5/2u}$ | 1.103 | 1.068 | 1.083 | 1.103 | 1.072 |
| $1_g$ | $45\%\ \sigma_{1/2u}\delta_{3/2u} + 20\%\ \pi_{1/2u}\delta_{3/2u}$ | 1.134 | 1.094 | 1.106 | 1.134 | 1.098 |
| $2_g$ | $30\%\ \sigma_{1/2u}\delta_{3/2u} + 15\%\ \pi_{1/2u}\delta_{3/2u}$ | 1.398 | 1.355 | 1.374 | 1.398 | 1.358 |
| $4_g$ | $49\%\ \sigma_{1/2u}\phi_{7/2u} + 23\%\ \pi_{1/2u}\phi_{7/2u} + 16\%\ \sigma_{1/2u}\phi'_{7/2u}$ | 1.699 | 1.645 | 1.680 | 1.698 | 1.646 |
| $3_g$ | $43\%\ \sigma_{1/2u}\delta_{5/2u} + 20\%\ \pi_{1/2u}\delta_{5/2u}$ | 1.757 | 1.704 | 1.739 | 1.757 | 1.705 |
| $3_g$ | $38\%\ \sigma_{1/2u}\phi_{7/2u} + 22\%\ \pi_{1/2u}\phi_{7/2u}$ | 2.076 | 2.028 | 2.059 | 2.076 | 2.029 |
| $2_g$ | $33\%\ \sigma_{1/2u}\delta_{5/2u} + 17\%\ \pi_{1/2u}\delta_{5/2u}$ | 2.519 | 2.476 | 2.502 | 2.519 | 2.478 |
| $1_u$ | $40\%\ \sigma_{1/2u}\delta_{3/2g} + 24\%\ \pi_{1/2u}\delta_{3/2g}$ | 2.669 | 2.696 | 2.680 | 2.669 | 2.696 |
| $0_u^+$ | $41\%\ \sigma_{1/2u}\sigma_{1/2g} + 30\%\ \pi_{1/2u}\sigma_{1/2g}$ | 2.709 | 2.757 | 2.740 | 2.709 | 2.755 |
| $1_u$ | $41\%\ \pi_{3/2u}\sigma'_{1/2g} + 29\%\ \pi'_{3/2u}\sigma''_{1/2g}$ | 2.711 | 2.759 | 2.743 | 2.711 | 2.757 |
| $1_g$ | $80\%\ \pi_{3/2u}\phi_{5/2u}$ | 2.711 | 2.675 | 2.690 | 2.711 | 2.679 |
| $2_u$ | $34\%\ \sigma_{1/2u}\delta_{3/2g} + 24\%\ \pi_{1/2u}\delta_{3/2g}$ | 2.749 | 2.767 | 2.758 | 2.749 | 2.768 |
| $4_g$ | $84\%\ \pi_{3/2u}\phi_{5/2u}$ | 2.844 | 2.808 | 2.823 | 2.844 | 2.811 |
| $2_u$ | $73\%\ \sigma_{1/2g}\phi_{5/2u}$ | 2.895 | 2.857 | 2.875 | 2.895 | 2.860 |

relatively low-lying state ($0_u^+$ at 2.7 eV) in which electrons in NUN can be transferred to spinors centered solely on the nitrogen atom. That said, for all species the eight lowest-lying electronic transitions involve the same unoccupied uranium spinors, with transition energies being rather similar (compare Tables II and III) and showing the same ordering with respect to Ω for NUN and $NUO^+$, but lower than the corresponding ones in $UO_2^{2+}$ by about 0.7–0.8 eV.

The distinctive features of the NUN electronic structure with respect to $NUO^+$ or $UO_2^{2+}$ originate from differences both in the occupied spinors (which are delocalized over the uranium and nitrogen atoms) and the lowest unoccupied spinors. The importance of the latter can be better understood by comparing the electronic structures of the Uranium (V) species: while the $UO_2^+$ and NUO molecules bear the same ground state electronic configuration ($\sigma^2_{1/2u}\phi^1_{5/2u}$ and $\sigma^2_{1/2}\phi^1_{5/2}$, respectively) and in both cases low-lying electronic excitations occur to the unoccupied uranium $\phi$- and $\delta$-spinors as in the Uranium (VI) species, the $NUN^-$ moiety exhibits a totally different ground state electronic configuration ($\sigma^2_{1/2u}\sigma^1_{1/2g}$) and other type of spinors are involved in low-lying excitations (see Tables S4– S6 of the supplementary material[19]). It is worth to note that the lowest-lying excited state in $NUN^-$ is much higher in energy than in the $UO_2^+$ and NUO molecules (0.95 eV vs. 0.25 eV). Conversely, the IH-FSCCSD excitation energies obtained for $NUN^+$ are quite comparable to those of $UO_2^{3+}$ in terms of energy and character; both of them are, however, again different from $NUO^{2+}$ (see Tables S7– S9 of the supplementary material[19]).

Excitation energies calculated from the DC Hamiltonian and those from the approximate X2C/AMF and X2C/MMF Hamiltonians are similar for all molecules investigated in this work. The overall agreement of the X2C/MMF data with the DC data is very good for all molecules and discrepancies are usually smaller than 0.005 eV (cf. Tables I–III and Tables S4– S9 of the supplementary material[19]). Yet, the advantage of X2C/MMF over DC is reduced CPU time required for the four-index transformation, a bottleneck of correlated calculations (see Figure S1 of the supplementary material[19]). The errors originating from the X2C/AMF approximation amount up to 0.03 eV which may be acceptable since this approach provides further computational savings. The presence of Gaunt interactions in the DC Hamiltonian (see DC(G) columns in Tables I–III and Tables S4– S9 of the supplementary material[19]) causes changes in the excitation energies on the order of 0.03–0.07 eV for all molecules. Notably, the X2C(G)/MMF results are almost indistinguishable from the reference DC(G) results.

The performance of TD-DFT for the electronic spectra of the $UO_2^{2+}$, $NUO^+$, and NUN molecules is summarized in Figure 1. The overall accuracy of the PBE0, M06, and CAM-B3LYP exchange–correlation functionals for the spin–orbit electronic spectra is acceptable. The deviations are no larger than 0.5 eV. In particular, the TD-DFT excitation energies calculated with the CAM-B3LYP exchange–correlation functional slightly overestimate, and with the PBE0 and M06 exchange–correlation functionals underestimate the reference data for $UO_2^{2+}$ and NUN. Interestingly, all functionals tend to shift the origin of the electronic spectra of the $NUO^+$ molecule towards lower energies (more negative errors) than for $UO_2^{2+}$ and NUN where the errors are in effect similar. The CAM-B3LYP exchange–correlation functional remains closest to IH-FSCCSD as it was observed for the spin–free electronic spectra (cf. Ref. 18). We believe that this is due to reduced self-interaction errors.[30,31] Furthermore, the inclusion of Gaunt interaction does not change the overall performance of CAM-B3LYP and its contribution to the electronic spectra is considerably smaller than the accuracy of TD-DFT (see Tables S10– S12 of the supplementary material[19]).

*IPs, EAs and related properties*. Table IV lists the first IPs and EAs calculated with the IH-FSCCSD approach. The largest IP (ca. 27 eV) and EA (ca. 14 eV) are obtained for the uranyl cation. These values are reduced approximately by factor of two in the $NUO^+$ molecule, and further scaled down in the NUN moiety. Thus, the overall trends in IPs and EAs

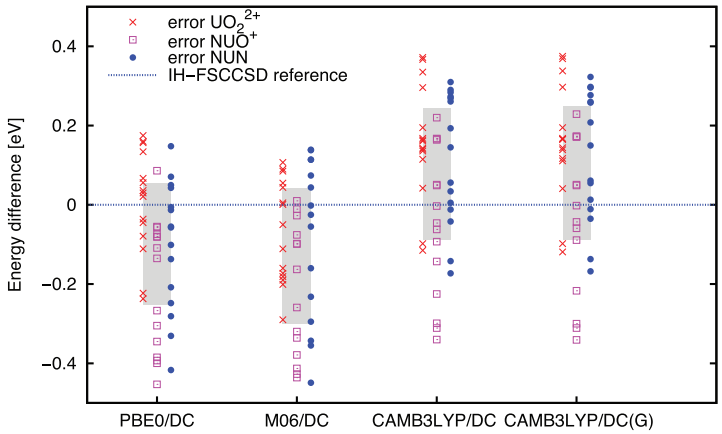



FIG. 1. Errors with respect to IH-FSCCSD for all excitations and all molecules. The gray boxes enclose a range of one sample standard deviation above and below the average error. Dots show individual errors for each excitation energy and are color-coded: red, violet, and blue correspond to $UO_2^{2+}$, $NUO^+$, and NUN, respectively. (Quantitative measure of the error can be found in Tables S10– S12 of the supplementary material.[19])

in the isoelectronic series of Uranium (VI) follow the charge decrease from +2 to 0.

The spin–free IH-FSCCSD EAs are significantly influenced by the presence of spin–orbit coupling. The effect is considerably larger for $UO_2^{2+}$ (ca. 0.5 eV) and $NUO^+$ (ca. 0.35 eV) than for NUN (ca. 0.06 eV). In the former two species, the electron is attached to the $\phi_{5/2}$ spinor in contrast to the $\sigma_{1/2}$-spinor in NUN for which spin–orbit coupling is less important[4] (cf. Tables S4– S6 of the supplementary material[19] for more details). Conversely, the effect of spin–orbit coupling on the first IPs of Uranium (VI) species is, in general, small and comparable for all molecules. This is easy to understand since an electron is removed from essentially the same $\sigma_{1/2}$-spinor in each system (cf. Tables S7– S9 of the supplementary material[19] for more details).

TABLE IV. IPs, EAs, electronegativity ($\chi$), hardness ($\eta$) and electrophilicity ($\omega$) obtained form the DC Hamiltonian (in eV).

| | Approach | $IP_1$ | $EA_1$ | $\chi$ | $\eta$ | $\omega$ |
|---|---|---|---|---|---|---|
| $UO_2^{2+}$ | IH-FSCCSD | | | | | |
| | spin–free | 27.09 | 14.42 | 20.76 | 6.34 | 34.00 |
| | spin–orbit | 26.99 | 14.91 | 20.95 | 6.04 | 36.33 |
| | Spin–orbit DFT | | | | | |
| | PBE0 | 23.74 | 17.89 | 20.82 | 2.93 | 74.06 |
| | M06 | 23.64 | 18.09 | 20.87 | 2.78 | 78.44 |
| | CAM-B3LYP | 25.27 | 16.72 | 21.00 | 4.28 | 51.55 |
| $NUO^+$ | IH-FSCCSD | | | | | |
| | spin–free | 16.74 | 6.25 | 11.50 | 5.25 | 12.60 |
| | spin–orbit | 16.71 | 6.59 | 11.65 | 5.06 | 13.41 |
| | Spin–orbit DFT | | | | | |
| | PBE0 | 13.83 | 9.39 | 11.61 | 2.22 | 30.36 |
| | M06 | 13.78 | 9.58 | 11.68 | 2.10 | 32.48 |
| | CAM-B3LYP | 15.33 | 8.21 | 11.77 | 3.56 | 19.46 |
| NUN | IH-FSCCSD | | | | | |
| | spin–free | 9.43 | 0.63 | 5.03 | 4.40 | 2.88 |
| | spin–orbit | 9.37 | 0.69 | 5.03 | 4.34 | 2.91 |
| | Spin–orbit DFT | | | | | |
| | PBE0 | 6.81 | 2.12 | 4.47 | 2.35 | 4.25 |
| | M06 | 6.69 | 2.11 | 4.40 | 2.29 | 4.23 |
| | CAM-B3LYP | 8.31 | 1.41 | 4.86 | 3.45 | 3.42 |

From EAs and IPs, one can calculate quantities such as electronegativity, chemical electrophilicity, or hardness, which are often used to rationalize the species' reactivity.[32,33] While it is beyond the scope of this communication to discuss the reactivity of larger uranium compounds with these descriptors, we nonetheless elaborate on the relative importance of spin–orbit coupling and the reliability of DFT to describe them.

Comparing first spin–free and spin–orbit IH-FSCCSD results listed in Table IV, we observe moderate changes in electronegativity ($\chi$), hardness ($\eta$), and electrophilicity ($\omega$) driven by the changes in EA in the spin–orbit case; this would suggest that a spin–free description could be sufficient, but due to the different nature of virtual spinors in the spin–free and spin–orbit description of $UO_2^{2+}$ and $NUO^+$, this requires further investigation. As one could expect due to the different molecular charges in the Uranium (VI) species, both $\chi$ and $\omega$ decrease when going from $UO_2^{2+}$ to $NUO^+$ to NUN; however, we find that $\eta$-values are rather similar for either molecule. Thus, knowing that molecules with small positive $\eta$ are most reactive,[32] we can speculate that the NUN and $NUO^+$ molecules are only slightly more reactive than $UO_2^{2+}$.

Comparing IH-FSCCSD and DFT, we conclude that, while CAM-B3LYP yields results close to IH-FSCCSD for all quantities, the other hybrid exchange–correlation functionals tend to significantly underestimate $\eta$ and overestimate $\omega$, which calls into question the validity of conclusions drawn from DFT calculations for these species.

To conclude, our work pointed out to some similarities in the electronic structures of $UO_2^{2+}$, $NUO^+$, and NUN which could be used to rationalize their electronic spectra. The occupied spinors in NUN and virtual spinors in $NUO^+$ closely resemble the electronic structure of the $UO_2^{2+}$ moiety.

P.T. gratefully acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada as well as computer time provided by the Reiher research group at ETHZ. P.T. and L.V. thank for financial support from the Netherlands Organization for Scientific Research (NWO). A.S.P.G. acknowledges support from the European Commission through the FP7 TALISMAN project, the French supercomputer facilities CINES and CCRT, and the French National Research Agency. S.K. acknowledges financial support (2010–2013) from the Danish Natural Science Research Council.

[1]K. Nash, Solvent Extr. Ion Exch. **11**, 729 (1993).
[2]G. J. Hutchings, C. S. Heneghan, I. D. Hudson, and S. H. Taylor, Nature (London) **384**, 341 (1996).
[3]R. G. Denning, J. Phys. Chem. A **111**, 4125 (2007).
[4]T. Fleig, Chem. Phys. **395**, 2 (2012).
[5]K. G. Dyall and J. K. Fægri, *Introduction to Relativistic Quantum Chemistry* (Oxford, New York, 2007).
[6]P. S. Bagus, E. S. Ilton, R. L. Martin, H. J. A. Jensen, and S. Knecht, Chem. Phys. Lett. **546**, 58 (2012).
[7]G. Schreckenbach and G. A. Shamov, Acc. Chem. Res. **43**, 19 (2010).
[8]K. Pierloot, E. van Besien, E. van Lenthe, and E. J. Baerends, J. Chem. Phys. **126**, 194311 (2007).
[9]H.-S. Hu, Y.-H. Qiu, X.-G. Xiong, W. H. E. Schwarz, and J. Li, Chem. Sci. **3**, 2786 (2012).
[10]P. Tecmer, R. Bast, K. Ruud, and L. Visscher, J. Phys. Chem. A **116**, 7397 (2012).
[11]P. Tecmer, N. Govind, K. Kowalski, W. A. de Jong, and L. Visscher, J. Chem. Phys **139**, 034301 (2013).
[12]F. Réal, V. Vallet, C. Marian, and U. Wahlgren, J. Chem. Phys. **127**, 214302 (2007).
[13]P. Tecmer, K. Boguslawski, O. Legaza, and M. Reiher, Phys. Chem. Chem. Phys. **16**, 719 (2014).
[14]A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **115**, 6862 (2001).
[15]I. Infante, A. S. P. Gomes, and L. Visscher, J. Chem. Phys. **125**, 074301 (2006).
[16]I. Infante, M. Vilkas, I. Ishikawa, U. Kaldor, and L. Visscher, J. Chem. Phys. **127**, 124308 (2007).
[17]A. S. P. Gomes, C. R. Jacob, and L. Visscher, Phys. Chem. Chem. Phys. **10**, 5353 (2008).
[18]P. Tecmer, A. S. P. Gomes, U. Ekström, and L. Visscher, Phys. Chem. Chem. Phys. **13**, 6249 (2011).
[19]See supplementary material at http://dx.doi.org/10.1063/1.4891801 for concerning computational details and TD-DFT excitation energies.
[20]P. Tecmer, H. van Lingen, A. S. P. Gomes, and L. Visscher, J. Chem. Phys. **137**, 084308 (2012).
[21]A. S. P. Gomes, C. R. Jacob, F. Réal, L. Visscher, and V. Vallet, Phys. Chem. Chem. Phys. **15**, 15153 (2013).
[22]M. Ilias and T. Saue, J. Chem. Phys. **126**, 064102 (2007).
[23]B. A. Hess, C. M. Marian, U. Wahlgren, and O. Gropen, Chem. Phys. Lett. **251**, 365 (1996).
[24]D. Peng, W. Liu, Y. Xiao, and L. Cheng, J. Chem, Phys. **127**, 104106 (2007).
[25]D. Peng and M. Reiher, J. Chem. Phys. **136**, 244108 (2012).
[26]D. Peng, N. Middendorf, F. Weigend, and M. Reiher, J. Chem. Phys. **138**, 184105 (2013).
[27]J. Sikkema, L. Visscher, T. Saue, and M. Ilias, J. Chem. Phys. **131**, 124116 (2009).
[28]J. A. Gaunt, Proc. R. Soc. Lon. A **122**, 159 (1929).
[29]F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet, and E. Eliav, J. Phys. Chem. A **113**, 12504 (2009).
[30]A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, J. Chem. Phys. **125**, 194112 (2006).
[31]P. Mori-Sánchez, A. J. Cohen, and W. T. Yang, J. Chem. Phys. **125**, 201102 (2006).
[32]R. G. Parr and R. G. Pearson, J. Am. Chem. Soc. **105**, 7512 (1983).
[33]P. W. Ayers and J. Melin, Theor. Chem. Acc. **117**, 371 (2006).

## C.6 Paper VI

PCCP

PAPER

View Article Online
View Journal | View Issue

# Electronic structure investigation of the evanescent $AtO^+$ ion†

Cite this: *Phys. Chem. Chem. Phys.*, 2014, **16**, 9238

André Severo Pereira Gomes,*[a] Florent Réal,[a] Nicolas Galland,[b] Celestino Angeli,[c] Renzo Cimiraglia[c] and Valérie Vallet[a]



Received 16th December 2013,
Accepted 10th February 2014

DOI: 10.1039/c3cp55294b

www.rsc.org/pccp

The electronic structure of the XO and $XO^+$ (X = I, At) species, as well that of a $AtO^+$–$H_2O$ complex have been investigated using relativistic wave-function theory and density functional theory (DFT)-based approximations (DFAs). The *n*-electron valence state perturbation method with the perturbative inclusion of spin–orbit coupling including spin–orbit polarization effects (SO-NEVPT2) was shown to yield transition energies within 0.1 eV of the reference four-component intermediate Fock-space coupled cluster (DC-IHFSCCSD) method at a significantly lower computational cost and can therefore be used as a benchmark to more approximate approaches in the case of larger molecular systems. These wavefunction calculations indicate that the ground state for the $AtO^+$ and $AtO^+$–$H_2O$ systems is the $\Omega = 0^+$ component of the $^3\Sigma^-$ LS state, which is quite well separated (by $\simeq$0.5 eV) from the $\Omega = 1$ components of the same state and from the $\Omega = 2$ state related to the $^1\Delta$ LS state (by $\simeq$1 eV). Time-dependent DFT calculations, on the other hand, place the $\Omega = 1$ below the $\Omega = 0^+$ component with the spurious stabilization of the former increasing as one increases the amount of Hartree–Fock exchange in the DFAs, while those employing the Tamm–Dancoff approximation and DFAs not including Hartree–Fock exchange yield transition energies in good agreement with SO-NEVPT2 or DC-IHFSCCSD for the lower-lying states. These results indicate the ingredients necessary for devising a DFA-based computational protocol applicable to the study of the properties of large $AtO^+$ clusters so that it may (at least) qualitatively reproduce reliable reference (SO-NEVPT2) calculations.

## 1 Introduction

Astatine (At), the heaviest naturally occurring halogen in the periodic table, is a rare radioelement due to all of its isotopes being short-lived. Among the isotopes the 100% alpha-particle emitter $^{211}$At is of considerable interest as a radiotherapeutic agent for targeted alpha therapy in nuclear medicine, due to its half-life of 7.2 h. This characteristic, together with its low availability (production at cyclotrons with maximum yields of $10^{-8}$ grams), make it extremely challenging for experimentalists to explore the basic chemical nature of this element, something that is recognized to be crucial for the development of targeted radiotherapy agents.[1] In that respect, theoretical investigations may be particularly valuable for determining the nature of At-containing species as well as their interactions with the environment (solution).

Unlike other halogens, astatine is thought to exist in different cationic forms in aqueous solutions as recently shown by experimental studies, either measuring distribution coefficients or high-performance ion-exchange chromatography, to determine the charge of the species.[2,3] Under acidic and oxidizing conditions, astatine exists as the $AtO^+$ cation. However, to assess the presence of a given species, it is important to predict its thermodynamic stability, as for instance the hydrolysis constant, a quantity which suffers from large experimental uncertainties as pointed out by Champion *et al.*[3] Quantum chemical calculations can therefore be used to investigate the existence of $AtO^+$ and to attempt to explain the unexpected reactivity with organic and inorganic ligands – which may seem odd given that a qualitative MO diagram constructed from the $At^+$ ($p^4$) and O ($p^4$) fragments would suggest a triplet ground state that, in analogy to the $O_2$ molecule, would not react readily with closed-shell species.

Ayed *et al.*,[4,5] based on an extensive set of scalar relativistic density functional theory (DFT)-based calculations employing the M06-2X exchange–correlation (XC) functional, have recently put forth the suggestion that the driving force behind the

[a] *Université de Lille 1, Laboratoire PhLAM, CNRS UMR 8523, CNRS FR 2416, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France. E-mail: andre.gomes@univ-lille1.fr; Fax: +33 3 2033 7020; Tel: +33 3 2043 4980*
[b] *Laboratoire CEISAM UMR CNRS 6230, Université de Nantes – Faculté des Sciences et des Techniques, BP 92208 44322 Nantes Cedex 3, France*
[c] *Dipartimento di Chimica, Università di Ferrara, Via Borsari 46, I-44121 Ferrara, Italy*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3cp55294b

observed reactivity of $AtO^+$ would be a change in the nature of the ground state, from a triplet in the gas phase to a singlet electronic state in solution, and have identified the point at which the singlet state becomes stabilized as occurring in the presence of four water molecules. Furthermore, they found that the inclusion of spin–orbit coupling (SOC) interactions does not alter the trends found at the scalar relativistic level. One may wonder, however, whether these calculations are sufficiently accurate to establish such changes in the electronic structure of the species, in the absence of more refined calculations or of direct experimental evidence such as electronic spectra, ionization potential (IP) or electron affinity (EA) measurements.

One cause for concern, apart from the multireference nature of singlet states in diradicals, is the well-known tendency of most density functional approximations (DFAs) to spuriously delocalize charge,[6–8] as well as of DFAs including the Hartree–Fock (HF) exchange (such as M06-2X) to favor high-spin states over those of lower multiplicity.[9,10] Another is the importance of spin–orbit coupling (SOC) interactions on the electronic structure of species containing heavy elements;[11–13] in particular, for astatine, SOC has been found to be quite important for structures and spectroscopic constants even for closed-shell species such as HAt[14] causing an increase in the bond length of about 0.03 Å, which is enough to offset the bond length contraction of about 0.02 Å found if only scalar relativistic effects are taken into account.

Thus, in this work our main goal is to study the electronic structure of the $AtO^+$ species with wavefunction theory (WFT)-based correlated electronic structure approaches such as multireference CI (with Davidson correction, MRCI + DC) and *n*-electron valence state perturbation (NEVPT2)[15–19] methods, capable of correctly accounting for the multireference character of the ground and excited-state wavefunctions, and to which SOC effects can be included *via* spin–orbit CI (SOCI) approaches.[20,21] Here we have chosen to investigate the performance of NEVPT2, for what we believe is the first time for heavy elements, instead of CASPT2 since by construction of its zeroth-order Hamiltonian (defined using Dyall's model Hamiltonian[22]) it does not suffer from the intruder state problem that plagues CASPT2.[15–19]

These will be compared to calculations with selected DFAs including SOC at the self-consistent field (SCF) level, and from that we expect to establish the most suitable approach, in terms of computational cost and accuracy, for subsequent studies of $AtO^+$ or other astatine complexes in solution. However, as there are no experimental measurements on the electronic spectra, IPs or EAs, we shall take an indirect route and perform benchmark calculations on the iodine homologue $IO^+$ as well as on the IO radical, for which experimental data and other theoretical studies are available. The approach serving as the reference was chosen to be the intermediate Hamiltonian Fock-space coupled-cluster method, employing the Dirac–Coulomb (DC) Hamiltonian (DC-IHFSCCSD), in light of its accuracy in prior studies on species containing heavy elements.[12,13,23–30]

This work is organized as follows: after outlining the computational details pertaining to each calculation, we begin the discussion by an assessment of the performance of IHFSCCSD for the $XO^+$ and XO (X = I, At) species, followed by the assessment of the more approximate approaches. We conclude by extending such a comparison to an $AtO^+$–water complex.

## 2 Computational methods

### 2.1 $XO^+$, $XO^-$ and XO systems

**2.1.1 Four-component coupled cluster.** All four-component (4C) calculations were carried out with a development version of the Dirac electronic structure code[31] (revision ab65b36), employing Dyall's basis sets[32,33] of triple-zeta quality for astatine and iodine, and Dunning's aug-cc-pVTZ sets[34] for oxygen, all of which are left uncontracted. The Dirac–Coulomb (DC) Hamiltonian along with the usual approximation of the (SS|SS)-type two-electron integrals by a point-charge model were used throughout.[35]

The electron correlation method employed was the DC-IHFSCCSD,[36–38] which allows for a proper description of a possible multiconfigurational nature of the ground and/or excited states. Whenever possible, single-reference coupled cluster calculations including the dynamical correlation with singles and doubles (DC-CCSD) with the inclusion of perturbative triples substitution (DC-CCSD(T))[38,39] were also performed.

In all calculations for the isolated species the $C_{\infty v}$ symmetry was used, and the correlation space ($Q$) consists of molecular spinors (MSs) with energies between −10.0 and 100.0 au. This corresponds to 32 correlated electrons and 248 virtual spinors for the systems containing iodine, and 46 correlated electrons and 340 virtual spinors for the systems containing astatine, respectively.

In the IHFSCCSD calculations the ground and excited states for the $XO^+$ were obtained with the (2h,0p) sector of Fock-space, starting from the closed-shell species $XO^-$ (sector (0h,0p)). In this process one also calculates the (1h,0p) sector of Fock-space, thus yielding the ground and excited states, electron affinities (EAs) and ionization potentials (IPs) for the XO radical as well.

The model ($P_m$) space used for the two isolated systems contains valence spinors arising form the p–p manifold ($\sigma_{1/2}$, $\pi_{3/2}$, $\pi_{1/2}$, $\pi_{3/2}^*$, $\pi_{1/2}^*$, see Fig. 1) with the exception of $\sigma_{1/2}^*$, which is unoccupied in the $XO^-$ reference. This implies that electronic states with important contributions from Slater determinants in which $\sigma_{1/2}^*$ is occupied will not be properly described (see discussion for more details). The remaining inner occupied spinors included in the correlation space (*e.g.* $ns_{1/2}$,$(n-1)d_{3/2}$,$(n-1)d_{5/2}$ *etc.* on the heavy centers) are part of the intermediate space ($P_i$). The resulting $P(=P_m + P_i)$ space is therefore analogous to those employed by Rota *et al.*[13] (''scheme (b)'') in the study of chalcogen diatomics.

Equilibrium geometries for the different electronic states were obtained by a polynomial fit to energies calculated for X–O internuclear distances ranging from 1.70 Å to 2.10 Å for $IO^-$/IO/$IO^+$ and from 1.80 Å to 2.20 Å for $AtO^-$/AtO/$AtO^+$, with increments of 0.01 Å between 1.75 Å and 1.95 Å, and between 1.85 Å and 2.05 Å, for $IO^-$/IO/$IO^+$ and $AtO^-$/AtO/$AtO^+$ respectively.

In addition to those, at selected geometries we have performed IHFSCCSD calculations with the spin-free Dirac–Coulomb (SFDC)

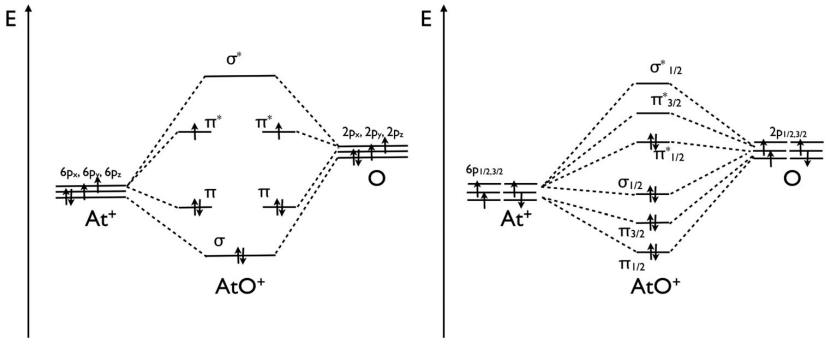



Fig. 1 Molecular orbital scheme for $AtO^+$ molecule in the spin–orbit (right) and spin–orbit-free (left) frameworks.



Hamiltonian[40,41] (SFDC-IHFSCCSD) calculations, employing the same $P$ and $Q$ spaces as above, and in $C_{2v}$ symmetry.

**2.1.2 Relativistic correlated calculations with SOC *a posteriori*.** An alternative to the 4C approaches above is to start from a scalar relativistic (SR) correlated approach and to add spin–orbit coupling (SOC) *a posteriori* (also referred to as perturbatively, in contrast to the variational treatment at the self-consistent field (SCF) level afforded by the use of the DC Hamiltonian above). In these calculations relativity was introduced *via* Relativistic Effective-Core-Pseudopotentials (RECPs) which replace, for both halides, either 28 (ECP28MDF)[42] or 60 core electrons (ECP60MDF).[43] The corresponding augmented quadruple-zeta basis sets by Peterson *et al.*[42,43] have been used for the halides, while the augmented quadruple-zeta basis set of Kendall *et al.*[34] was used for the oxygen atom.

Using $C_{2v}$ symmetry, the molecular orbitals (MOs) of $XO^+$ were optimized using a state-averaged complete active space self-consistent field calculations (CASSCF)[44–47] including 13 singlet states and 11 triplet states. The molecular active space encompasses the σ, σ*, π and π* relativistic molecular orbitals, as illustrated in Fig. 1.

Dynamical correlation effects beyond the CASSCF treatment are taken into account using either a variational method, namely the size-extensive corrected multireference configuration interaction method[48,49] with Davidson correction (DC)[50] (MRCI + DC), or using the partially contracted *n*-electron valence state perturbation (PC-NEVPT2) approach,[15–19] by computing the second-order correction to the energy of the various electronic states with the quasi-degenerate (QD-NEVPT2) formalism.[18] In all post-CASSCF treatments, the oxygen 1s core orbital and the sub-valence $(n-1)$s, $(n-1)$p and $(n-1)$d halide orbitals were kept frozen.

The spin–orbit coupled states are then computed in two ways. A first approach, hereafter referred to as "contracted spin–orbit configuration interaction" (c-SOCI), couples by the spin–orbit interaction (using the spin–orbit counterparts of the scalar RECPs) all the 24 spin–orbit free correlated states within the contracted state-interacting method implemented in the molpro2012[51] quantum chemistry suite of programs. The diagonal elements of the spin–orbit coupling matrices are either MRCI + DC or QD-NEVPT2 energies. The second approach, hereafter referred to as the "uncontracted SOCI" (uc-SOCI), diagonalizes the total relativistic Hamiltonian over a configurational space corresponding to the CAS plus all single-excitations (directly coupled by the one-electron spin–orbit ECP to the CAS configurations), projecting the effect of dynamic correlation (QD-NEVPT2) onto that SOCI model space by an effective spin–orbit Hamiltonian.[21,52]

All scalar relativistic correlated calculations and c-SOCI calculations were performed using the molpro2012[51] quantum chemistry suite of programs, while the uc-SOCI calculations were performed with the EPCISO[21] code, interfaced to the molcas78[53] quantum chemistry package. For the MRCI + DC calculations, the same geometries as in the four-component coupled cluster approach were employed (see above), whereas uc-SOCI/QD-NEVPT2 calculations were performed at $R(IO^+)$ = 1.806 Å and $R(AtO^+)$ = 1.903 Å.

**2.1.3 DFT-based approximations.** We have employed the ADF program[54–56] for all DFT-based calculations. We considered the spin–orbit Zeroth-Order Regular Approximation (ZORA) Hamiltonian,[57–59] and employed basis sets of triple-zeta plus polarization (TZ2P) quality for all atoms without freezing core orbitals.

Excitation energies were obtained from time-dependent DFT (TDDFT) calculations, both in the full form and in the Tamm–Dancoff approximation (TDA), and employing the adiabatic LDA (ALDA) approximation.

The density functionals considered were the Minnesota family[60] (the meta-GGA M06-L and the meta-Hybrids, M06 and M06-2X), as well as the statistical average of orbital potentials (SAOP)[61–63] model potential.

### 2.2 $AtO^+$–water complex

The structure for the $AtO^+$–water complex was obtained with the ADF program, employing TZ2P basis sets, the SO-ZORA Hamiltonian and the M06-2X functional, using the same reference as discussed above. The geometry optimization was carried out in $C_{2v}$ symmetry. Coordinates for the optimized structure can be found in the ESI.†

The electronic spectra calculations were performed at the optimized geometry. Both the WFT-based and DFT-based

calculations were performed roughly as above; for the WFT-based calculations, apart from obvious adjustments (*e.g.* due to the additional water molecular orbitals/spinors), we only used the uncontracted SOCI/QD-NEVPT2 method where the number of spin–orbit free states correlated were identical to that of $AtO^+$ calculations, noting that here we also freeze the 1s water oxygen orbital in the NEVPT2 correlation step. To quantify the effect of the water molecule on the $AtO^+$ spectrum, the u-SOCI wave-functions of $AtO^+$–water are projected on the $AtO^+$ wave-functions using the TRANSO program.[64]

## 3 Results and discussion

### 3.1 Benchmark calculations on XO and $XO^+$

As mentioned above, we will employ the DC-IHFSCCSD method as a reference to which we compare our results of more approximate approaches for astatine, due to the absence of experimental spectroscopic data. We now proceed to its evaluation for the IO radical, as its spectra, electron affinity and ionization potentials are well characterized experimentally using different approaches such as photoelectron spectroscopy,[66] high-resolution rotational spectroscopy,[67] as well as fluorescence spectroscopy.[69]

**3.1.1 Iodine oxide radical.** The DC-IHFSCCSD spectra and spectroscopic constants for IO are summarized in Table 1, whereas the states' composition in terms of the determinants of the model space $P$ is shown in Table S1 of the ESI.†

The calculated equilibrium distance for the ground-state compares rather well with both experimental values,[66,67] showing a slight ($\simeq$0.007 Å) elongation with respect to the latter. It also compares well with the CCSD(T)/augmented quintuple-zeta basis calculations of Peterson *et al.*,[42] as they differ only by 0.0027 Å. Peterson *et al.* found out two important contributions to the equilibrium distance arising from extrapolation to the basis set limit and higher excitations in the coupled cluster treatment, but they have opposite signs and similar magnitudes so that they cancel each other. Our results are, on the other hand, clearly superior to the SO-MRCI results carried out by Roszak *et al.*[65] with relatively small basis sets, which show differences with respect to the experimental data of roughly one order of magnitude ($\simeq$0.05 Å) larger.

The calculated harmonic frequency for the ground-state, on the other hand, is not in very good agreement with experiment or the CCSD(T) calculations of Peterson *et al.*[42] Where the latter shows discrepancies of about 2 $cm^{-1}$ with respect to experiment, in our case these are of roughly an order of magnitude larger, at about 40 $cm^{-1}$ which, in absolute terms, is close to the discrepancy between the SO-MRCI results of Roszak *et al.*[65] and experiment. We can attribute the difference between DC-IHFSCC and experiment partially to basis set incompleteness, but also to the lack of triple excitations which, for $IO^+$, decrease the harmonic frequency by 60 $cm^{-1}$ (see Section 3.1.3).

The same applies to the bond length and harmonic frequency of the first excited state ($\Omega$ = 1/2), which arises from the spin–orbit splitting of the spin–orbit free $^2\Pi$ state. A more interesting difference, however, is that the DC-IHFSCCSD excitation energy compares better to the experimental one (overestimating it by about 0.02 eV) than both the CCSD(T) and SO-MRCI calculations (which underestimate it by 0.03 and 0.05 eV, respectively).

If we set the limit of acceptable accuracy to about 0.1 eV, these differences appear to be small. They might arise from the way spin–orbit coupling has been introduced in these calculations.

Table 1 Bond lengths ($R_e$ in Å), harmonic frequencies ($\omega_e$, in $cm^{-1}$), vertical ($T_v$, in eV) and adiabatic ($T_e$, in eV) excitation energies calculated with the DC-IHFSCCSD method for XO (X = I, At). Additionally, we present the CBS/CCSD(T) results of Peterson *et al.*[42] the SOCI/MRCI results of Roszak,[65] as well as the available experimental data for IO

| | | IO | | | | AtO | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\Omega$ | $R_e$ (Å) | $\omega_e$ ($cm^{-1}$) | $T_e$ (eV) | $T_v$ (eV) | $R_e$ (Å) | $\omega_e$ ($cm^{-1}$) | $T_e$ (eV) | $T_v$ (eV) |
| DC-IHFSCCSD | X 3/2 | 1.875 | 722 | 0.00 | 0.00 | 1.973 | 644 | 0.00 | 0.00 |
| | a 1/2 | 1.887 | 702 | 0.28 | 0.28 | 2.018 | 585 | 0.70 | 0.72 |
| | a 3/2 | 2.095 | 492 | 2.76 | 3.22 | 2.209 | 435 | 2.34 | 2.78 |
| | b 1/2 | 2.077 | 497 | 3.09 | 3.48 | 2.166 | 445 | 2.78 | 3.07 |
| | c 1/2 | 2.022 | 514 | 4.02 | 4.22 | 2.108 | 493 | 4.97 | 5.12 |
| CCSD(T)[42] | X 3/2 | 1.872 | 684 | 0.00 | | | | | |
| | a 1/2 | 1.889 | 651 | 0.23 | | | | | |
| SO-MRCI[65] | X 3/2 | 1.922 | 650 | 0.00 | | | | | |
| | a 1/2 | 1.939 | 626 | 0.21 | | | | | |
| | b 1/2 | 2.249 | 288 | 1.69 | | | | | |
| | a 3/2 | 2.250 | 287 | 1.69 | | | | | |
| | b 3/2 | 2.115 | 514 | 2.82 | | | | | |
| Exp. | X 3/2 | 1.8677[a,c] | 681.6[a,c] | 0.00 | | | | | |
| | | 1.86762[b] | 681.7[b] | 0.00 | | | | | |
| | a 1/2 | 1.887[a] | 658[a] | 0.2593[a] | | | | | |
| | | 1.88468[b] | 645.3[b] | 0.2593[b] | | | | | |
| | a 3/2 | 2.072[d] | 514.5[d] | 2.6729[d] | | | | | |

[a] Ref. 66. [b] Ref. 67. [c] Ref. 68. [d] Ref. 71.

Peterson *et al.*[42] estimated spin–orbit effects with an uncontracted SO-MRCI calculations including spin–orbit coupling in a full-valence active space, omitting important contributions from single excitations.[21] In the SO-MRCI calculations of Roszak *et al.*[65] the RECPs differs from the one used by Peterson *et al.*[42] and us, and has been found by Roszak *et al.*[65] to underestimate by about 0.037 eV the spin–orbit splitting for the iodine $^2$P atomic ground state. The Gaunt interaction has previously been found to change the energy of this state by only about 0.004 eV,[11] so disregarding it here does not affect our conclusions on the accuracy of the DC-IHFSCCSD method.

Beyond the first two electronic states we obtained three additional states, with vertical excitation energies over 3 eV. These have significantly longer equilibrium distances than the former, and lower harmonic frequencies, reflecting their experimentally known predissociative character.[67,69] Of these, only the $^2\Pi_{3/2}$ has been measured experimentally,[68–71] notably by fluorescence studies ($T_e$ = 2.67 eV). This means that the apt comparison is between them and the DC-IHFSCCSD adiabatic value ($T_e$ = 2.82 eV; which, as for the vertical excitations, does not take into account any vibrational corrections). Doing so we obtain a discrepancy of 0.09 eV which is, as for the first excited state, about 0.06 eV smaller than the one obtained by Roszak *et al.* ($T_e$ = 2.82 eV).[65] Interestingly, here Roszak's harmonic frequencies show somewhat better agreement to experiment than ours, and discrepancies on bond lengths are now of the same order of magnitude (0.02 Å for DC-IHFSCCSD and 0.04 Å for SO-MRCI, respectively).

One should note in Roszak's calculations[65] that one finds some additional dissociative $\Omega$ = 1/2 electronic states, corresponding to configurations where the $\sigma_{1/2}$* spinor is also occupied by one electron coming from the $\pi$* spinor manifold. These states are, unfortunately, inaccessible with the present DC-IHFSCCSD calculations, since $\sigma_{1/2}$* could not be included in the model spaces considered here. However, as in Roszak *et al.* calculations they seem to be fairly well-separated from the other $\Omega$ = 1/2 state, we may speculate they do not significantly affect our results.

Finally, with respect to the ionization potential (IP) and electron affinity (EA) for IO, we observe very good agreement between DC-IHFSCCSD and experiment, shown in Table 2. However, while the calculated values for IP are in excellent agreement with the adiabatic ionization energy measured by Zhang *et al.*[74] from the recording of the photoionization efficiency spectra of IO, for EA the calculations are still 0.1 eV too low. The reason for this is likely due to an imbalance in the description of the $IO^-$ with respect to IO and $IO^+$, in particular, concerning the need of more flexible basis sets for the anion, as well as the lack of high-order excitations (triple, quadruple) in the coupled-cluster treatment. This is underscored by the EA value of 2.374 eV computed by Peterson *et al.*,[42] which is within the experimental error bars and by the fact that their estimate of SOC contributions (0.15 eV) is similar to the difference between our DC-IHFSCCSD and SFDC-IHFSCCSD results, the latter calculated at the ground-state bond length in Table 1.

Table 2 Vertical and adiabatic ionization potentials (IP, in eV) and electron affinities (EA, in eV) for IO and AtO, calculated using the SFDC-IHFSCCSD and DC-IHFSCCSD methods, along with experimental values for IO. Here "Vertical" values are calculated at the DC-IHFSCCSD ground-state equilibrium structures

| | | EA | |
|---|---|---|---|
| | Hamiltonian | Vertical | Adiabatic |
| IO | SFDC | 2.38 | |
| | DC | 2.23 | 2.28 |
| | Exp.[66,72] | 2.378 ± 0.005 | |
| AtO | SFDC | 2.33 | |
| | DC | 1.88 | 1.94 |
| | | IP | |
| | Hamiltonian | Vertical | Adiabatic |
| IO | SFDC | 9.79 | |
| | DC | 9.78 | 9.72 |
| | Exp.[73] | 9.66 ± 0.10 | |
| | Exp.[74] | | 9.735 ± 0.017 |
| AtO | SFDC | 9.46 | |
| | DC | 9.02 | 8.98 |

**3.1.2 Astatine oxide radical.** In Table 1 we present the spectroscopic constants and excitation energies obtained for AtO with the DC-IHFSCCSD. The composition of these states is roughly the same for both radicals, with $\Omega$ = 3/2 states being dominated by single determinants whereas $\Omega$ = 1/2 states exhibit some multideterminantal character (see Table S1 of the ESI†).

However, due to the larger magnitude of spin–orbit splitting for astatine than for iodine we see that the energy difference between ground and first excited state for AtO is more than twice that of IO, and almost as large as what would be expected from an estimation based on the atomic splitting (a factor of three).[43]

The increased spin–orbit splitting, which translates into a larger difference between the $\langle r \rangle$ values for the $p_{3/2}$ over the $p_{1/2}$ spinors, is also responsible for the roughly 0.1 Å increases in bond lengths of all electronic states; this is half of what would be expected purely by the change in radii between the atomic 5p and 6p shells.[75] This elongation is accompanied by an overall decrease in harmonic frequencies, indicating a weakening of the bond for AtO.

The importance of SOC for AtO is also clearly seen when comparing the EA and IP for DC and SFDC calculations, shown in Table 2. Whereas for the IO EA and IP one had, respectively, SOC contributions of 0.15 eV and 0.01 eV, for AtO these are both of the order of 0.45 eV, three times as large as in the lighter radical.

**3.1.3 The halide oxide cations.** The electronic spectra and spectroscopic constants for ground and excited states of $IO^+$ and $AtO^+$ are summarized in Table 3. Both molecules have been found to have a $\Omega = 0^+$ ground state, corresponding to one of the components of a $^3\Sigma^-$ state in LS coupling and whose configuration is predominantly one in which the $\pi_{3/2}$* is now empty (see Table S2 in the ESI†). Due to the smaller occupation of the

Table 3 Bond lengths ($R_e$ in Å), harmonic frequencies ($\omega_e$, in $cm^{-1}$), vertical ($T_v$, in eV) and adiabatic ($T_e$, in eV) excitation energies calculated using the DC-IHFSCCSD method for $XO^+$ (X = I, At), along with single-reference coupled cluster (DC-CCSD, DC-CCSD(T)) results for the ground state

| | | $IO^+$ | | | | $AtO^+$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\Omega$ | $R_e$ (Å) | $\omega_e$ ($cm^{-1}$) | $T_e$ (eV) | $T_v$ (eV) | $R_e$ (Å) | $\omega_e$ ($cm^{-1}$) | $T_e$ (eV) | $T_v$ (eV) |
| DC-IHFSCCSD | X $0^+$ | 1.812 | 818 | 0.00 | 0.00 | 1.916 | 707 | 0.00 | 0.00 |
| | a 1 | 1.812 | 821 | 0.13 | 0.13 | 1.939 | 646 | 0.63 | 0.63 |
| | a 2 | 1.835 | 750 | 0.71 | 0.72 | 1.986 | 562 | 0.97 | 1.02 |
| | a $0^+$ | 1.854 | (761)[a] | 1.35 | 1.38 | 2.132 | 523 | 1.44 | 1.94 |
| | a 3 | 2.030 | 575 | 1.84 | 2.45 | 2.133 | 519 | 1.48 | 1.98 |
| | b $0^-$ | 2.013 | (787)[a] | 1.77 | 2.39 | 1.944 | 672 | 2.74 | 2.75 |
| DC-CCSD | X $0^+$ | 1.809 | 825 | | | 1.903 | 730 | | |
| DC-CCSD(T) | X $0^+$ | 1.829 | 763 | | | 1.930 | 676 | | |

[a] Parentheses for states with an avoided crossing at 2.07 Å.



$\pi^*$ manifold for the ground-state, in comparison to IO and AtO, one can expect a strengthening of the X–O bond which is confirmed by shorter bond lengths (by about 0.05 Å for both species) and higher harmonic frequencies (by roughly 90 $cm^{-1}$ for $IO^+$ and 60 $cm^{-1}$ for $AtO^+$).

It is also interesting to note, from Table S2 in the ESI,† that while the ground states of both cations are dominated by a single configuration where the $\pi_{3/2}^*$ is unoccupied for $IO^+$ there is also an important (about 20%) contribution from a determinant in which $\pi_{1/2}^*$ is unoccupied, whereas for $AtO^+$ no such thing occurs, and instead one has contributions from configurations where the $\pi_{3/2}$ and $\pi_{3/2}^*$ are singly occupied. In any case, as noted by Rota *et al.*,[13] the dominant configuration is closed-shell with the $\pi_{1/2}^*$ spinor Kramers pair doubly occupied, making both cations relativistic closed-shell systems.

The (relativistic) single-reference character of the ground state, due to the variational inclusion of SOC, also has the advantage that standard approaches such as CCSD or CCSD(T) can be readily employed, something that may be interesting for subsequent studies of complexes of these cations. The values for the $T_1$ diagnostic in the DC-CCSD calculations, around the equilibrium geometry, are 0.018 for $AtO^+$ and 0.020 for $IO^+$, and these rise to 0.03 and 0.08 at $R$ = 2.00 Å, respectively. We thus present spectroscopic constants for both approaches in Table 3. We observe that the inclusion of triples markedly lowers the harmonic frequencies (by about 50–60 $cm^{-1}$ −1 in both cases), in line with the discussion for the radicals, and here increase the bond lengths by about 0.02 Å. Interestingly, we see that the DC-IHFSCCSD results lie roughly in between the DC-CCSD and DC-CCSD(T) results, something that indicates that the multi-reference character of the first acts in a way which partially mimics the effect of the perturbative triples.

Focusing now on the excited states, we see that the first ($\Omega$ = 1) correspond to the other components of the LS $^3\Sigma^-$ state. Due to changes in the magnitude of SOC in both species, we see that it goes from roughly 0.13 eV for $IO^+$ to 0.63 eV for $AtO^+$. Unlike the ground-state, its wavefunction for $IO^+$ has significantly more single-reference character than for $AtO^+$. The second ($\Omega$ = 2) state, which correlates with the $^1\Delta$ state in LS coupling, has somewhat less single-reference character than the previous states and, more importantly, is nearly 1 eV higher than the ground state for $AtO^+$. This is nearly twice as much as one would obtain in a spin-free calculation (see Table S3 in the ESI†) for the $^3\Sigma^- - ^1\Delta$ splitting, and might be of importance in the light of the mechanism proposed by Ayed *et al.*,[4,5] involving a change in ordering of these two states in solution.

The ($\Omega$ = 3) excited state is the only purely single-reference state for both species. It is close to two other $\Omega$ = 0 ($0^+$, $0^-$) states that, at the respective equilibrium geometries, change order for $AtO^+$ and $IO^+$. Among these, it is important to note that the wavefunction for the $\Omega$ = $0^+$ state is dominated by a doubly excited determinant with respect to the ground-state (see Table S2 of the ESI†).

### 3.2 Comparison to approximate methods

The DC-IHFSCCSD results discussed above give us confidence that the method can be used to benchmark the more approximate wavefunction (WFT)-based methods (MRCI, NEVPT2), as well as the DFT-based ones.

**3.2.1 WFT approaches.** As discussed before, the approximate WFT-based approaches considered here consist of two steps: in the first one obtains spin-free states, and in the second these are coupled by spin–orbit interactions. Therefore, before discussing the final outcome it is also of interest to briefly discuss the performance of NEVPT2 and MRCI + DC with respect to reference SFDC-IHFSCCSD results.

The results of these calculations, performed at the respective MRCI + DC equilibrium geometries for $IO^+$ ($R_e$ = 1.806 Å) and $AtO^+$ ($R_e$ = 1.903 Å), are summarized in Table S3 in the ESI.† There, we can see that all three methodologies yield spectra in very good agreement with each other, with discrepancies between methods usually being on the order of 0.1 eV, but often smaller for the lowest states. That makes the NEVPT2 approach particularly interesting in this context, as it is computationally much less costly than MRCI, by at least a factor of two.

It should be noted that some higher-lying $\Pi$ states ($^3\Pi$, $^1\Pi$) found using the approximate methods are absent from the SFDC-IHFSCCSD calculation. This is due to the fact that these states correspond to configurations for which the reference determinants have the $\sigma^*$ orbital occupied; orbital, which is not included in the SFDC-IHFSCCSD model space.

Table 4 Electronic excitation energies (in eV) for different correlation and SOC treatments for $IO^+$ and $AtO^+$ calculated at $R(IO^+)$ = 1.806 Å and $R(AtO^+)$ = 1.903 Å respectively. Here (u)c-SOCI denote (un)contracted SOCI, and the TDDDFT calculations employ the Tamm–Dancoff approximation (TDA)

| | | | | QD-NEVPT2 | | TDDFT/TDA | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Ω | IHFSCCSD | MRCI + DC c-SOCI | c-SOCI | uc-SOCI | SAOP | M06-L | M06 | M06-2X |
| $IO^+$ | X $0^+$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | a 1 | 0.13 | 0.10 | 0.09 | 0.09 | −0.19 | −0.21 | −0.33 | −0.45 |
| | a 2 | 0.72 | 0.69 | 0.68 | 0.69 | 0.59 | 0.57 | 0.51 | 0.48 |
| | a $0^+$ | 1.39 | 1.26 | 1.34 | 1.35 | [3.77][a] | [3.95][a] | [3.67][a] | [3.80][a] |
| | b $0^-$ | 2.42 | 2.44 | 2.54 | 2.52 | 2.70 | 2.89 | 1.94 | 1.26 |
| | a 3 | 2.48 | 2.51 | 2.62 | 2.57 | 2.71 | 2.90 | 1.93 | 1.23 |
| $AtO^+$ | X $0^+$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | a 1 | 0.64 | 0.53 | 0.54 | 0.49 | 0.47 | 0.41 | 0.29 | 0.25 |
| | a 2 | 1.05 | 0.97 | 0.99 | 0.93 | 1.12 | 1.09 | 1.01 | 1.02 |
| | a $0^+$ | 2.03 | 2.12 | 2.22 | 1.98 | [3.78][a] | [3.90][a] | [3.54][a] | [3.70][a] |
| | a 3 | 2.07 | 2.37 | 2.47 | 2.00 | 2.41 | 2.54 | 1.66 | 0.88 |
| | b $0^-$ | 2.77 | 2.29 | 2.32 | 2.45 | 2.40 | 2.53 | 1.66 | 0.90 |

[a] States with singly excited character, whereas WFT methods indicate doubly excited character.



Having shown that electron correlation can be properly captured by the NEVPT2 and MRCI approaches, we investigate the effect of including SOC by comparing the SOCI/MRCI + DC and SOCI/QD-NEVPT2 to DC-IHFSCCSD. The corresponding spectra, calculated at the equilibrium spin-free MRCI + DC geometries, are shown in Table 4. For NEVPT2 we also present vertical energies computed with a contracted SOCI approach with different diagonal elements of the spin–orbit coupling matrices.

In general, when spin–orbit coupling is treated *a posteriori*, it is recommended to use uncontracted SOCI methods to account for the spatial and energetic splitting of the open-shell spinors.[21,28,76,77] Our results show that these spin–orbit polarization effects are not important for $IO^+$, given that contracted and uncontracted SOCI/QD-NEVPT2 results are nearly identical. However, they are crucial in the case of $AtO^+$; we get smaller deviations with respect to the IHFSCCSD numbers with the uncontracted SOCI/QD-NEVPT2, than with the contracted SOCI/QD-NEVPT2. The largest deviation of 0.30 eV appears for the b $0^-$ state, a state which, according to the analysis of the SOCI/QD-NEVPT2 wave-function exhibits some contribution from the first $^3\Pi$ state that is absent from the IHFSCCSD calculation.

The conclusions drawn for these particular geometries are by and large valid along a wide range of internuclear distances, as one can see from the comparison of the SOCI/MRCI + DC and DC-IHFSCCSD potential energy curves in Fig. 2 (the SOCI/MRCI + DCs spectroscopic constants can be found in Table S4 in the ESI†). We see that for lower-lying states there is quantitative agreement between the two approaches, whereas for higher-lying states there are more significant differences not only in

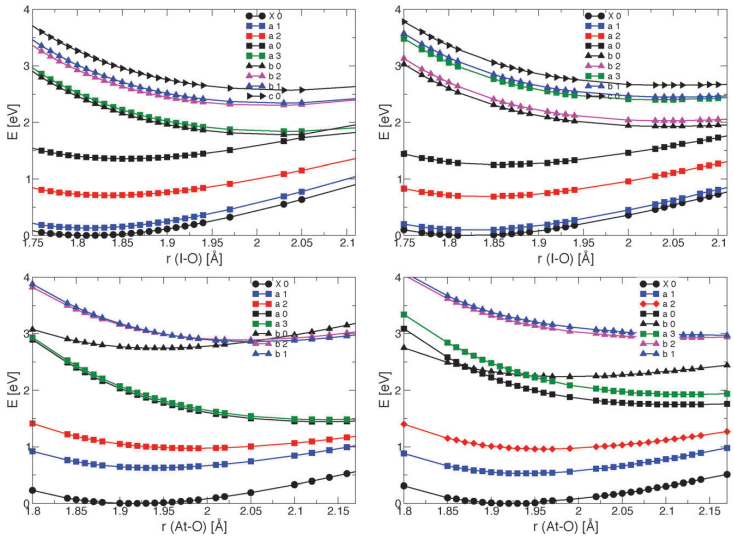



Fig. 2 Potential energy curves of the low-lying states of $IO^+$ (top) and $AtO^+$ (bottom) calculated using the IHFSCCSD (left) and SOCI/MRCI + DC (right) methods.

terms of energies but also for certain features such as avoided crossings, which are located at different places.

One important feature of the curves is that the energy differences between the $\Omega = 0^+$, 1, 2, which relate to the $^3\Sigma^-$ ground-state and first $^1\Delta$ excited state, and the higher-lying $\Omega = 0^+$ state, remain more or less constant over a wide range of internuclear distances, and the first three only approach each other towards the dissociation limit. The higher-lying $\Omega = 0^-$, 3 states, on the other hand, are very much destabilized compared to the first three for shorted bond lengths, but approach the others towards the dissociation limit.

It is also interesting to note that the $\Omega = 3$, $0^-$ excited states are reversed in $IO^+$ and $AtO^+$, and that is captured by both approaches.

**3.2.2 Density functional approximations.** The calculated spectra with different DFAs are also shown in Table 4. There, one may note that only TDDFT results employing the Tamm–Dancoff approximation (TDDFT/TDA) are shown. The reason for doing so is that, in the course of our TDDFT calculations (whose results are shown in Table S5 in the ESI†), we have detected the presence of triplet instabilities[78–83] for all DFAs considered and for both molecules.

It can be clearly seen from the results in Table S5 that such instabilities manifest themselves here by strongly stabilizing the first $\Omega = 1$ excited state with respect to the $\Omega = 0$ ground-state, and that such stabilization becomes stronger as the fraction of Hartree–Fock exchange employed is increased, as one goes from M06-L (0%) to M06 (27%) and M06-2X (54%). Additionally, for the latter a dramatic change occurs in which almost all of the low-lying states considered become stabilized with respect to the ground-state, a situation clearly at odds with the reference WFT results. While TDDFT is clearly unreliable for these particular systems, a number of recent studies suggest the use of TDDFT/TDA as a pragmatic way to circumvent these instabilities while improving excitation energies.[84–87]

The results in Table 4 indicate that this is indeed the case here. We observe, for $AtO^+$, that both the SAOP model potential and the M06-L functional yield results which are usually within 0.1 eV of the WFT results for the $\Omega = 1$, 2 excited states, while showing slightly larger discrepancies for the $\Omega = 3$, $0^-$ states. The $\Omega = 0^+$ excited state, on the other hand, is quite far from the WFT results, but this is due to the fact that the two sets of calculations do not represent the same state; in the WFT ones, this corresponds to a doubly-excited state whereas TDDFT, irrespective of the use of the Tamm–Dancoff approximation, can only represent singly excited states in the ALDA approximation. The performance of SAOP and M06-L for $IO^+$ is similar to that of $AtO^+$. There, however, the DFAs are not able to properly capture the small splitting between the ground-state and the $\Omega = 1$ excited-state, effectively placing the latter below the former.

As was the case for plain TDDFT calculations, the TDDFT/TDA results for DFAs containing Hartree–Fock exchange show a tendency to spuriously stabilize most the excited states; for instance, the splitting between the ground-state and the $\Omega = 1$ excited-state for $AtO^+$ is nearly halved when passing from SAOP or M06-L to M06-2X, as are the energies for the $\Omega = 3$, $0^-$ states. Interestingly, the $\Omega = 2$ state, which relates to the $^1\Delta$ state in LS coupling, does not show a similar reduction, and remains roughly 1 eV (0.5 eV) higher than the $\Omega = 0$ ground-state of $AtO^+$ ($IO^+$).

### 3.3 The $AtO^+$–water complex

While the results for the isolated molecules indicate a clear hierarchy of methods with respect to the reliability of the calculations, and put into question the ability of DFAs in general to properly describe the electronic structure of $AtO^+$, one can argue that interest ultimately lies in describing the species in solution.

As a step towards that, we investigate here the electronic structure of a complex between $AtO^+$ and a single water molecule. We note, however, that an extensive evaluation along the lines of that of Ayed *et al.*[4,5] is beyond the scope of our work, due to constraints on computational resources currently available to us.

The optimized structure for this complex (in $C_{2v}$ symmetry) can be found in Table S6 in the ESI.† The distance between astatine and the oxygen in the water molecule (2.655 Å) is somewhat longer (by 0.08 Å) than that found by Ayed *et al.*[4,5] for their spin-free triplet electronic ground-state and much longer (by 0.289 Å) than the structure for the singlet state. The At–O distance (1.900 Å), on the other hand, is nearly the same as the optimized geometry for the isolated molecule with the M06-2X functional (1.885 Å, compared to 1.866 Å for a spin-free unrestricted calculation with two unpaired electrons), indicating that the perturbation to the bonding due to the presence of one single water is not very significant. This distance is nevertheless smaller (by 0.054 Å) than that found Ayed *et al.*[4,5] for the triplet electronic ground-state, but nearly identical to that found for the singlet structure (1.901 Å). This difference is probably due to use of different basis sets and PPs *vs.* frozen-core AE.

That the water does not strongly perturb the electronic structure of $AtO^+$ can also be seen from the electronic spectra for the complex, calculated using different approaches and whose results are found in Table 5, in comparison to the results for the isolated molecule in Table 4 (at a slightly different internuclear distance). We observe that the uncontracted-SOCI/QD-NEVPT2 places the low-lying electronic states at roughly the same transition energies, even if one can observe a small

Table 5 Electronic excitation energies (in eV) for the $AtO^+$–water complex. All TDDFT calculations employ the Tamm–Dancoff approximation

| $\Omega(C_{\infty v})$ | $C_{2v}$ | NEVPT2 | SAOP | M06-L | M06 | M06-2X |
|---|---|---|---|---|---|---|
| X $0^+$ | X $A_1$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| a 1 | a $B_1$ | 0.36 | 0.43 | 0.37 | 0.24 | 0.15 |
| | a $B_2$ | 0.36 | 0.43 | 0.37 | 0.24 | 0.15 |
| a 2 | a $A_2$ | 0.83 | 1.08 | 1.05 | 0.96 | 0.93 |
| | a $A_1$ | 0.84 | 1.08 | 1.05 | 0.96 | 0.93 |
| a $0^+$ | b $A_1$ | 1.97 | [2.85][a] | [2.73][a] | [2.67][a] | [2.26][a] |
| a 3 | b $B_2$ | 2.02 | 2.38 | 2.34 | 1.65 | 0.86 |
| | b $B_1$ | 2.04 | 2.38 | 2.34 | 1.65 | 0.86 |
| b $0^-$ | b $A_2$ | 2.14 | 2.37 | 2.35 | 1.65 | 0.88 |

[a] States with singly excited character, whereas WFT methods indicate doubly excited character.

downwards shift in the $AtO^+$–water complex, of about 0.13 eV for the lowest states correlated with the X $0^+$, a 1, a 2 states.

The TDDFT/TDA calculations also present trends which are much in line with those already discussed for the isolated molecule: first, a good agreement between the SAOP model potential and the M06-L functional, and those two with wavefunction-based results, whereas for the higher states there is a larger discrepancy between the isolated molecule and the complex for M06-L than for SAOP.

For DFAs including Hartree–Fock exchange (M06, M06-2X), on the other hand, we observe that apart from the state related to the $\Omega = 2$ state in the isolated species, there is a more pronounced decrease of the excitation energies of the lowest-lying excited states, as the amount of Hartree–Fock exchange is increased, for the $AtO^+$–water complex. The states corresponding to the $\Omega = 3$, $0^-$, however, appear to change very little between the isolated $AtO^+$ and $AtO^+$–water complex.

That certain states are affected very differently for the isolated and $AtO^+$–water complexes depending on the amount of Hartree–Fock exchange in the calculation, and do so in disagreement with correlated wavefunction-based approaches, does not bode well for the application of such functionals to solvation studies on $AtO^+$, and poses questions as to the accuracy of the conclusions of Ayed *et al.*[4,5]

## 4 Conclusion

In this work we have studied the electronic structure of the XO and $XO^+$ (X = I, At) species, as well as the complex between $AtO^+$ and a water molecule, with wavefunction-based correlated methods as well as density functional theory-based approaches, both with and without the inclusion of spin–orbit coupling (SOC) interactions, with the aim of establishing which electronic structure approach is the most suitable one for the study of $AtO^+$ in the gas phase and in solution.

Since little experimental data is available for At-containing species, we have first benchmarked the DC-IHFSCCSD against experimental results and prior calculations for iodine-containing species, and found that the method can reproduce with a reasonable accuracy the experimental bond lengths (to $\simeq$0.01 Å), excitation energies (to $\simeq$0.05 eV), electron affinities and ionization potentials (to $\simeq$0.1 eV) while overestimating harmonic frequencies (by about $\simeq$40 $cm^{-1}$). With that, we believe it is reasonable to employ the DC-IHFSCCSD results to assess the accuracy of other methods for $AtO^+$.

Furthermore, these reference calculations on the isolated cationic species indicate that due to SOC the stabilization of $\pi_{1/2}^*$ spinors with respect to $\pi_{3/2}^*$ is such that electrons will preferentially occupy the latter and one is justified, for $IO^+$, $AtO^+$ and the $AtO^+$–water complex, to consider the ground-state as a (relativistic) closed-shell reference. This allows the use of standard single-reference approaches such as CCSD and CCSD(T), or approaches based on a linear response formalism for excited states. The strong SOC effects for $AtO^+$ also means that it makes little physical sense to discuss the electronic structure of these systems in a spin-free formalism.

The approximate wavefunction approaches, which took SOC into account perturbatively, yielded results which were generally very close to the reference calculations (*e.g.* usually within 0.1 eV), with NEVPT2 showing an overall identical performance than MRCI, but at a significantly lower computational cost. This makes NEVPT2 the method of choice, among those considered, for the study of larger systems. It is fundamental, on the other hand, to include spin–orbit polarization effects in the SOCI step in order to better match the DC-IHFSCCSD results for $AtO^+$.

The performance of DFAs, which included SOC at the ZORA level for $AtO^+$ and $IO^+$ is, on the other hand, rather disappointing for approaches including Hartree–Fock exchange, as one finds that the higher the fraction of Hartree–Fock exchange used the more severely the low-lying excited states are underestimated, both for the isolated species and in the case of a $AtO^+$–water complex.

In particular, our findings suggest that the M06-2X functional, employed in prior studies on the solvation of $AtO^+$, is the least suitable of the DFAs investigated here since for it such spurious stabilization of the excited states is largest. The SAOP model potential and the M06-L functional, on the other hand, yielded excitation energies in good agreement with the wavefunction-based approaches, but like all other DFAs considered, suffer from a triplet instability problem that places the $\Omega = 1$ excited state below the $\Omega = 0$ ground-state in TDDFT calculations, effectively calling for the use of the Tamm–Dancoff approximation in order to obtain reliable results. We believe, therefore, that in situations where the SOCI/NEVPT2 approach would be too costly, a protocol based on TDA calculations with either of these functionals might provide results that are likely to be (at least) qualitatively correct.

## Acknowledgements

Most of the computations were performed on the large scale facilities of the "Centre de Calcul et de Recherche Technologique" (CCRT, grants 2012-081859 and 2013-081859) of the French Nuclear Agency (CEA). Additional computational resources have also been provided by the "Institut de Développement et de Ressources en Informatique Scientifique du Centre National de la Recherche Scientifique" (IDRIS-CNRS, grant 71859) and by the "Centre Informatique National de l'Enseignement Supérieur" (CINES, Project No. phl2531). The authors are thankful for the support from the French National Research Agency under the contracts ANR 2010-BLAN-0807-01 and ANR-11-LABX-0005 Chemical and Physical Properties of the Atmosphere (CaPPA).

## References

1 D. Scott Wilbur, *Nat. Chem.*, 2013, **5**, 246.
2 J. Champion, C. Alliot, E. Renault, B. M. Mokili, M. Chérel, N. Galland and G. Montavon, *J. Phys. Chem. A*, 2010, **114**, 576.
3 J. Champion, A. Sabatié-Gogova, F. Bassal, T. Ayed, C. Alliot, N. Galland and G. Montavon, *J. Phys. Chem. A*, 2013, **117**, 1983.

4 T. Ayed, M. Seydou, F. Réal, G. Montavon and N. Galland, *J. Phys. Chem. B*, 2013, **117**, 5206.
5 T. Ayed, F. Réal, G. Montavon and N. Galland, *J. Phys. Chem. B*, 2013, **117**, 10589–10595.
6 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Science*, 2008, **321**, 792.
7 A. M. Teale, F. De Proft and D. J. Tozer, *J. Chem. Phys.*, 2008, **129**, 044110.
8 W. Yang, A. J. Cohen and P. Mori-Sánchez, *J. Chem. Phys.*, 2012, **136**, 204111.
9 M. Swart, A. R. Groenhof, A. W. Ehlers and K. Lammertsma, *J. Phys. Chem. A*, 2004, **108**, 5479.
10 C. P. Constantinides, P. A. Koutentis and J. Schatz, *J. Am. Chem. Soc.*, 2004, **126**, 16232.
11 J. Sikkema, L. Visscher, T. Saue and M. Iliaš, *J. Chem. Phys.*, 2009, **131**, 124116.
12 A. S. P. Gomes, L. Visscher, H. Bolvin, T. Saue, S. Knecht, T. Fleig and E. Eliav, *J. Chem. Phys.*, 2010, **133**, 064305.
13 J.-B. Rota, S. Knecht, T. Fleig, D. Ganyushin, T. Saue, F. Neese and H. Bolvin, *J. Chem. Phys.*, 2011, **135**, 114106.
14 A. S. P. Gomes and L. Visscher, *Chem. Phys. Lett.*, 2004, **399**, 1.
15 C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger and J.-P. Malrieu, *J. Chem. Phys.*, 2001, **114**, 10252.
16 C. Angeli, R. Cimiraglia and J.-P. Malrieu, *Chem. Phys. Lett.*, 2001, **350**, 297.
17 C. Angeli, R. Cimiraglia and J.-P. Malrieu, *J. Chem. Phys.*, 2002, **117**, 9138.
18 C. Angeli, S. Borini, M. Cestari and R. Cimiraglia, *J. Chem. Phys.*, 2004, **121**, 4043.
19 C. Angeli, M. Pastore and R. Cimiraglia, *Theor. Chem. Acc.*, 2007, **117**, 743.
20 B. A. Heß and C. M. Marian, Relativistic effects in the calculation of electronic energies, in *Computational Molecular Spectroscopy*, ed. P. Jensen and P. R. Bunker, John Wiley & Sons, Sussex, 2000, pp. 169–219.
21 V. Vallet, L. Maron, C. Teichteil and J.-P. Flament, *J. Chem. Phys.*, 2000, **113**, 1391.
22 K. G. Dyall, *J. Chem. Phys.*, 1995, **102**, 4909.
23 I. Infante, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2006, **125**, 074301.
24 I. Infante, E. Eliav, M. J. Vilkas, Y. Ishikawa, U. Kaldor and L. Visscher, *J. Chem. Phys.*, 2007, **127**, 124308.
25 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353.
26 F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet and E. Eliav, *J. Phys. Chem. A*, 2009, **113**, 12504.
27 F. Ruipérez, C. Danilo, F. Réal, J.-P. Flament, V. Vallet and U. Wahlgren, *J. Phys. Chem. A*, 2009, **113**, 1420.
28 A. Weigand, X. Cao, V. Vallet, J.-P. Flament and M. Dolg, *J. Phys. Chem. A*, 2009, **113**, 11509.
29 P. Tecmer, A. S. P. Gomes, U. Ekström and L. Visscher, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6249.
30 A. S. P. Gomes, C. R. Jacob, F. Réal, V. Vallet and L. Visscher, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15153.
31 DIRAC, a relativistic ab initio electronic structure program, Release DIRAC11 (2011), written by R. Bast, H. J. A. a. Jensen, T. Saue and L. Visscher, with contributions from V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, T. Fleig, O. Fossgaard, A. S. P. Gomes, T. Helgaker, J. K. Lærdahl, J. Henriksson, M. Iliaš, C. h. R. Jacob, S. Knecht, C. V. Larsen, H. S. Nataraj, P. Norman, G. Olejniczak, J. Olsen, J. K. Pedersen, M. Pernpointner, K. Ruud, P. Sałek, B. Schimmelpfennig, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, S. Villaume, O. Visser, T. Winther, and S. Yamamoto (see http://dirac.chem.vu.nl).
32 K. G. Dyall, *Theor. Chem. Acc.*, 2002, **108**, 335.
33 K. G. Dyall, *Theor. Chem. Acc.*, 2003, **109**, 284.
34 R. A. Kendall, T. H. Dunning, Jr and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.
35 L. Visscher, *Theor. Chem. Acc.*, 1997, **98**, 68.
36 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2000, **113**, 9905.
37 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 6862.
38 L. Visscher, E. Eliav and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 9720.
39 L. Visscher, T. J. Lee and K. G. Dyall, *J. Chem. Phys.*, 1996, **105**, 8769.
40 K. G. Dyall, *J. Chem. Phys.*, 1994, **100**, 2118.
41 L. Visscher and T. Saue, *J. Chem. Phys.*, 2000, **113**, 3996.
42 K. A. Peterson, B. C. Shepler, D. Figgen and H. Stoll, *J. Phys. Chem. A*, 2006, **110**, 13877.
43 K. A. Peterson, D. Figgen, E. Goll, H. Stoll and M. Dolg, *J. Chem. Phys.*, 2003, **119**, 11113.
44 K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, *J. Phys. Chem.*, 1990, **94**, 5483.
45 H.-J. Werner and W. Meyer, *J. Chem. Phys.*, 1980, **73**, 2342.
46 H.-J. Werner and W. Meyer, *J. Chem. Phys.*, 1980, **74**, 5794.
47 H.-J. Werner, *Adv. Chem. Phys.*, 1987, **1**, LXIX.
48 H.-J. Werner and P. Knowles, *J. Chem. Phys.*, 1988, **89**, 5803.
49 P. Knowles and H.-J. Werner, *Chem. Phys. Lett.*, 1988, **145**, 514.
50 S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.*, 1974, **8**, 61.
51 Molpro, version 2012.1, a package of ab initio programs, 2012, see http://www.molpro.net.
52 R. Llusar, M. Casarrubios, Z. Barandiarán and L. Seijo, *J. Chem. Phys.*, 1996, **105**, 5321.
53 G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P. O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady and L. Seijo, *Comput. Mater. Sci.*, 2003, **28**, 222, DOI: 10.1016/S0927-0256(03)00109-5.
54 C. Fonseca Guerra, J. G. Snijders, G. te Velde and E. J. Baerends, *Theor. Chem. Acc.*, 1998, **99**, 391.
55 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931.
56 Adf2013, http://www.scm.com.
57 E. van Lenthe, E. J. Baerends and J. G. Snidjers, *J. Chem. Phys.*, 1993, **99**, 4597.
58 E. van Lenthe, E. J. Baerends and J. G. Snidjers, *J. Chem. Phys.*, 1994, **101**, 9783.

59 E. van Lenthe, A. Ehlers and E. J. Baerends, *J. Chem. Phys.*, 1999, **110**, 8943.
60 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215.
61 P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344.
62 O. V. Gritsenko, P. R. T. Schipper and E. J. Baerends, *Chem. Phys. Lett.*, 1999, **302**, 199.
63 O. V. Gritsenko, P. R. T. Schipper and E. J. Baerends, *Int. J. Quantum Chem.*, 2000, **76**, 407.
64 TRANSO is a companion program to EPCISO[21] computing observables and transition matrix elements for spin–orbit free and spin–orbit states.
65 S. Roszak, M. Krauss, A. B. Alekseyev, H.-P. Liebermann and R. J. Buenker, *J. Phys. Chem. A*, 2000, **104**, 2999.
66 M. K. Gilles, M. L. Polak and W. C. Lineberger, *J. Chem. Phys.*, 1991, **95**, 4723.
67 C. E. Miller and E. A. Cohen, *J. Chem. Phys.*, 2001, **115**, 6459.
68 J. P. Bekooy, W. L. Meerts and A. Dymanus, *J. Mol. Spectrosc.*, 1983, **102**, 320.
69 T. J. Gravestock, M. A. Blitz and D. E. Heard, *Phys. Chem. Chem. Phys.*, 2010, **12**, 823.
70 S. M. Newman, W. H. Howie, I. C. Lane, M. R. Upson and A. J. Orr-Ewing, *J. Chem. Soc., Faraday Trans.*, 1998, **94**, 2681.
71 R. A. Durie, F. Legay and D. A. Ramsay, *Can. J. Phys.*, 1960, **38**, 444.
72 M. K. Gilles, M. L. Polak and W. C. Lineberger, *J. Chem. Phys.*, 1992, **96**, 8012.
73 P. S. Monks, L. J. Stief, D. C. Tardy, J. F. Liebman, Z. Zhang, S.-C. Kuo and R. B. Klemm, *J. Chem. Phys.*, 1995, **99**, 16566.
74 Z. Zhang, P. S. Monks, L. J. Stief, J. F. Liebman, R. E. Huie, S.-C. Kuo and R. B. Klemm, *J. Phys. Chem.*, 1996, **100**, 63.
75 J. P. Desclaux, *At. Data Nucl. Data Tables*, 1973, **12**, 311.
76 E. Fromager, V. Vallet, B. Schimmelpfennig, P. Macak, T. Privalov and U. Wahlgren, *J. Phys. Chem. A*, 2005, **109**, 4957.
77 C. Danilo, V. Vallet, J.-P. Flament and U. Wahlgren, *J. Chem. Phys.*, 2008, **128**, 154310.
78 W. H. Adams, *Phys. Rev.*, 1962, **127**, 1650.
79 J. Čížek and J. Paldus, *J. Chem. Phys.*, 1967, **47**, 3976.
80 R. Seeger and J. A. Pople, *J. Chem. Phys.*, 1977, **66**, 3045.
81 R. Bauernschmitt and R. Ahlrichs, *J. Chem. Phys.*, 1996, **104**, 9047.
82 K. J. H. Giesbertz and E. J. Baerends, *Mol. Phys.*, 2010, **108**, 2579.
83 O. B. Lutnæs, T. Helgaker and M. Jaszuński, *Mol. Phys.*, 2010, **108**, 2579.
84 M. J. G. Peach, M. J. Williamson and D. J. Tozer, *J. Chem. Theory Comput.*, 2011, **7**, 3578.
85 J. S. Sears, T. Koerzdoerfer, C.-R. Zhang and J.-L. Brédas, *J. Chem. Phys.*, 2011, **135**, 151103.
86 M. J. G. Peach and D. J. Tozer, *J. Phys. Chem. A*, 2012, **116**, 9783.
87 M. J. G. Peach, N. Warner and D. J. Tozer, *Mol. Phys.*, 2013, **111**, 1271.

## C.7 Paper VI

IOP Publishing New J. Phys. 17 (2015) 043005 doi:10.1088/1367-2630/17/4/043005

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
6 December 2014

REVISED
23 February 2015

ACCEPTED FOR PUBLICATION
2 March 2015

PUBLISHED
2 April 2015



PAPER

# Theoretical study on $ThF^+$, a prospective system in search of time-reversal violation

**Malika Denis**[1], **Morten S Nørby**[2], **Hans Jørgen Aa Jensen**[2], **André Severo Pereira Gomes**[3], **Malaya K Nayak**[1,4], **Stefan Knecht**[5] **and Timo Fleig**[1]

[1] Laboratoire de Chimie et Physique Quantiques, IRSAMC, Université Paul Sabatier Toulouse III, 118 Route de Narbonne, F-31062 Toulouse, France
[2] University of Southern Denmark, Department of Physics, Chemistry and Pharmacy, Campusvej 55, DK-5230, Odense, Denmark
[3] Laboratoire de Physique des Lasers, Atomes et Molécules (PhLAM), CNRS UMR 8523, Université de Lille, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France
[4] Bhabha Atomic Research Centre, Trombay, Mumbai-400085, India
[5] ETH Zürich, Laboratorium für Physikalische Chemie, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland

**E-mail:** stefan.knecht@phys.chem.ethz.ch **and** timo.fleig@irsamc.ups-tlse.fr

**Keywords:** electron electric dipole moment, *P, T* symmetry violation, relativistic electronic structure, hyperfine interaction, electronically excited states, wavefunction theory

## Abstract

The low-lying electronic states of $ThF^+$, a possible candidate in the search for $\mathcal{P}$- and $\mathcal{T}$-violation, have been studied using high-level correlated relativistic *ab initio* multi-reference coupled-cluster and configuration interaction approaches. For the $^3\Delta$ state component with $\Omega = 1$ (electron electric dipole moment 'science state') we obtain an effective electric field of $E_{\text{eff}} = 35.2$ GV cm$^{-1}$, a $\mathcal{P}$- and $\mathcal{T}$-odd electron–nucleon interaction constant of $W_{P,T} = 48.4$ kHz, a magnetic hyperfine interaction constant of $A_{\parallel} = 1833$ MHz for $^{229}$Th ($I = 5/2$), and a very large molecular dipole moment of 4.03 D. The $\Omega = 1$ state is found to be more than 300 cm$^{-1}$ lower in energy than $\Omega = 0^+$ ($^1\Sigma^+$), challenging the state assignment from an earlier theoretical study on this species (Barker *et al* 2012 *J. Chem. Phys.* **136** 104305).

## 1. Introduction

The enormous surplus of matter over antimatter in our Universe is a fact that remains unexplained by the standard model (SM) of elementary particles [1]. A microscopic violation of the combined symmetries charge ($C$) conjugation and spatial parity ($\mathcal{P}$) has been identified as one of several conditions [2] which can give rise to an appreciable baryon number and explain this asymmetry. It is expected that flavor-diagonal $\mathcal{CP}$ violation, absent in the SM, must be sought for [3] and that electric dipole moments (EDMs) [4] constitute a sensitive probe of such new physics beyond the SM. Given the validity of the $\mathcal{CPT}$ theorem [5], the measurement of an EDM would be the first direct signature of the violation of time-reversal ($\mathcal{T}$) invariance [6].

The electron's EDM has, despite a vigorous search for over half a century, still not been detected. The most constraining upper bounds on the electron EDM have for some time been obtained from experimental and theoretical investigations on atoms [7, 8], and such upper bounds are useful guiding constraints on beyond SM theories [9]. However, polar diatomic molecules have become the major players in this quest, since they offer an orders of magnitude larger enhancement [7, 10] of the ensuing energy shift than what could be achieved with an atom [11, 12]. This means that, for a given measurement on a molecular system, the possible magnitude of the electron EDM is constrained to a smaller value, or conversely, that the effect of a smaller electron EDM can be detected through the measurement. The corresponding enhancement factor is not accessible by experimental means and has to be determined—preferrably—via a molecular relativistic many-body calculation.

According to the most recent findings using the polar molecule ThO [13–16] the upper bound on the electron EDM is $|d_e| < 9.6 \times 10^{-29}$ $e$ cm. This value is more than 16 times smaller than the most constraining upper bound from an atomic study [12]. Charged molecules offer an experimental advantage over neutral

systems in that ion traps can be used which allow for long interrogation times. High-resolution spectroscopy employing rotating electric fields has been presented recently as a viable technique for symmetry violation searches in charged molecular ions [17, 18]. The ionic systems used in these experiments are $HfF^+$ and, as a perspective molecular ion, $ThF^+$.

What the afore-mentioned molecules, and several others such as $HfH^+$, $PtH^+$ [19] and WC [20, 21], have in common is an energetically low-lying $^3\Delta$ electronic state (in $\Lambda$–$S$ coupling picture). In the fluorides and oxides this state is deeply bound which is an experimental advantage. The magnetic moment in the $\Omega = 1$ component of this term is approximately zero which helps reduce the vulnerability of the experiment to decoherence and systematic errors [17].

$HfF^+$ and $ThF^+$ exhibit a considerably large EDM effective electric field in the relevant 'science' state [22–24] and, at the same time, a small $\Lambda$ (or $\Omega$) doublet splitting. This latter property is an asset for efficient mixing of rotational parity eigenstates through the external electric laboratory field. While $HfF^+$ has been characterized in detail [22, 23, 25–28] considerably less is known for $ThF^+$ [24, 29, 30]. The joint experimental and theoretical work of Barker *et al* [29] left some uncertainty as to whether the $\Omega = 1$ state is the ground-state or the first excited state, as there is an $\Omega = 0^+$ state ($^1\Sigma_0^+$) separated from it by only 315 $cm^{-1}$. The experimental resolution was not sufficient to unequivocally assign those states and, unlike $HfF^+$, the $\Omega = 1$ and $0^+$ states of $ThF^+$ possess similar vibrational frequencies at around 658 $cm^{-1}$. Accompanying theoretical calculations were also inconclusive, but from the best estimate the $\Omega = 0^+$ state was proposed as ground state with the $\Omega = 1$ state higher by 65 $cm^{-1}$ in [29] and 202 $cm^{-1}$ in [31], respectively.

Turning to the EDM effective electric field in $\Omega = 1$ of $ThF^+$, the work of Meyer *et al* [24] suggests an extremely large value of $E_{\rm eff} = 90$ GV $cm^{-1}$. Recent and more rigorous relativistic many-body calculations on the isoelectronic ThO molecule have shown [14, 15] that the model calculation of Meyer *et al* yields a signifcantly overestimated $E_{\rm eff}$ for the case of ThO (by more than 35%). It can therefore be expected that for this kind of molecules and electronic states the model of Meyer *et al* contains a systematic error that is also present in the above prediction for $E_{\rm eff}$ in $ThF^+$.

We pursue two major goals in this work. Using spinor-based many-body methods which treat dynamic electron correlations and electronic spin–orbit interactions on the same footing, a rigorous determination of the electronic ground state of $ThF^+$ and some of its low-lying electronically excited states is carried out. Second, with the same uncompromising techniques we determine with high accuracy properties of the $^3\Delta_1$ ($\Omega = 1$) state which are of direct relevance for proposed measurements of the electron EDM. In particular, we present the first rigorous calculation of the eEDM $\mathcal{P}, \mathcal{T}$-odd interaction constant and of the molecular static electric dipole moment. Furthermore, the magnetic hyperfine interaction constant is calculated for $\Omega = 1$ along with the scalar–pseudoscalar $\mathcal{P}, \mathcal{T}$-odd interaction constant, both of which play an important role in the interpretation of corresponding and ongoing experimental measurements [18, 32]. We also calculated static molecular EDMs and electric transition moments, the latter of which are of interest regarding state preparation in an EDM experiment. Molecular dipole moments are directly related to the EDM effective electric field, since they are a measure of the mixing of parity eigenstates.

The manuscript is structured as follows: section 2 summarizes underlying theory and gives a concise account of the employed electronic-structure methods. In section 3 we present results and their discussion, preceded by a brief summary of relevant technical details for the calculations that have been carried out. In the final section 4 we summarize the major findings and draw conclusions for future work.

## 2. Theory and methods

### 2.1. Theory

The potential energy due to the electron EDM interaction in the molecule is determined as an expectation value over the effective one-body Hamiltonian $\hat{H}_{\rm edm}^{\rm eff}$ in accord with stratagem II of Lindroth *et al* [33]

$$\left\langle \hat{H}_{\rm edm}^{\rm eff} \right\rangle_\psi = \frac{2ic d_e}{e\hbar} \left\langle \sum_{j=1}^{n} \gamma_j^0 \gamma_j^5 \, \mathbf{p}_j^2 \right\rangle_\psi , \tag{1}$$

where $n$ is the number of electrons, $\gamma$ are the standard Dirac matrices, $d_e$ is the electron electric dipole moment, and $\mathbf{p}_j$ the momentum operator for electron $j$.

The parallel magnetic hyperfine interaction constant $A_\parallel$ is defined as the $z$ projection of the expectation value of the corresponding perturbative Hamiltonian in Dirac theory

$$A_{\parallel} = \frac{\mu_{\mathrm{Th}}}{I\Omega}\left\langle \sum_{i=1}^{n}\left(\frac{\vec{\alpha}_i \times \vec{r}_i}{r_i^3}\right)_z \right\rangle_{\psi}, \tag{2}$$

where $I$ is the nuclear spin quantum number, $\alpha_k$ is a Hamiltonian-form Dirac matrix for particle $k$, and $n$ is the number of electrons. Details on the implementation of the two afore-mentioned operators can be found in references [14, 23].

The $\mathcal{P}$, $\mathcal{T}$-odd interaction constant $W_{P,T}$ arising from the electron–nucleon scalar–pseudoscalar (S–PS) interaction is determined as the expectation value over the $\mathcal{P}$, $\mathcal{T}$-odd Hamiltonian

$$W_{P,T} = \frac{1}{k_s \Omega}\langle H_s \rangle_{\psi}, \tag{3}$$

where $k_s$ is the electron–nucleus S–PS coupling constant. The interaction Hamiltonian $H_s$ is defined [34] as

$$H_s = \imath\frac{G_{\mathrm{F}}}{\sqrt{2}} Z k_s \sum_{j=1}^{n} \gamma_j^0 \gamma_j^5 \rho_N(r_j), \tag{4}$$

where $\rho_N(r_j)$ is the nuclear charge density at position $r_j$ normalized to unity, $G_{\mathrm{F}}$ is the Fermi constant and $k_s$ is a dimensionless S–PS interaction constant. The latter is defined as $Zk_s = (Zk_{s,p} + Nk_{s,n})$, where $k_{s,p}$ and $k_{s,n}$ are electron–proton and electron–neutron coupling constants, respectively. Equations (2) and (3) are evaluated as expectation values over the CI wavefunction for the $\psi_{\Omega=1}$ state.

### 2.2. Methods

All approaches employed in the present work are spinor-based molecular many-body methods in the framework of the four-component *no-virtual-pair* approximation (see [35] and references therein). For the treatment of dynamic interelectronic correlations linear and nonlinear (exponential) expansions of the molecular wavefunctions have been employed. As to the latter, we applied the coupled cluster approach [36] in a modern implementation, the intermediate Hamiltonian Fock-space coupled cluster including single and double excitations (IHFSCC) [37–41]. The linear expansions have been carried out with the general-active-space configuration interaction (GASCI) method [42, 43], implemented as the `KR-CI` module [44] in the `DIRAC12` [45] relativistic electronic-structure program package.

## 3. Application to $ThF^+$

### 3.1. Technical details

Calculations of spectroscopic properties were performed with the `DIRAC12` program [45], except for vertical excitation energies of the $\Omega = 1$ state using refined active spinor spaces, which were carried out with a modified local version of the `DIRAC11` program package [46]. This latter program version was also used for determining all $\mathcal{P}$, $\mathcal{T}$-odd and magnetic hyperfine expectation values.

We employed uncontracted atomic Gaussian basis sets for the description of both atoms' electronic shells. For thorium, Dyall's [47, 48] double-$\zeta$ (DZ, dyall.cv2z, [26s23p17d13f1g]) triple-$\zeta$ (TZ, dyall.cv3z, [33s29p20d15f5g1h] for IHFSCC and CI models $\mathcal{I}^{CI}$ and $\mathcal{II}^{CI}$ and TZ', [33s29p20d14f4g1h] for all other CI models), and quadruple-$\zeta$ (QZ, dyall.cv4z, [37s34p26d23f9g5h1i] for IHFSCC and CI models $\mathcal{I}^{CI}$ and $\mathcal{II}^{CI}$ and QZ', [37s34p24d19f7g4h] for all other CI models) basis sets were used. For the latter basis set, QZ', all 5d, 6d, 7s correlating functions, except for the *i* function, have been added. For the fluorine atom, aug-cc-pV$n$Z ($n = T, Q$) and cc-pV$n$Z ($n = D, T, Q$) [49] basis sets have been used.

For all wavefunction models of type $\mathcal{N}^{CC}$, $\mathcal{I}^{CI}$ and $\mathcal{II}^{CI}$ (see next paragraph for model definitions) we used the exact two-component Hamiltonian scheme of Iliaš and Saue [50] where two-electron spin-same-orbit (SSO) and spin-other-orbit (SOO) corrections were either obtained by means of atomic mean-field integrals [51, 52] (amf) or in a molecular mean-field approach [53] (mmf), based on the X2C transformation of the converged four-component Fock operator. For the models $\mathcal{III}^{CI}$ and $\mathcal{IV}^{CI}$ molecular spinors were optimized through all-electron four-component Dirac–Coulomb Hartree–Fock calculations. We based the open-shell calculations on an average-of-configuration Fock operator for two electrons in the three Kramers pairs of Th(7s, 6d$\delta$), the other 96 electrons are restricted to closed shells.

The IHFSCC calculations on $ThF^+$ were performed via the (0h, 2p) sector of Fock-space, by taking as closed-shell reference the $ThF^{3+}$ cation and following the route:

$$ThF^{3+}(0h, 0p) \rightarrow ThF^{2+}(0h, 1p) \rightarrow ThF^{+}(0h, 2p).$$

Table 1 summarizes the three different model ($P = P_m + P_i$) and correlating ($Q$) spaces which we explored in the IHFSCC singles and doubles calculations. The main model ($P_m$) space always comprised the virtual $ThF^{3+}$

**Table 1.** Active ($P_m$, $P_i$) and correlation spaces ($Q$) in terms of thorium spinors (the F 2s, 2p spinors are always included in the occupied $Q$ space). In case of MRCI the maximum excitation level for a given space is denoted as singles ($S$) and singles–doubles (SD), respectively.

| | | Th spinor distribution on spaces | | | | |
|---|---|---|---|---|---|---|
| | Model | 4f5s5p | 5d | 6s6p | 5f6d7s | 7p7d8s8p6f |
| IHFSCC | $\mathcal{I}^{CC}$ | frozen | frozen | $Q$ | $P_m$ | $P_i$ |
| | $\mathcal{II}^{CC}$ | frozen | $Q$ | $Q$ | $P_m$ | $P_i$ |
| | $\mathcal{III}^{CC}$ | $Q$ | $Q$ | $Q$ | $P_m$ | $P_i$ |
| MRCI | $\mathcal{I}^{CI}$ | frozen | $Q-S$ | $Q-S$ | $P_m$ | $Q-$SD |
| | $\mathcal{II}^{CI}$ | frozen | $Q-$SD | $Q-$SD | $P_m$ | $Q-$SD |

spinors with Th 5f, 6d, 7s character while the Th 7p spinors were included in the intermediate space $P_i$. Spinors with Th 7d, 8s, 8p and 6f character were added to the $P_i$ and all remaining virtuals were kept in the $Q$ space. A total number of 462 (818) virtual spinors were considered in the correlation treatment using the TZ (QZ) basis set combinations which corresponds in each case to an energy cutoff in the virtual space of ≈95 $E_H$. We have explored three different occupied $Q$ spaces, all of which comprised the spinors with F 2s, 2p character. The first $Q$ space contains in addition the spinors with Th 6s, 6p character, and corresponds to correlating 16 electrons (denoted in the following as $\mathcal{I}^{CC}$); the second ($\mathcal{II}^{CC}$) and third one ($\mathcal{III}^{CC}$) includes the spinors with Th 6s, 6p, 5d and Th 6s, 6p, 5d, 4f, 5s, 5p character, respectively, correlating in total 26 (48) electrons.

The first motivation for choosing these active spaces is the attempt to obtain the highest possible accuracy while avoiding intruder states in the calculations. Accuracy in IHFSCC calculations is linked to both the dimension of the model space $P$ (it has been argued [54] that large $P$ spaces may alleviate the need of considering triple or higher excitations in the dynamical correlation treatment due to the inclusion of corresponding excited determinants in the effective Hamiltonian) as well as of the intermediate space $P_m$ (states with their largest components in $P_m$ are described more accurately than those for which the largest components are in $P_i$ [41]). Secondly, considering a purely ionic model ($Th^{2+}F^-$), one may expect from the electronic structure of the $Th^{2+}$ ion in particular states to arise from the 6d5f manifold [55], whereas in the covalent ($Th^{1+}F$) case electronic states arising from the $6d^27s$ manifold of atomic $Th^+$ will play an important role.

The MRCI expansions for the spectroscopic studies, summarized in the lower part of table 1, are based on an average-of-configuration Fock operator for two electrons in 12 spinors (Th 7s6d) restricting all other electrons to closed shells. We considered two active space choices, $\mathcal{I}^{CI}$ and $\mathcal{II}^{CI}$, respectively, which differ by the maximum allowed excitation level of singles ($S$) and singles–doubles (SD) from the Th 5d6s6p + F 2s2p space into the model space $P_m$ comprising the Th 5f, 6d, 7s spinors. Due to limited resources, the virtual space was restricted to spinors below an energy of 20 $E_H$.

We defined models of varying quality to perform GASCI calculations of the effective electric field $E_{\text{eff}}$, the parallel magnetic hyperfine interaction constant $A_\parallel$, the scalar–pseudoscalar electron–nucleon interaction constant $W_{P,T}$ in the $\Omega = 1$ state and the vertical excitation energy $T_v$ of the $\Omega = 0$ state.

We used two principal CI models, denoted as $\mathcal{III}^{CI}$ and $\mathcal{IV}^{CI}$, the former of which has been further refined to accomodate for varying size of the valence spinor space and for the inclusion of determinants with more than two particles in the virtual spinor space. This elaborate choice of models is motivated by earlier findings on the ThO system [14]. The models are defined in full detail in table 2, using the nomenclature from table 1, for coherence.

For the calculation of the nuclear magnetic hyperfine coupling constant we use the thorium isotope $^{229}$Th for which the nuclear magnetic moment has been determined to be $\mu = 0.45\mu_N$ [56]. Its nuclear spin quantum number is $I = 5/2$. In all calculations the speed of light was set to 137.0359998 a.u.

### 3.2. Results and discussion

*3.2.1. Spectroscopic properties*

In order to settle the question about which state is the ground state we have carefully investigated convergence of the results with respect to inclusion of electron correlation effects and with respect to basis set extent. Theoretical excitation energies obtained from IHFSCC and a subset of GASCI calculations are compiled in table 3 along with theoretical and experimental results from Barker *et al* [29]. All data was calculated at $R = 1.981$ Å which corresponds to the calculated CCSDT($Q$) equilibrium geometry [29] of the $^1\Sigma_0^+$ state. Starting with the GASCI results a clear trend emerges which—nearly independent of the choice of active space composition and/or choice of Hamiltonian—places the $\Omega = 1$ state below the $\Omega = 0^+$ state by about 600–850 $cm^{-1}$. This finding seems to be in contradiction to the experimental data by Barker *et al* [29] whereas for the remaining $\Omega = 2, 3$ components, which primarily derive from the $^3\Delta$ state by first-order spin–orbit coupling (SOC), a good

**Table 2.** CI wavefunction models using refined active spinor spaces; the size of the $P_m$ active space is given, in the name of the models, by an upper index $X$, which is the number of Kramers pairs in the active space. The model $\mathit{III}^{CI,3}$ thus comprises the minimal active space to describe the $\Omega = 0$ et $\Omega = 1$ states corresponding to $^1\Sigma_0$ and $^3\Delta_1$ in the $A$–$S$ coupling picture. Models $\mathit{III}^{CI,3}$ and $\mathit{III}^{CI+T,3}$ differ by the highest excitation rank allowed from the hole spaces to the particle space. In the latter, triple excitations are included.

| Model | Th 6s,6p F 2s,2p | Th 7s,6d$\delta$ | Th 6d$\pi$ | Th 6d$\sigma$ | Th 7p$\pi$ | Th 7p$\sigma$,8s | Below 10 a.u. |
|---|---|---|---|---|---|---|---|
| $\mathit{III}^{CI,3}$ | $Q-SD$ | $P_m$ | $Q-SD$ | $Q-SD$ | $Q-SD$ | $Q-SD$ | $Q-SD$ |
| $\mathit{III}^{CI+T,3}$ | $Q-SD$ | $P_m$ | $Q-SDT$ | $Q-SDT$ | $Q-SDT$ | $Q-SDT$ | $Q-SDT$ |
| $\mathit{III}^{CI,5}$ | $Q-SD$ | $P_m$ | $P_m$ | $Q-SD$ | $Q-SD$ | $Q-SD$ | $Q-SD$ |
| $\mathit{III}^{CI,6}$ | $Q-SD$ | $P_m$ | $P_m$ | $P_m$ | $Q-SD$ | $Q-SD$ | $Q-SD$ |
| $\mathit{III}^{CI,8}$ | $Q-SD$ | $P_m$ | $P_m$ | $P_m$ | $P_m$ | $Q-SD$ | $Q-SD$ |
| $\mathit{III}^{CI,10}$ | $Q-SD$ | $P_m$ | $P_m$ | $P_m$ | $P_m$ | $P_m$ | $Q-SD$ |
| $\mathit{IV}^{CI}$ | frozen | $P_m$ | $P_m$ | $P_m$ | $P_m$ | $P_m$ | $Q-SD$ |

**Table 3.** Electronic spectra obtained with IHFSCC and MRCI for different model spaces at $R = 1.981$ Å. Unless otherwise noted, results are for TZ basis sets and include the spin-same orbit (SSO) and spin-other orbit interaction (SOO) in an atomic mean-field fashion. Subscripts on the electronic state labels indicate the value of $\Omega$. All energies are given in cm$^{-1}$. Our best estimate is displayed in boldface.

| | | | Electronic state energy | | | | |
|---|---|---|---|---|---|---|---|
| Method | Model[a,f] | Hamiltonian | $^1\Sigma^+_{0^+}$ | $^3\Delta_1$ | $^3\Delta_2$ | $^3\Delta_3$ | $^3\Pi_{0^-}$ |
| IHFSCC | $\mathit{I}^{cc}$ | 2c | 285.29 | 0.00 | 1063.29 | 3096.14 | 5228.76 |
| | $\mathit{II}^{cc}$ | 2c | 27.89 | 0.00 | 1070.40 | 3166.36 | 4690.68 |
| | $\mathit{II}^{cc}$[d] | 2c | 42.16 | 0.00 | 1062.01 | 3146.00 | 4499.13 |
| | $\mathit{III}^{cc}$[d] | 2c | 15.25 | 0.00 | 1062.22 | 3149.47 | 4510.50 |
| | $\mathit{III}^{cc}$[e] | 2c | 190.85 | 0.00 | 1048.27 | 3156.71 | 4123.14 |
| | $\mathit{III}^{cc}$[a] | 2c | 0.00 | 108.26 | 1157.05 | 3235.93 | 4415.96 |
| | $\mathbf{\mathit{III}^{cc}}$[f] | **2c** | **318.99** | **0.00** | **1038.94** | **3161.99** | **3841.17** |
| MRCI | $\mathit{I}^{CI}$ | 2c | 854.32 | 0.00 | 1154.40 | 3188.81 | 3387.74 |
| | $\mathit{II}^{CI}$ | 2c | 630.04 | 0.00 | 1166.86 | 2986.27 | - |
| CCSD(T)+SO[b] | | | 500.7 | 0.0 | 889.5 | 2156.8 | |
| CCSDT+SO[b] | | | 143.3 | 0.0 | 889.7 | 2157.1 | |
| CCSDT(Q)+SO[b] | | | 0.0 | 65.5 | 955.3 | 2222.9 | |
| MRCI+Q/SO[c] | | | 0.0 | 202 | 1047 | 2163 | |
| Experiment[b] | | | 0.00 | 315.0(5) | 1052.5(5) | 3150(15) | 3395(15) |

[a] 2c-mmf approach and modified QZ basis set for Th ((37s34p26d23f5g1h)).
[b] Reference [29].
[c] Reference [31].
[d] 2c-mmf approach.
[e] 2c-mmf approach (QZ basis set, (37s34p26d23f9g5h1i)).
[f] 2c-mmf approach and extrapolation to the basis set limit.

agreement is obtained. We also note that accounting in the GASCI expansion for correlation ($\mathit{SD}$) rather than mere polarization effects ($\mathit{S}$) from the 5d shell of Th decreases the gap between $\Omega = 0^+$ and $\Omega = 1$ states by approximately 30%.

Turning to the IHFSCC results, the smallest active space calculation ($\mathit{I}^{cc}$) yields the same picture as was obtained from the MRCI data: the $\Omega = 1$ state is predicted to be the ground state, although the energy gap to the $\Omega = 0^+$ state now reduces to only 285 cm$^{-1}$. In accord with the observed qualitative trend in the GASCI results, the inclusion of the Th 5d spinors in the correlation space (model $\mathit{II}^{cc}$) leads to a strong stabilization of the $\Omega = 0^+$ state by ≈260 cm$^{-1}$ with respect to the $\Omega = 1$ state. Interestingly, differential correlation effects of the Th 5d shell hardly affect the relative energy separation between the remaining spin–orbit split states $\Omega = 2, 3$ of the $^3\Delta$ term while the vertical excitation energy from the ground state to the $\Omega = 0^-$ component of the $^3\Pi$ state is lowered by ≈10%. Next, changing the treatment of the two-electron SSO and SOO corrections to a molecular mean-field approach enlarges the energetic gap between the $\Omega = 1$ and the $\Omega = 0^+$ states to 42 cm$^{-1}$, a small but perhaps non-negligible effect given the close proximity of the two states. Therefore, all further IHFSCC data have been obtained from 2c-calculations based on the mmf approach.

Using an extended $Q$ correlation space (model $\mathit{III}^{CC,\dagger}$) underlines the above-encountered stabilization of the $\Omega = 0^+$ state compared to the $\Omega = 1$ ground state with respect to the inclusion of core-valence correlations. The additional 4f5s5p correlation brings both states closer by about 30 cm$^{-1}$ which amounts to roughly 10% of

the effect of solely correlating the 5d shell. All remaining $\Omega$ states under consideration essentially retain their relative energetic separation. Going from a TZ to a QZ basis set description largely counteracts the correlation trends found above, as the $\Omega = 0^+$ state is now ≈190 cm$^{-1}$ above the $\Omega = 1$ state. To emphasize that this change is primarily a result of the quality of the correlating basis can be seen from table 3 by comparing the original QZ results ($\mathit{III}^{\mathrm{CC},\ddagger}$) to those for the modified QZ basis ($\mathit{III}^{\mathrm{CC},\S}$). In the latter case we employ the same QZ spdf function set whereas the higher angular momentum correlation functions resemble the extent of the TZ composition (5g1h). Using this combination places the $\Omega = 0^+$ state below the $\Omega = 1$ state by ≈100 cm$^{-1}$. It becomes apparent that basis set and electron correlation effects strongly influence the relative energetic separation between the $^1\Sigma$ and the $^3\Delta$ terms, but hardly affect the splitting of the $\Omega$ components of the $^3\Delta$ term. This is due to differential correlation effects arising from the presence of a Fermi hole in a triplet term and the absence of a Fermi hole in a singlet term, making the latter more sensitive to the description of interelectronic correlations than the former. These, in turn, depend on the quality of the one-particle basis set.

Due to the apparent sensitivity of the system's electronic excitation energies to the quality of the basis set, we performed an extrapolation to the complete basis set limit (model $\mathit{III}^{CC,\star}$ in table 3) based on the $\mathit{III}^{CC,\dagger}$ and $\mathit{III}^{CC,\ddagger}$ data, respectively, according to the expression [57]

$$E[(0\mathrm{h},\ 2\mathrm{p})]_{\Omega}^{\infty} = \frac{3^3 E[(0\mathrm{h},\ 2\mathrm{p})]_{\Omega}^{\mathrm{TZ}} - 4^3 E[(0\mathrm{h},\ 2\mathrm{p})]_{\Omega}^{\mathrm{QZ}}}{3^3 - 4^3}, \tag{5}$$

where $E[(0\mathrm{h},\ 2\mathrm{p})]_{\Omega}^{\mathrm{X}}$ (X = TZ, QZ) denotes the absolute energy in the sector (0h, 2p) of each of the states under consideration. The extrapolated relative term energies are roughly within 10 cm$^{-1}$ of the experimental ones for the $^1\Sigma_0^+$ and the components of the $^3\Delta$ state, if one assumes an inversion of the experimental assignment for the two lowest electronic states. For the $^3\Pi_0^-$ state the discrepancy with experiment is considerably larger (about 440 cm$^{-1}$).

The results of Barker *et al* [29]—based on spin-free open-shell coupled-cluster (CC) up to perturbative quadruples (CCSDT(Q)) combined *a posteriori* with SOC parameters obtained from a spin–orbit multireference configuration interaction (MRCI) calculation—show that cluster excitation ranks higher than Doubles play an important role in determining the relative energies of $^1\Sigma_0^+$ and $^3\Delta^1$, which is not surprising as this is a showcase for differential correlation effects (for the above-mentioned reasons). The IHFSCC models do not contain Triples and Quadruples excitations in the projection manifold, and so a downward correction on the $^1\Sigma_0^+$ energy from model $\mathit{III}^{\mathrm{CC},\star}$ is to be expected. On the other hand, all models of Barker *et al*, including the MRCI+Q/SO result from [31], yield a $^3\Delta 1$–$^3\Delta_2$ separation of roughly 900 cm$^{-1}$, whereas the experimental splitting is 740 cm$^{-1}$(or 1050 cm$^{-1}$ assuming an inverted assignment). Even poorer is the $^3\Delta_2$–$^3\Delta_3$ splitting for the model CCSDT(Q)+SO of 1270 cm$^{-1}$ compared with the experimental splitting of about 2100 cm$^{-1}$. Such a large error suggests that the perturbative treatment of spin–orbit interaction in a framework of $\Lambda$–$S$-coupled states is questionable for the respective state manifold of $ThF^+$. For instance, a larger splitting of the $^3\Delta$ components would shift the $^3\Delta_1$ state to an energy lower than that of $^1\Sigma_0^+$ in the perturbative CC calculations of Barker *et al*.

In contrast to this, the $^3\Delta_2$–$^3\Delta_3$ splitting from the rigorous non-perturbative CC model $\mathit{III}^{CC,\star}$ of 2120 cm$^{-1}$is in excellent agreement with the experimental value of 2100 cm$^{-1}$. The rigorous non-perturbative CI models yield similar results. Therefore, the remaining uncertainty in the calculations of Barker *et al* seem to be greater than those in the present calculations. Although their electron correlation treatment is of higher order than ours, it is obvious that our non-perturbative treatment of SOC is essential in obtaining the correct ground state.

In order to shed further light on the relative energies of the $^1\Sigma_0^+$ and $^3\Delta_1$ states and to verify whether these states may cross in the vicinity of their respective minima we calculated spectroscopic constants for both states from a quartic polynomial fit (program `twofit` provided by `DIRAC`) for a series of five equally-spaced (±0.025 Å and ±0.050 Å) data points around 1.981 Å using the computational model $\mathit{III}^{CC,\star}$. The resulting minimum internuclear distances ($R_e$), harmonic frequencies ($\omega_e$) and anharmonicity constants ($\omega_e x_e$) are compiled in table 4 together with the data of Barker *et al* [29].

Considering the equilibrium internuclear distance $R_e$ our IHFSCC as well as the other theoretical results are close enough to be consistent with the average value of $R_0 = 1.98(2)$ Å determined from the measurements of the rotational constants $B_0$ [29]. We further note that our calculated harmonic vibrational frequencies are similar to the CCSD(T)+SO results, in particular for the $\Omega = 0^+$ state, but overestimate both the experimental values and the theoretical benchmark CCSDT(Q)+SO of Barker *et al* by about 10–14 cm$^{-1}$. Although not very large, the discrepancy is not unexpected, since this property is in general sensitive to the extent to which dynamic electron correlations are accounted for.

Assuming a correct experimental assignment, we find rather large discrepancies for the anharmonicity constants, deviating by about 0.3 cm$^{-1}$ for the $^1\Sigma_0^+$ and −1 cm$^{-1}$ for the $^3\Delta_1$ state. Reversing the experimental assignment would improve the picture and reduce the discrepancies to −0.2 cm$^{-1}$ and −0.6 cm$^{-1}$ for the two

**Table 4.** Spectroscopic constants for the lowest two electronic states of $ThF^+$ in comparison with other theoretical and experimental work by Barker *et al* [29].

| State | | $\omega_e$/cm$^{-1}$ | $\omega_e x_e$/cm$^{-1}$ | $R_e$ (Å) |
|---|---|---|---|---|
| $^3\Delta_1$ | CCSDT(Q) + SO[a] | 651.1 | | 1.993 |
| | CCSD(T) + SO[a] | 654.1 | | 1.992 |
| | IHFSCC $\mathcal{III}^{cc,\star}$ | 667.3 | 1.268 | 1.984 |
| | Experiment[a] | 658.3(10) | 2.3(5) | |
| $^1\Sigma^+_{0^+}$ | CCSDT(Q) + SO[a] | 659.8 | | 1.981 |
| | CCSD(T) + SO[a] | 672.3 | | 1.975 |
| | CCSD(T)[b] | 675.7 | | 1.973 |
| | IHFSCC $\mathcal{III}^{cc,\star}$ | 670.8 | 2.088 | 1.974 |
| | Experiment[a] | 656.8(10) | 1.85(25) | |

[a] Reference [29].

[b] Reference [59]; calculations on the $^1\Sigma^+_{0^+}$ state are assumed.

**Table 5.** Vertical excitation energy for $\Omega = 0^+$, electron EDM effective electric field, magnetic hyperfine interaction constant, and scalar–pseudoscalar electron–nucleon interaction constant for $\Omega = 1$ at an internuclear distance of $R = 3.779\ a_0$ using basis sets with increasing cardinal number and the wavefunction model $\mathcal{III}^{CI,5}$.

| Basis set | $T_v$(cm$^{-1}$) | $E_{\rm eff}$ (GV cm$^{-1}$) | $A_\parallel$(MHz) | $W_{P,T}$ (kHz) |
|---|---|---|---|---|
| DZ | 378 | 37.8 | 1824 | 51.90 |
| TZ' | 787 | 36.9 | 1836 | 50.73 |
| QZ' | 877 | 36.9 | 1830 | 50.77 |

states, respectively. However, this average discrepancy is still about as large as the difference between the experimental anharmonicity constants for the two states in question. We, therefore, regard this finding as an indication but not as conclusive.

Based on the calculated spectroscopic constants, the potential minima of the discussed $\Omega$ states spread over a range of only 0.01 Å and are also very close to the reference value used for the calculations presented in table 3. This in turn means that the corresponding vertical excitation energies have to be close to the adiabatic ones. Based on the FSCC energies extrapolated to the basis-set limit, we estimate an adiabatic separation of 313.6 cm$^{-1}$ between the $\Omega = 1$ and $\Omega = 0^+$ states, or 317 cm$^{-1}$ for the 0–0 transition between the vibrational ground states.

*3.2.2. $\mathcal{P}$, $\mathcal{T}$-odd and magnetic hyperfine interaction constants*

We now turn to the discussion of our results of direct relevance to the search for $\mathcal{P}$, $\mathcal{T}$-odd effects in $ThF^+$. We have used a series of one-particle basis sets and CI models, all of which are defined in subsection 3.1 and table 2.

In order to minimize error bars we test the influence of several criteria, the first of which is the quality of the basis set. The results in table 5 demonstrate that the effective electric field and the parallel magnetic hyperfine interaction constant (for $^{229}$Th ($I = 5/2$)) are rather insensitive to the size of the basis set employed. Increasing the basis set cardinal number changes the value of the hyperfine interaction constant $A_\parallel$ by less than 0.6% in magnitude. Likewise, the correction yielded by the TZ' basis set for the effective electric field $E_{\rm eff}$ is smaller than 2% and the use of the QZ' basis set leads to a further change of less than 0.1%. The latter very small correction is also found for the electron–nucleon interaction constant. The vertical excitation energy for $\Omega = 0$ ($T_v$) undergoes a slightly larger change. Replacing the DZ by the TZ' basis set doubles the value of $T_v$, an increase of 409 cm$^{-1}$ on the absolute. Using the set of QZ' quality yields a correction of 11% in magnitude, less than 90 cm$^{-1}$. The sensitivity of this excitation energy to basis set extent was already observed in the results in table 3. However, based on the results in table 5 we conclude that the values of $E_{\rm eff}$, $A_\parallel$ and $W_{P,T}$ for $\Omega = 1$ are sufficiently converged with the TZ' basis set, allowing us to use this basis set for further analysis.

The results in table 6 show that wavefunctions accounting only for correlation effects among the two outermost valence electrons ($\mathcal{IV}^{CI}$) are too approximate for determining $E_{\rm eff}$, $A_\parallel$ and $W_{P,T}$ for $\Omega = 1$, although they do yield a correct qualitative description of the low-lying electronic valence states of the molecule and, in some cases, benefit from favorable error cancellations. It has been shown in [14] on the isoelectronic ThO molecule that these properties are essentially unaffected by accounting for electron correlations arising from Th core shells, and the reason for this has been explained via orbital (more precisely, spinor) perturbation theory.

**Table 6.** Vertical excitation energy for $\Omega = 0^+$, electron EDM effective electric field, magnetic hyperfine interaction constant, and scalar–pseudoscalar electron–nucleon interaction constant for $\Omega = 1$ at an internuclear distance of $R = 3.779\ a_0$ using the TZ' basis set, varying number of correlated electrons and varying active spinor spaces.

| CI model (TZ basis) | $T_v$(cm$^{-1}$) | $E_{\rm eff}$ (GV cm$^{-1}$) | $A_\parallel$(MHz) | $W_{P,T}$(kHz) |
|---|---|---|---|---|
| $IV^{CI}$ | 274 | 35.4 | 1749 | 49.44 |
| $III^{CI,3}$ | 1029 | 47.5 | 1842 | 65.78 |
| $III^{CI,5}$ | 787 | 36.9 | 1836 | 50.73 |
| $III^{CI,6}$ | 709 | 36.2 | 1836 | 49.90 |
| $III^{CI,8}$ | 598 | 35.6 | 1834 | 49.04 |
| $\mathbf{III^{CI,10}}$ | 538 | **35.2** | **1833** | **48.35** |

**Table 7.** Vertical excitation energy for $\Omega = 0^+$, electron EDM effective electric field, magnetic hyperfine interaction constant, and scalar–pseudoscalar electron–nucleon interaction constant for $\Omega = 1$ at an internuclear distance of $R = 3.779\ a_0$ using the DZ basis set and varying maximum excitation rank.

| CI model (DZ basis) | $T_v$ (cm$^{-1}$) | $E_{\rm eff}$ (GV cm$^{-1}$) | $A_\parallel$(MHz) | $W_{P,T}$ (kHz) |
|---|---|---|---|---|
| $III^{CI,3}$ | 654 | 47.0 | 1830 | 64.92 |
| $III^{CI,10}$ | 88 | 37.1 | 1832 | 51.06 |
| $III^{CI+T,3}$ | 247 | 35.4 | 1834 | 48.64 |

We have therefore carried out a study of the influence of the active spinor space, models of type $III^{CI,X}$, and restricting the electron correlation treatment to the outermost electronic shells (Th 6s, 6p, 7s, 6d, F 2s, 2p).

Our findings are very similar to those obtained for ThO in [14]. Increasing the size of the active space leads to significant corrections to the vertical excitation energy. The greatest change occurs when adding the energetically following $\pi$-type spinors to the minimal active space (step from $X = 3$ to $X = 5$). A similar drop of the values of the effective electric field and the hyperfine interaction constant is here observed. Including the energetically following spinors entails further decrease of all studied properties, but significantly less pronounced than the previous ones.

In view of the significant changes of the results when increasing the active spinor space, one could ponder the necessity to include excitation ranks higher than Doubles into the set of virtual spinors. We investigated this using the DZ basis (due to computational cost), and the results can be found in table 7. The hyperfine interaction constant $A_\parallel$ is insensitive to these higher excitations, allowing triple excitations to the virtual space changes the value by only 0.2%. However, the effective electric field as well as the S–PS interaction constant exhibit a strong dependence on higher excitations. The inclusion of triple excitations yields a drop of 25% in magnitude, respectively. Interestingly, this dramatic decrease is also observed when excluding triple excitations and augmenting the active spinor space by seven additional $\sigma$- and $\pi$-type Kramers pairs. Such an augmentation introduces a subset of triple and a subset of quadruple excitations but avoids terms with three or more particles in the external spinor space, therefore leading to a much shorter CI expansion. The additional excitation classes can be written symbolically as $(\text{core})^h$, $(\text{active})^p$, $(\text{external})^q$, where 'active' denotes the *additional* active-space spinors, $h$ denotes the number of holes and $p$ and $q$ the number of particles in the respective spinor space. In case of the triples, the additional sets of configurations then read as $(h = 1, p = 3, q = 0)$, $(h = 1, p = 2, q = 1)$ and $(h = 2, p = 1, q = 2)$. For the quadruples one obtains only $(h = 2, p = 2, q = 2)$. Evidently, the augmentation of the active space largely covers the set of Triple excitations that are required for obtaining accurate values of $E_{\rm eff}$, $A_\parallel$, and $W_{P,T}$. In case of the excitation energies we observe that the additional Quadruple excitations, which are not present in the model $III^{CI+T,3}$, have a significant effect of stabilizing the $\Sigma$ state relative to the $\Delta$ state, in accord with the discussion in the previous section.

In order to gain insight into the character of the excitations leading to important corrections, we carried out a detailed analysis of the wavefunction expansions referred to as $III^{CI,3}$ and $III^{CI,5}$. They turn out to be very similar, the expansion coefficients remaining almost unchanged with the exception of a determinant that is the next-to-leading contributor with a coefficient $c \approx 0.046$ in the expansion of $III^{CI,5}$ whereas its coefficient is much smaller in the $III^{CI,3}$ expansion ($c < 0.01$). This respective determinant can be written as a $\delta^1_{6d}(\pi\sigma)^1$ occupation which corresponds to a single excitation with respect to the leading determinant $\sigma^1_{7s}\delta^1_{6d}$ for this $\Omega = 1$ state. Since $\delta^1_{6d}(\pi\sigma)^1$ is already contained in the $III^{CI,3}$ expansion, it is necessarily the additional higher excitations included in the $III^{CI,5}$ expansion which lend amplitude to the $\delta^1_{6d}(\pi\sigma)^1$ determinant.

We carried out a Mulliken population analysis of the spinors occupied in this decisive determinant. $\pi\sigma$ denotes a spinor of $\pi$-character with significant admixture of $\sigma$-character (see the fourth spinor in table 8). In

**Table 8.** Characterization of active-space Kramers pairs in the TZ' basis. Orbital angular momentum projection, spinor energy and principal atomic shell character.

| Spinor($\lambda$) | $\langle \hat{l}_z \rangle_{\varphi_i}$ | $\varepsilon_{\varphi_i}[E_H]$ | atomic character |
|---|---|---|---|
| $\sigma$ | −0.001 | −0.43 | 85% Th(s) |
| $\delta$ | 1.966 | −0.42 | 98% Th(d) |
| $\delta$ | −2.000 | −0.41 | 99% Th(d) |
| $\pi\sigma$ | −0.720 | −0.14 | 50% Th(p), 45% Th(d) |
| $\pi$ | 1.025 | −0.13 | 60% Th(d), 36% Th(p) |
| $\sigma\pi$ | −0.290 | −0.12 | 47% Th(d), 43% Th(p) |
| $\pi$ | −0.980 | −0.10 | 55% Th(p), 36% Th(d) |
| $\pi$ | 1.011 | −0.09 | 64% Th(p), 29% Th(d) |
| $\sigma$ | −0.005 | −0.07 | 59% Th(p), 19% Th(f), 15% Th(d) |
| $\sigma$ | −0.014 | −0.06 | 90% Th(s) |
| $\pi\sigma$ | −0.894 | −0.03 | 89% Th(p) |
| $\pi$ | 1.006 | −0.03 | 94% Th(p) |
| $\sigma$ | −0.097 | −0.02 | 65% Th(p), 29% Th(f) |

this case, the spinor is of Th(p) character with a 45% contribution of Th(d). Hence, the drop of the values of $E_{\text{eff}}$, $A_{\parallel}$ and $W_{P,T}$ is related to a shift of electron density from Th(7s) to Th(p) and Th(d), the two latter of which have an approximate angular node at the nucleus. It is, therefore, physically plausible that the mentioned higher excitations which lead to a shift of electron (spin) density from Th(s) to higher angular momentum entail a reduction of the EDM effective electric field and of the magnetic hyperfine interaction constant.

*3.2.3. EDMs and transition dipole moments*

The electron EDM effective field stands in direct relationship with the static molecular electric dipole moment. We calculated this latter quantity as an expectation value over relativistic CI wavefunctions, and in addition, electric dipole transition moments between different electronic states in an energy window of up to roughly 8000 $cm^{-1}$. The results are compiled in tables 9 and 10. Concerning the notation for electronic states we have here added information on dominant and minor contributors in $\Lambda$–$S$ coupling to a given well-defined $\Omega$ state.

The absolute molecular dipole moment is very large for $ThF^+$, especially for the low-lying electronic states, and reaches into the range of the largest dipole moments for diatomic molecules. Concerning transition dipole moments we observe a generally good agreement with expected selection rules for transitions between different $^{2S+1}\Lambda_{\Omega}$ states. For example, the largest matrix element, $||\langle {}^3\Phi_2 ({}^3\Pi_2)|\hat{\vec{D}}|{}^3\Delta_1\rangle|| = 1.34$ [D], is spin-allowed ($\Delta S = 0$) and also orbital angular momentum allowed (here $\Delta\Lambda = \pm 1$). In addition, $\Delta\Omega = \pm 1$ is also satisfied. On the other hand, very small transition moments are typically found for spin-forbidden transitions. Our study of transition moments covers a few more states than those reported in a recent study on actinide bonding by Heaven *et al* [31], and the results agree quite well with the values obtained in that reference.

The comparison of tables 9 and 10 shows that our larger set of results obtained with the smaller two-electron CI expansion agrees quite well with the results from the more elaborate CI model, $\mathcal{II}^{CI}$. We therefore consider the values in table 9 as a good approximation to the accurate values.

## 4. Conclusion

In the earlier work of Barker *et al* [29, 31] $\Omega = 0^+$ had been proposed as the electronic ground state of $ThF^+$, supported by the measured intensities of the lowest band compared to those of other bands in a pulsed field ionization — zero kinetic energy experiment. Accompanying many-body electronic structure calculations were judged to be inconclusive in this regard. From our detailed discussion of relativistic many-body calculations, including those from [29] for excitation energies, we conclude that the assignment of the ground electronic state of $ThF^+$ remains an open issue. The models of Barker *et al* suffer from the incomplete account of spin–orbit interaction and its intertwining with dynamic electron correlations, which becomes manifest in the poor description of the energetic splitting of the $^3\Delta$ state into its $\Omega$ components. Our present study takes these effects into account rigorously which leads to a $^3\Delta_1$ ground state. Our models lack excitation ranks with three or more external particles in the wave operator, the inclusion of which may have the effect of inverting the energetic order of $^3\Delta_1$ and $^1\Sigma_0^+$ due to considerable differential correlation effects. On the other hand, the model spaces we have used in the IHFSCC calculations do give rise to a subset of excitation ranks higher than doubles in the projection manifold, which should mitigate the uncertainty in our best present calculation (boldface in table 3). Giving preference to assigning the ground state as $^1\Sigma_0^+$ is, therefore, no longer tenable from a theoretical point of view,

**Table 9.** Molecular static electric dipole moments $\langle {}^{M}\Lambda_{\Omega} | \hat{D}_{\perp} | {}^{M}\Lambda_{\Omega} \rangle$, transition dipole moments $||\langle {}^{M}\Lambda'_{\Omega} | \hat{\vec{D}} | {}^{M}\Lambda_{\Omega} \rangle||$, with $\hat{\vec{D}}$ the electric dipole moment operator (both in [D] units), and vertical transition energies for low-lying electronic states using the TZ' basis set and the CI model $TV^{CI}$. The origin is at the center of mass, and the internuclear distance is $R = 3.779\ a_0$. $({}^{M}\Lambda_{\Omega})$ denotes a term contributing at least 10% to the state. $^{1,3}$ denotes cases where $\Lambda$–$S$ coupling breaks down significantly according to the analysis of our spinor-based $\omega$–$\omega$ coupled wavefunctions.

| ${}^{M}\Lambda_{\Omega}$ state | $T_v$ (cm$^{-1}$) | $^1\Sigma_0^+$ | $^3\Delta_1$ | $^3\Delta_2$ | $^3\Delta_3$ | $^1\Sigma_0\,(^3\Pi_0)$ | $^3\Pi_0$ | $^{1,3}\Pi_1\,(^3\Sigma_1)$ | $^3\Pi_0\,(^1\Sigma_0)$ | $^{1,3}\Delta_2\,(^3\Pi_2)$ | $^3\Sigma_1$ | $^{1,3}\Pi_1$ | $^3\Phi_2\,(^3\Pi_2)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $^1\Sigma_0^+$ | 274 | −4.004 | | | | | | | | | | | |
| $^3\Delta_1$ | 0 | 0.012 | −4.075 | | | | | | | | | | |
| $^3\Delta_2$ | 724 | 0.000 | 0.070 | −4.022 | | | | | | | | | |
| $^3\Delta_3$ | 2198 | 0.000 | 0.000 | 0.052 | −4.075 | | | | | | | | |
| $^1\Sigma_0\,(^3\Pi_0)$ | 6344 | 0.439 | 0.455 | 0.000 | 0.000 | −3.752 | | | | | | | |
| $^3\Pi_0$ | 6528 | 0.000 | 0.571 | 0.000 | 0.000 | 0.000 | −2.116 | | | | | | |
| $^{1,3}\Pi_1\,(^3\Sigma_1)$ | 6639 | 0.868 | 0.142 | 0.218 | 0.000 | 0.197 | 0.000 | −2.375 | | | | | |
| $^3\Pi_0\,(^1\Sigma_0)$ | 6747 | 0.003 | 0.391 | 0.000 | 0.000 | 0.929 | 0.000 | 0.094 | −2.717 | | | | |
| $^{1,3}\Delta_2\,(^3\Pi_2)$ | 7008 | 0.000 | 0.473 | 0.334 | 0.298 | 0.000 | 0.000 | 0.529 | 0.000 | −2.734 | | | |
| $^3\Sigma_1$ | 7490 | 0.226 | 0.069 | 0.221 | 0.000 | 0.136 | 0.197 | 0.451 | 0.145 | 0.087 | −4.463 | | |
| $^{1,3}\Pi_1$ | 7918 | 0.667 | 0.052 | 0.801 | 0.000 | 0.011 | 0.064 | 0.107 | 0.043 | 0.444 | 0.209 | −2.708 | |
| $^3\Phi_2\,(^3\Pi_2)$ | 8245 | 0.000 | 1.338 | 0.234 | 0.272 | 0.000 | 0.000 | 0.134 | 0.000 | 0.384 | 0.018 | 0.099 | −2.271 |

**Table 10.** Molecular static electric dipole moments $\left\langle {}^{M}\Lambda_{\Omega} \left| \hat{D}_z \right| {}^{M}\Lambda_{\Omega} \right\rangle$, with $\hat{\vec{D}}$ the electric dipole moment operator, using the TZ basis set and the CI model $\mathit{III}^{CI}$. The origin is at the center of mass, and the internuclear distance is $R = 3.779\ a_0$ (F nucleus at $z\vec{e}_z$ with $z < 0$). ${}^{M}\Lambda$ is an approximate notation and refers to the term derived from the leading Slater determinants and the information in table 8.

| ${}^{M}\Lambda_{\Omega}$ State | $T_v$ (cm$^{-1}$) | $\left\langle {}^{M}\Lambda_{\Omega} \left\| \hat{D}_z \right\| {}^{M}\Lambda_{\Omega} \right\rangle$ ($D$) |
|---|---|---|
| $^1\Sigma_0^+$ | 630 | 3.941 |
| $^3\Delta_1$ | 0 | 4.029 |
| $^3\Delta_2$ | 1167 | 3.970 |
| $^3\Delta_3$ | 2986 | 4.034 |

**Table 11.** Effective electric field for the science states of selected diatomic candidate molecules in search of parity- and time-reversal violation.

| Molecule | Electronic state | $E_{\rm eff}$ (GV cm$^{-1}$) |
|---|---|---|
| ThO | $^3\Delta_1$ | 75.2[a], 81.5[b] |
| YbF | $^2\Sigma_{1/2}^+$ | 26[c], 25[d], 24[e] |
| PbO | $^3\Sigma_1^+$ | 25[f] |
| $ThF^+$ | $^3\Delta_1$ | 35.2[g], 90[h] |
| WC | $^3\Delta_1$ | −36[i] |

[a] Reference [14].
[b] Reference [60].
[c] Reference [61].
[d] Reference [62].
[e] Reference [63].
[f] Reference [64].
[g] This work.
[h] Reference [24].
[i] Reference [20].

based on our present findings. In any case, it is beyond reasonable doubt that the two respective states are the lowest-lying electronic states and that they are so close in energy that an eEDM experiment could be carried out irrespective of their ordering [32].

We conclude that our best model for the determination of $\mathcal{P}, \mathcal{T}$-odd and magnetic hyperfine interaction constants is $\mathit{III}^{CI,10}$ in the TZ' basis set (boldface in table 6), which displays property values nearly converged with respect to the different degrees of freedom in the models we have tested. Our best prediction for the hyperfine constant in the $\Omega = 1$ 'science state' is 1833 MHz, which awaits confirmation from an experimental measurement. The obtained effective electric field of $E_{\rm eff} = 35.2$ GV cm$^{-1}$ in this same state is more than 60% smaller than the value of $E_{\rm eff} = 90$ GV cm$^{-1}$ obtained earlier by Meyer *et al* [24]. The large deviation is very likely to be due to the limited set of electronic configurations and further model-inherent approximations used in the approach of Meyer *et al*. The smaller value of $E_{\rm eff}$ is a setback for potential electron EDM searches with this molecular ion, but given the body of other favorable properties (low-lying $^3\Delta_1$ state, large molecular dipole moment) of $ThF^+$ still large enough to retain the system as a promising candidate in search of $\mathcal{P}, \mathcal{T}$ violation. In table 11 we provide a summary of $E_{\rm eff}$ values in the respective science states of some diatomic molecules of current interest in this search. Our $E_{\rm eff}$ presently determined for $ThF^+$ is still larger than $E_{\rm eff}$ in the science state of the YbF molecule, in which a new upper bound to the electron EDM had been determined in 2011 [58]. The static electric transition dipole moments we have determined for a set of states below 9000 cm$^{-1}$ in $ThF^+$ may also be helpful in devising a route for state preparation for an EDM measurement in this promising molecular ion.

## Acknowledgments

MD, MKN, and TF thank the *Agence Nationale de la Recherche* (ANR) through grant no. ANR-BS04-13-0010-01, project 'EDMeDM', for financial support. This work was granted access to the HPC resources of CALMIP under the allocation 2014-p1424. ASPG acknowledges support from PhLAM (Unité Mixte de Recherche de l'Université de Lille 1 et du CNRS), the French National Research Agency (Contract ANR-11-LABX-0005 'Chemical and Physical Properties of the Atmosphere' (CaPPA)), and the European Commission through the FP7 TALISMAN project (Contract No. 323300). We are grateful for HPC resources from GENCI-CCRT (Grant 2013-081859) and the Danish Center for Scientific Computing (DCSC) at SDU Odense. We thank Eric Cornell (JILA, Boulder) for helpful comments and Radovan Bast (Stockholm) for technical support.

## References

[1] Dine M and Kusenko A 2004 Leptonic CP violation *Rev. Mod. Phys.* **76** 1
[2] Sakharov A D 1967 Violation of CP invariance, *C* asymmetry, and baryon asymmetry of the Universe *JETP Lett.* **5** 24
[3] Pospelov M and Ritz A 2005 Electric dipole moments as probes of new physics *Ann. Phys.* **318** 119
[4] Fortson N, Sandars P and Barr S 2003 *Phys. Today* **56** 33–39
[5] Streater R F and Wightman A S 1964 *PCT, Spin and Statistics, and All That* (New York: W A Benjamin)
[6] Sachs R G 1987 *The Physics of Time Reversal* (Chicago, IL: University of Chicago Press)
[7] Roberts B L and Marciano W J 2009 The electric dipole moment of the electron *Lepton Dipole Moments* (*Advances Series on Directions in High Energy Physics* vol 20) ed E D Commins and D DeMille (New Jersey: World Scientific) chapter 14, pp 519–81
[8] Liu Z W and Kelly H P 1992 Analysis of atomic electric dipole moment in thallium by all-order calculations in many-body perturbation theory *Phys. Rev.* A **45** R4210
[9] Bernreuther W and Suzuki M 1991 The electric dipole moment of the electron *Rev. Mod. Phys.* **63** 313
[10] Sushkov O P and Flambaum V V 1978 Parity breaking effects in diatomic molecules *Sov. Phys.—JETP* **48** 608
[11] Sandars P G H 1965 The electric dipole moment of an atom *Phys. Lett.* **14** 194
[12] Regan B C, Commins E D, Schmidt C J and DeMille D 2002 New limit on the electron electric dipole moment *Phys. Rev. Lett.* **88** 071805
[13] Baron J *et al* The ACME Collaboration 2014 Order of magnitude smaller limit on the electric dipole moment of the electron *Science* **343** 269
[14] Fleig T and Nayak M K 2014 Electron electric dipole moment and hyperfine interaction constants for ThO *J. Mol. Spectrosc.* **300** 16
[15] Skripnikov L V, Petrov A N and Titov A V 2013 Communication: theoretical study of ThO for the electron electric dipole moment search *J. Chem. Phys.* **139** 221103
[16] DeMille D and On behalf of the ACME collaboration 2014 ASME: search for the electron's electric dipole moment *24th Int. Conf. on Atomic Physics (ICAP 2014), (Washington, DC, 3-8 August 2014)* (http://icap2014.org/book-abstracts-download)
[17] Leanhardt A E, Bohn J L, Loh H, Maletinsky P, Meyer E R, Sinclair L C, Stutz R P and Cornell E A 2011 High-resolution spectroscopy on trapped molecular ions in rotating electric fields: a new approach for measuring the electron electric dipole moment *J. Mol. Spectrosc.* **270** 1
[18] Cossel K C *et al* 2012 Broadband velocity modulation spectroscopy of $HfF^+$: towards a measurement of the electron electric dipole moment *Chem. Phys. Lett.* **546** 1
[19] Meyer E R, Bohn J L and Deskevich M P 2006 Candidate molecular ions for an electron electric dipole moment experiment *Phys. Rev.* A **73** 062108
[20] Lee J, Chen J, Skripnikov L V, Petrov A N, Titov A V, Mosyagin N S and Leanhardt A E 2013 Optical spectroscopy of tungsten carbide for uncertainty analysis in electron electric dipole moment search *Phys. Rev.* A **87** 2013
[21] Wang F and Steimle T C 2011 Communication: electric dipole moment and hyperfine interaction of tungsten monocarbide, WC *J. Chem. Phys.* **134** 201106
[22] Petrov A N, Mosyagin N S, Isaev T A and Titov A V 2007 Theoretical study of $HfF^+$ in search for the electron electric dipole moment *Phys. Rev.* A **76** 030501
[23] Fleig T and Nayak M K 2013 Electron electric-dipole-moment interaction constant for $HfF^+$ from relativistic correlated all-electron theory *Phys. Rev.* A **88** 032514
[24] Meyer E R and Bohn J L 2008 Prospects for an electron electric-dipole moment search in metastable ThO and $ThF^+$ *Phys. Rev.* A **78** 010502
[25] Adam A G, Hopkins W S and Tokaryk D W 2004 High-resolution visible laser spectroscopy of HfF *J. Mol. Spectrosc.* **225** 1
[26] Barker B J, Antonov I O, Bondybey V E and Heaven M C 2011 Communication: spectroscopic measurements for $HfF^+$ of relevance to the investigation of fundamental constants *J. Chem. Phys.* **134** 201102
[27] Petrov A N, Mosyagin N S and Titov A V 2009 Theoretical study of low-lying electronic terms and transition moments for $HfF^+$ for the electron electric dipole moment search *Phys. Rev.* A **79** 012505
[28] Skripnikov L V, Mosyagin N S, Petrov A N and Titov A V 2009 *JETP Lett.* **88** 578
[29] Barker B J, Antonov I O, Heaven M C and Peterson K A 2012 Spectroscopic investigations of ThF and $ThF^+$ *J. Chem. Phys.* **136** 104305
[30] Irikura K K 2013 *J. Phys. Chem.* A **117** 1276
[31] Heaven M C, Barker B J and Antonov I O 2014 Spectroscopy and structure of the simplest actinide bonds *J. Phys. Chem.* A **118** 10867
[32] Cornell E A and Gresh D N 2014 private communication
[33] Lindroth E, Lynn B W and Sandars P G H 1989 Order $\alpha^2$ theory of the atomic electric dipole moment due to an electric dipole moment on the electron *J. Phys.* B **22** 559
[34] Kozlov M G 1985 Semiempirical calculations of *P*-and *P, T*-odd effects in diatomic molecules-radicals *Zh. Eksp. Teor. Fiz.* **89** 1933
[35] Saue T and Visscher L 2003 Four-component electronic structure methods for molecules *Theoretical Chemistry and Physics of Heavy and Super Heavy Elements* ed U Kaldor and S Wilson (Dordrecht: Kluwer)
[36] Čársky P, Paldus J and Pittner J (ed) 2010 *Recent Progress in Coupled Cluster Methods* (Heidelberg, Germany: Springer)
[37] Visscher L, Eliav E and Kaldor U 2001 Formulation and implementation of the relativistic Fock-space coupled-cluster method for molecules *J. Chem. Phys.* **115** 9720

[38] Landau A, Eliav E, Ishikawa Y and Kaldor U 2001 Intermediate Hamiltonian Fock-space coupled cluster method in the one-hole one-particle sector: excitation energies of xenon and radon *J. Chem. Phys.* **115** 6862–5
[39] Landau A, Eliav E, Ishikawa Y and Kaldor U 2000 Intermediate Hamiltonian Fock-space coupled-cluster method: excitation energies of barium and radium *J. Chem. Phys.* **113** 9905–10
[40] Landau A, Eliav E and Kaldor U 1999 Intermediate Hamiltonian Fock-space coupled-cluster method *Chem. Phys. Lett.* **313** 399
[41] Landau A, Eliav E, Ishikawa Y and Kaldor U 2004 Mixed-sector intermediate Hamiltonian Fock-space coupled-cluster approach *J. Chem. Phys.* **121** 6634
[42] Fleig T, Olsen J and Marian C M 2001 The generalized active space concept for the relativistic treatment of electron correlation: I. Kramers-restricted two-component configuration interaction *J. Chem. Phys.* **114** 4775
[43] Fleig T, Olsen J and Visscher L 2003 The generalized active space concept for the relativistic treatment of electron correlation: II. Large-scale configuration interaction implementation based on relativistic 2- and 4-spinors and its application *J. Chem. Phys.* **119** 2963
[44] Knecht S, Aa Jensen H J and Fleig T 2010 Large-scale parallel configuration interaction: II. Two- and four-component double-group general active space implementation with application to BiH *J. Chem. Phys.* **132** 014108
[45] Jensen H J A *et al* 2012 DIRAC, a relativistic ab initio electronic structure program, Release DIRAC12 www.diracprogram.org
[46] Bast R *et al* 2011 DIRAC, a relativistic *ab initio* electronic structure program, Release DIRAC11 http://dirac.chem.vu.nl
[47] Dyall K G 2006 Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the actinides ac-lr *Theor. Chim. Acta* **90** 491
[48] Dyall K G 2012 Core correlating basis functions for elements 31-118 *Theor. Chim. Acta* **131** 1217
[49] Dunning T H Jr 1989 Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen *J. Chem. Phys.* **90** 1007
[50] Iliaš M and Saue T 2007 An infinite-order two-component relativistic hamiltonian by a simple one-step transformation *J. Chem. Phys.* **126** 064102
[51] Hess B A, Marian C M, Wahlgren U and Gropen O 1996 A mean-field spin–orbit method applicable to correlated wavefunctions *Chem. Phys. Lett.* **251** 365–71
[52] Schimmelpfennig B, Maron L, Wahlgren U, Teichteil C, Fagerli H and Gropen O 1998 On the combination of ECP-based CI calculations with all-electron spin–orbit mean-field integrals *Chem. Phys. Lett.* **286** 267
[53] Sikkema J, Visscher L, Saue T and Iliaš M 2009 The molecular mean-field approach for correlated relativistic calculations *J. Chem. Phys.* **131** 124116
[54] Eliav E, Borschevsky A, Shamasundar K R, Pal S and Kaldor U 2009 *Int. J. Quantum Chem.* **109** 2909
[55] Kramida A, Ralchenko Y, Reader J and NIST ASD Team 2013 NIST Atomic Spectra Database (version 5.1) (Gaithersburg, MD.: National Institute of Standards and Technology) (http://physics.nist.gov/asd)
[56] Kazakov G A, Litvinov A N, Romanenko V I, Yatsenko L P, Romanenko A V, Schreitl M, Winkler G and Schumm T 2012 Performance of a $^{229}$Thorium solid-state nuclear clock *New J. Phys.* **14** 083019
[57] Helgaker T, Jørgensen P and Olsen J 2000 *Molecular Electronic Structure Theory* (Chichester: Wiley)
[58] Hudson J J, Kara D M, Smallman I J, Sauer B E, Tarbutt M R and Hinds E A 2011 Improved measurement of the shape of the electron *Nature* **473** 493
[59] Thanthiriwatte K S, Wang X, Andrews L, Dixon D A, Metzger J, Vent-Schmidt T and Riedel S H 2014 *J. Phys. Chem.* A **118** 2107
[60] Skripnikov L V and Titov A V 2015 Theoretical study of thorium monoxide for the electron electric dipole moment search: electronic properties of $H^3\Delta_1$ in ThO *J. Chem. Phys.* **142** 024301
[61] Kozlov M G 1997 Enhancement of the electric dipole moment of the electron in the YbF molecule *J. Phys.* B **30** L607
[62] Mosyagin N S, Kozlov M G and Titov A V 1998 *J. Phys.* B **31** L763
[63] Parpia F 1998 *Ab initio* calculation of the enhancement of the electric dipole moment of an electron in the YbF molecule *J. Phys.* B **31** 1409
[64] Kozlov M G and DeMille D 2002 Enhancement of the electric dipole moment of the electron in PbO *Phys. Rev. Lett.* **89** 133001

# Appendix D

# Development of FDE and Applications to Heavy Elements

This chapter presents the papers concerning the development of FDE, which are by nature applicable to both light and heavy elements, and their application to systems containing heavy elements. Here, the use of the "static" embedding potential is assessed for uranyl and neptunyl chloride is discussed in D.1 and D.6, whereas the shortcomings of approximate kinetic energy density functionals for complexes of CUO and noble gases is presented in D.4.

The work on the PyADF scripting framework, which was developed in order to facilitate the use of one or more electronic structure codes in computational chemistry workflows, is presented in D.2. While of a more technical nature, work of this kind is of immense value to the case of embedding calculations such as those carried out in the aforementioned works, as it serves to automate the execution and data exchange, e.g. the construction of embedding potentials with ADF and its use with Dirac.

Also, though not specifically discussing heavy elements, I present in D.3 the general formalism through which it is possible to define WFT-in-DFT for both variational and non-variational wavefunctions for the ground state, and for excited states and general molecular properities via time-dependent response theory. This formalism has been applied to define the CC-in-DFT approach presented in D.5.

This response-based formalism has also served as a basis for the work in D.7 which outlines the definition of magnetic properties for FDE in a four-component framework at the DFT-in-DFT level. The ease with which we can change electronic structure approach and Hamiltonia serves to underscore the power of the framework defined in D.3, and the fact that expanding it while focusing on relativistic approaches automatically opens us perspectives for exploring non-relativistic approaches and vice-versa. The focus on magnetic properties in a four-component framework of this last paper also represents a new research direction in for me, which as discussed above I plan to consolidate in the coming years.

## D.1 Paper VIII

PAPER www.rsc.org/pccp | Physical Chemistry Chemical Physics

# Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory

André Severo Pereira Gomes,*[a] Christoph R. Jacob*[ab] and Lucas Visscher*[a]

*Received 4th April 2008, Accepted 21st May 2008*
*First published as an Advance Article on the web 4th July 2008*
**DOI: 10.1039/b805739g**



We present a simple and efficient embedding scheme for the wave-function based calculation of the energies of local excitations in large systems. By introducing an embedding potential obtained from density-functional theory (DFT) it is possible to describe the effect of an environment on local excitations of an embedded system in wave-function theory (WFT) calculations of the excitation energies. We outline the implementation of such a WFT-in-DFT embedding procedure employing the ADF, Dalton and DIRAC codes, where the embedded subsystem is treated with coupled cluster methods. We then evaluate this procedure in the calculation of the solvatochromic shift of acetone in water and of the *f*–*f* spectrum of $NpO_2^{2+}$ embedded in a $Cs_2UO_2Cl_4$ crystal and find that our scheme does effectively incorporate the environment effect in both cases. A particularly interesting finding is that with our embedding scheme we can model the equatorial $Cl^-$ ligands in $NpO_2Cl_4^{2-}$ quite accurately, compared to a fully wavefunction-based calculation, and this opens up the possibility of modeling the interaction of different ligands to actinyl species with relatively high accuracy but at a much reduced computational cost.

## I. Introduction

Electronic excitations play an important role in several biological processes, such as photosynthesis and vision, as well as in technological applications like lighting materials. The information that computational chemistry is able to provide on such phenomena is helpful in the interpretation of complex experimental data (for reviews see, *e.g.*, ref. 1 and 2) and can be used in the development of new materials (see, *e.g.*, ref. 3).

To be of practical use, calculations should yield an accurate, preferably quantitatively correct, picture, but also be computationally efficient so that real-life systems can be tackled. Given its good balance between accuracy and computational efficiency, time-dependent density-functional theory (TDDFT)[1,4,5] has become the standard *ab initio* approach for treating excited states of large-scale systems. Even though there are efforts towards having efficient wave-function based methods to calculate excitation energies,[6–8] these remain computationally very expensive compared to TDDFT.

Full TDDFT calculations are, however, also limited to systems with up to a few hundred atoms, bringing calculations for still larger systems out of reach for routine application. This leaves the domain of large (biological) systems to subsystem methods that assume localization of the electronic excitations. The most employed approaches are hybrid quantum mechanics/molecular mechanics (QM/MM) methods,[9–11] that treat a central part in which the excitations of interest take place using a quantum mechanical method such as TDDFT, while its environment is described using molecular mechanics (for examples see, *e.g.*, ref. 12–16). However, in order to obtain accurate results, the force field used in the MM part has to be parametrized carefully, which is particularly difficult for heavy elements that display a variety of bonding interactions due to the many chemically accessible valence orbitals.

This weakness of the MM description can be overcome by considering a subsystem method with a QM description of the environment, the so-called QM/QM embedding schemes.[17–20] Among these schemes, the ONIOM methods by Morokuma and co-workers[17,18] are very popular. However, for the calculation of molecular properties these do not include the effect of the environment on the electronic structure of the embedded system, and are, therefore, only applicable as long as the property of interest can be adequately described by the lower-level method. Another example of such QM/QM schemes, that is particularly suited for studying localized excitations in solids is the *ab initio* model potential (AIMP) method.[21–23] In this method, the effect of atoms or ions (or in a recent extension also larger fragments[24]) in the environment of a subsystem of interest is included in the calculation of this active subsystem by means of nonlocal model potential obtained from Hartree–Fock theory. These model potentials contain, in addition to the electrostatic potential of the environment, projection operators to ensure the orthogonality between the wave functions of the active subsystem and the fragments in the environment.

A different QM/QM approach is taken in the frozen-density embedding (FDE) scheme by Wesolowski and Warshel,[25] in which both the system of interest and its environment are described using DFT (DFT-in-DFT). By basing the

[a] *Department of Theoretical Chemistry, Faculty of Sciences, Amsterdam Center for Multiscale Modeling, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands. E-mail: gomes@few.vu.nl; visscher@chem.vu.nl*
[b] *Present address: ETH Zurich, Laboratorium für Physikalische Chemie, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland. E-mail: christoph.jacob@phys.chem.ethz.ch*

formulation on DFT, it is possible to define a local embedding potential that only depends on the electron densities in the active subsystem and in the environment and does not contain any nonlocal projection operators. The FDE scheme has been shown to be both accurate and efficient for the calculation of solvent effects on electronic excitation energies[26–28] as well as for the description of induced circular dichroism in guest–host complexes[29] and the electronic spectra of transition metal- and lanthanide-containing solids.[30,31] Recently, it has been generalized to include the description of couplings between excitations in different subsystems,[32] and it has been shown that this can be used to calculate excitation energy transfer couplings in natural light-harvesting systems.[33]

Notwithstanding the success of the applications mentioned, the use of TDDFT in a pure DFT-in-DFT embedding scheme will encounter the limitations of TDDFT itself. Most important is the well-known problem in describing charge-transfer (CT) excitations,[34–38] that are of particular importance in many biological systems and other systems exhibiting interesting photophysical properties.[39–43] The DFT-in-DFT frozen density embedding scheme is able to remove spurious solvent–solute CT excitations but cannot solve the problem in cases where an intramolecular charge-transfer occurs. A second problem of a general nature is that, within the adiabatic approximation, TDDFT can only describe single excitations.[44,45] Another problematic area for the application of TDDFT concerns excitations for heavy open-shell systems where inclusion of spin–orbit (SO) coupling is necessary already for a qualitative description of the system. Given the increasing interest in the organometallic and inorganic chemistry of molecules containing lanthanides, actinides and heavy transition elements,[46,47] it is desirable to have alternative methods available that can handle such difficult cases.

An attractive way forward would be to combine the flexibility and accuracy of wave function theory (WFT) based methods with the efficiency of DFT. With subsystem methods one could then tackle the problems mentioned above, provided that the subsystem of interest can be chosen small enough to employ an accurate (relativistic) WFT approach. In this case, the FDE scheme offers a very distinct advantage over other schemes, since it employs the electron density, an observable quantity, and thereby avoids complicated problems like the definition of projection operators in calculations where the environment is to be treated by one-component DFT while the active system is described by a 4-component relativistic WFT method. Such a WFT-in-DFT embedding scheme based on the DFT-in-DFT frozen-density embedding scheme has first been proposed by Carter and coworkers,[48–50] where DFT and variational methods such as Hartree–Fock, CASSCF or (multireference)-CI were combined. This approach was mainly used to describe localized properties in solids or surfaces, *e.g.* for the calculation of excitation energies of CO adsorbed on a platinum surface.[51,52]

In this paper, we aim to introduce a simplified and computationally less involved version of this WFT-in-DFT embedding scheme for the calculation of local excitations in large systems. In particular, we apply coupled cluster methods for the treatment of an embedded system and describe the environment by DFT. We test this approach for two different systems. First, as a benchmark application, we revisit the calculation of the solvatochromic shift of acetone in water that was previously performed by Neugebauer *et al.*[26] Second, as an example for a system where the WFT-in-DFT treatment of the excitations is essential, we investigate the *f*–*f* spectrum of neptunyl ($NpO_2^{2+}$) embedded in the $Cs_2UO_2Cl_4$ crystal, following our previous study of the isolated neptunyl ion with relativistic coupled cluster methods.[53]

The outline of the paper is as follows. In section II we present the essential theoretical aspects of WFT-in-DFT embedding methods and outline the proposed scheme. In section III, we then describe our implementation, as well as other computational details. This is followed by the two sample applications of the proposed scheme. In section IVA, the results obtained for acetone solvated in water are presented, and in section IVB, the spectrum of neptunyl embedded in a $Cs_2UO_2Cl_4$ crystal is discussed. Finally, concluding remarks are given in section V.

## II. Theory

In the formulation of the WFT-in-DFT frozen density embedding scheme proposed by Carter and coworkers,[48–51] the total system is partitioned into an embedded subsystem I and its environment, so that the total energy for the system can be expressed as

$$E[\Psi_{\mathrm{I}}^{\mathrm{WFT}}, \rho_{\mathrm{II}}^{\mathrm{DFT}}] = E_{\mathrm{I}}[\Psi_{\mathrm{I}}^{\mathrm{WFT}}] + E_{\mathrm{II}}[\rho_{\mathrm{II}}^{\mathrm{DFT}}] + E_{\mathrm{int}}[\rho_{\mathrm{I}}^{\mathrm{WFT}}, \rho_{\mathrm{II}}^{\mathrm{DFT}}], \tag{2.1}$$

where $E_{\mathrm{I}}$ is the energy of the embedded subsystem I, described using a wavefunction-based method and characterized by its wave function $\Psi_{\mathrm{I}}^{\mathrm{WFT}}$, while $E_{\mathrm{II}}$ is the energy of the environment (subsystem II), described using DFT and characterized by its electron density $\rho_{\mathrm{II}}^{\mathrm{DFT}}$. The interaction energy $E_{\mathrm{int}}$ is defined within DFT as

$$E_{\mathrm{int}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] = E_{NN} + \int \rho_{\mathrm{I}}(\boldsymbol{r}) v_{\mathrm{II}}^{\mathrm{nuc}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \int \rho_{\mathrm{II}}(\boldsymbol{r}) v_{\mathrm{I}}^{\mathrm{nuc}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \frac{1}{2}\int \frac{\rho_{\mathrm{I}}(\boldsymbol{r})\rho_{\mathrm{II}}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{r}' + E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] + T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] \tag{2.2}$$

where the density of subsystem I ($\rho_{\mathrm{I}} \equiv \rho_{\mathrm{I}}^{\mathrm{WFT}}$) is the density obtained from the wave function treatment, while the environment density ($\rho_{\mathrm{II}} \equiv \rho_{\mathrm{II}}^{\mathrm{DFT}}$) is to be obtained from a DFT calculation. In this expression for the interaction energy, $E_{NN}$ is the nuclear–nuclear repulsion energy, $v_{\mathrm{I}}^{\mathrm{nuc}}$ and $v_{\mathrm{II}}^{\mathrm{nuc}}$ are the electrostatic potentials of the nuclei in subsystems I and II, respectively, $E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] = E_{\mathrm{xc}}[\rho_{\mathrm{I}} + \rho_{\mathrm{II}}] - E_{\mathrm{xc}}[\rho_{\mathrm{I}}] - E_{\mathrm{xc}}[\rho_{\mathrm{II}}]$ is the nonadditive part of the exchange–correlation energy, and $T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}}, \rho_{\mathrm{II}}] = T_s[\rho_{\mathrm{I}} + \rho_{\mathrm{II}}] - T_s[\rho_{\mathrm{I}}] - T_s[\rho_{\mathrm{II}}]$ is the nonadditive kinetic energy, where $T_s$ is the kinetic energy of the Kohn–Sham noninteracting reference system.

To include the effect of the environment in the WFT calculation of subsystem I, Carter and coworkers proposed[48] to include an embedding potential given, in analogy to the DFT-in-DFT frozen-density embedding scheme of

Wesolowski and Warshel,[25] by the functional derivative of $E_{\mathrm{int}}^{\mathrm{DFT}}$ with respect to $\rho_I$,

$$v_{\mathrm{eff}}^{\mathrm{emb}}[\rho_{\mathrm{I}},\,\rho_{\mathrm{II}}](\boldsymbol{r}) = v_{\mathrm{II}}^{\mathrm{nuc}}(\boldsymbol{r}) + \int \frac{\rho_{\mathrm{II}}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}\boldsymbol{r}' + \frac{\delta E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_{\mathrm{I}},\,\rho_{\mathrm{II}}]}{\delta\rho_{\mathrm{I}}} + \frac{\delta T_s^{\mathrm{nadd}}[\rho_{\mathrm{I}},\,\rho_{\mathrm{II}}]}{\delta\rho_{\mathrm{I}}}. \tag{2.3}$$

This embedding potential includes the electrostatic potentials of the nuclei and the electron density in the environment, as well as contributions arising from the nonadditive part of the exchange–correlation energy and from the nonadditive kinetic energy. Since this embedding potential depends on the density of subsystem I it has to be updated iteratively during the WFT calculation.

While the theoretical justification of the use of such an embedding potential derived from DFT in wave function based calculations was previously debated,[54,55] Wesolowski recently showed[56] that for exact density functionals and a full CI expansion of the wave function, the above approach will indeed lead to the exact total electron density $\rho_{\mathrm{I}} + \rho_{\mathrm{II}}$. If a truncated expansion of the wave function is employed, an additional term that corrects for the deficiency of the WFT description, could be included in the embedding potential to keep the theory formally exact. Since such a term will be the less important the closer the WFT description is to full CI and, moreover, since it is neither possible nor desirable to correct the error introduced by a poor WFT description by means of DFT, we consider it well justified to neglect this term and simply employ the uncorrected embedding potential given above.

In the practical application of the WFT-in-DFT embedding scheme described above, several additional approximations will be introduced. First, neither the DFT treatment nor the WFT treatment can be exact. In the DFT case, this is due to the well-known deficiencies in the currently available exchange–correlation functionals, in the WFT case due to the infeasibility to employ a full CI expansion in a large basis for anything but the smallest model systems. Second, the nonadditive kinetic energy component of the embedding potential is evaluated using an approximate kinetic energy functional. Several applications of the DFT-in-DFT frozen-density embedding scheme show that the available kinetic energy functionals provide a reliable description, in the case of weakly bonded complexes[57–59] and solute–solvent interactions[26–28,60,61] or in simple solid-state systems.[30,31,62–66] For strongly bound covalent systems one needs to resort to other solutions, *e.g.*, employ a 3-partitioning scheme introducing capping groups with a constrained electron density that has recently been developed by two of us.[67]

For efficiency reasons it is furthermore common to employ approximations in the construction of the electron density of the environment $\rho_{\mathrm{II}}$. In the simplest case, the electron density is obtained as a sum of fragment densities that is kept completely frozen in the following calculations. This choice is still in principle exact as long as the density of the environment is everywhere smaller than the exact total density of the full system. In case of strong polarization, the simple sum-of-fragments approximation for the environment density may easily violate the latter condition. In such cases it is necessary to consider the relaxation of the environment under the influence of the embedded subsystem. As for DFT-in-DFT embedding, this is possible by employing so-called freeze-and-thaw cycles,[68] *i.e.*, by interchanging the roles of the two subsystems and updating the density of the environment in a DFT calculation that includes the effect of subsystem I *via* the embedding potential $v_{\mathrm{eff}}^{\mathrm{emb}}[\rho_{\mathrm{II}}, \rho_{\mathrm{I}}]$. This can be repeated iteratively until convergence is reached.

For the application of WFT-in-DFT embedding to atoms and molecules absorbed on surfaces, Carter and coworkers proposed different simplified schemes for performing the freeze-and-thaw cycles and obtaining a self-consistent embedding potential. In their initial work,[49,51,52] the total density $\rho_{\mathrm{tot}} = \rho_{\mathrm{I}} + \rho_{\mathrm{II}}$ was obtained from a DFT calculation with periodic boundary conditions. This total density was then kept fixed, while $\rho_{\mathrm{I}}$ and $\rho_{\mathrm{II}}$ were updated subsequently, based on the density $\rho_{\mathrm{I}}^{\mathrm{WFT}}$ of the WFT calculation of the absorbed molecule. In later work,[50] they modified this scheme such that the density $\rho_{\mathrm{II}} = \rho_{\mathrm{tot}} - \rho_{\mathrm{I}}^{\mathrm{bare}}$ is kept frozen, where $\rho_{\mathrm{tot}}$ is again obtained from a DFT calculation, while $\rho_{\mathrm{I}}^{\mathrm{bare}}$ is taken from a WFT calculation on the isolated subsystem I.

However, for the calculation of excitation energies using WFT-in-DFT embedding, it may be that a simpler approach can be employed. In many cases where TDDFT is known to fail, such as for charge-transfer excitations or for open-shell systems with close-lying excited states, it is still possible to obtain an accurate description of the ground-state density from a DFT treatment. When the ground state is well described by DFT, it should be possible to first calculate an approximation of both $\rho_{\mathrm{I}}$ and $\rho_{\mathrm{II}}$ using DFT-in-DFT embedding (possibly employing freeze-and-thaw cycles) and to subsequently use the embedding potential constructed using these densities in the WFT calculation of the excitation energies of the embedded system. This is quite different from the cases studied by Carter and coworkers, who investigated excitation energies of molecules absorbed on surfaces,[51,52] where already the description of the ground-state in DFT is problematic.[49]

This means that in eqn (2.3), instead of the density $\rho_{\mathrm{I}}^{\mathrm{WFT}}$ obtained from the WFT calculation, the DFT density for subsystem I, $\rho_{\mathrm{I}}^{\mathrm{DFT}}$, obtained from a DFT-in-DFT embedding calculation, is used. If such a simplified treatment is justified, it offers several advantages over the self-consistent WFT-in-DFT schemes described above. First, only one computationally expensive WFT calculation is required, compared to the multiple calculations required to converge a freeze-and-thaw procedure. Second, it is not necessary to generate the density $\rho_{\mathrm{I}}^{\mathrm{WFT}}$, thereby avoiding a non-trivial and computationally expensive step in non-variational WFT approaches that employ intermediate normalization. Finally, also the embedding potential has to be calculated only once, instead of updating it during the iterative solution of the Hartree–Fock and CC equations. This latter simplification is similar to the linearization of the embedding potential recently proposed in DFT-in-DFT frozen-density embedding.[69]

The calculation of excitation energies in particular (or of response properties in general) can be carried out with a generalization of the theory outlined above. We again make the assumption that the WFT method will be able to yield a

close enough approximation to the exact density response of the subsystems. While in the case of static embedding theory this is in principle true for a full CI calculation in a saturated basis, we now have the additional limitation that we can only account for the response of the active system.

This limitation is usually also present in implementations of DFT-in-DFT, although Neugebauer[32] has recently presented a sophistication of the theory by Casida and Wesolowski[70] that makes it possible to treat also the case of coupled excitations in extended systems. For our initial application we consider the latter scheme and even make a further approximation in neglecting any variations on the embedding potential due to the response of the active subsystem to the external perturbations, which for TDDFT excitation energies are incorporated by including the derivative of the embedding potential in the kernel.[70] This means that in effect we always consider a fixed embedding potential based on the ground-state density.

This additional approximation is valid as long as the non-additive kinetic and exchange–correlation potentials of the initial and final states are not too different, which will be the case if the excitation is truly localized in the interior of the embedded subsystem, as is known to be the case in the applications considered in this work.

## III. Implementation and computational details

In our implementation, we use a combination of different quantum chemical program packages. In the first step, ADF[71,72] and the implementation of the DFT-in-DFT frozen-density embedding scheme in this package[27,73] are employed for obtaining an approximation to the density of subsystem I, $\rho_{\mathrm{I}}^{\mathrm{DFT}}$, and to calculate the embedding potential $v_{\mathrm{eff}}^{\mathrm{emb}}[\rho_{\mathrm{I}}^{\mathrm{DFT}}, \rho_{\mathrm{II}}]$. The values of this embedding potential on the points of the integration grid generated by ADF are stored in a file for later use in the WFT calculation.

In the present work, we will always start by approximating the density of the environment as a sum of the densities of molecular or atomic fragments that were calculated for the isolated fragments. As fragments we either used the distinct water molecules in the solvent environment of acetone in water, or the $UO_2Cl_4^{2-}$, $Cs^+$, and $Cl^-$ units that constitute the $Cs_2UO_2Cl_4$ crystal environment. Details on the construction of the environment density will be given below. In the case of acetone solvated in water, the simplest approximation of the solvent environment as a sum of fragments was previously shown to be a sufficiently accurate description.[26] For the crystal environment, a detailed study of the different possibilities for constructing the environment density is outside the scope of this work, and will be addressed in another publication that deals with excitations in the pure $Cs_2UO_2Cl_4$ crystal.[74] We first consider the sum of fragments approach described above and improve on this simple approximation, by employing the freeze-and-thaw procedure only for the nearest-neighbor $Cl^-$ ions.

In the calculation of the fragments used for the construction of the frozen density, we used for the solvent water environment the local-density approximation (LDA)[75] for the exchange–correlation potential and a DZP basis set from the ADF basis set library.[71,76] For the fragments used to model the $Cs_2UO_2Cl_4$ crystal environment, we used the statistical averaging of model orbital potentials (SAOP),[77] in combination with a TZ2P basis set. In this case, scalar relativistic effects were included *via* the zeroth-order regular approximation (ZORA).[78] In all cases, we employed a spin-restricted closed-shell description of the frozen environment density.

To model the density of the active system in the DFT-in-DFT frozen-density embedding calculations for acetone and neptunyl we employed the SAOP functional and the TZ2P basis set. The nonadditive kinetic-energy component of the embedding potential was modeled by the PW91k functional,[79] while the exchange–correlation component was treated using the Becke–Perdew–Wang (BPW91) exchange–correlation functional.[80,81] For the neptunyl calculations, spin-unrestricted DFT calculations were performed, and scalar relativistic effects were included *via* the zeroth-order regular approximation (ZORA).[78]

After generating the embedding potential, we used a locally modified version of DALTON[82] for nonrelativistic WFT calculations and a development version of DIRAC[83] for relativistic WFT calculations. In the nonrelativistic case, matrix elements of $v_{\mathrm{eff}}^{\mathrm{emb}}(r)$ are constructed as

$$v_{ij} = \langle \phi_i | v_{\mathrm{eff}}^{\mathrm{emb}} | \phi_j \rangle \approx \sum_k w_k v_{\mathrm{eff}}^{\mathrm{emb}}(\boldsymbol{r}_k) \varphi_i(\boldsymbol{r}_k) \varphi_j(\boldsymbol{r}_k), \qquad (3.4)$$

where $\varphi_i(\boldsymbol{r}_k)$ is the value of orbital $\varphi_i$ evaluated at grid point $\boldsymbol{r}_k$ and $w_k$ is the integration weight associated with this grid point. It should be noted that we employ the accurate numerical integration grid used by ADF, which is also able to integrate the electrostatic terms appearing in the embedding potential accurately.[27,84] In a relativistic framework the embedding potential, being a scalar potential, enters the one-electron Hamiltonian as a diagonal operator expressed (in 2-component form) as

$$\boldsymbol{v}_{\mathrm{eff}}^{\mathrm{emb}}(\boldsymbol{r}) = \begin{pmatrix} v_{\mathrm{eff}}^{\mathrm{emb}}(\boldsymbol{r}) & \mathbf{0} \\ \mathbf{0} & v_{\mathrm{eff}}^{\mathrm{emb}}(\boldsymbol{r}) \end{pmatrix}. \qquad (3.5)$$

Therefore, the matrix elements of the embedding potential with respect to spinors $\phi_i(r)$ and $\phi_j(r)$ are given by

$$\begin{aligned} v_{ij} &= \langle \phi_i | \boldsymbol{v}_{\mathrm{eff}}^{\mathrm{emb}} | \phi_j \rangle \\ &\approx \sum_k w_k v_{\mathrm{eff}}^{\mathrm{emb}}(\boldsymbol{r}_k) [\varphi_i^L(\boldsymbol{r}_k) \varphi_j^L(\boldsymbol{r}_k) + \varphi_i^S(\boldsymbol{r}_k) \varphi_j^S(\boldsymbol{r}_k)] \end{aligned} \qquad (3.6)$$

where $\varphi_{i,j}^L(\boldsymbol{k})$ and $\varphi_{i,j}^S(\boldsymbol{k})$ are scalar functions for the large and small components, respectively. Once the embedding potential matrix is set up, it is added to the one-electron Fock matrix in the WFT calculation, like any other one-electron operator.

For the calculation of the excitation energies of acetone solvated in water, we employed the CC2 method[85–87] as implemented in DALTON.[82] The aug-cc-pVDZ basis set was used for acetone,[88] as it has previously been shown[89] that CCSD excitation energies are essentially converged with this basis sets, and the same behavior is expected for CC2. In the CC2 calculations the 1s orbitals of all atoms are kept frozen.

The calculations of the excitation spectrum of neptunyl were performed with DIRAC[83] using the exact two-component (X2C) approach recently outlined by Iliaš and Saue,[90] with spin–orbit coupling being included in the latter *via* mean-field

integrals, calculated with the AMFI code.[91,92] A valence double zeta basis by Dyall[93] was used for Np whereas the aug-cc-pVTZ basis[88] was used for O (and for Cl where applicable). In the correlated calculations the intermediate Hamiltonian Fock-space coupled-cluster (IHFSCC) method[94–96] was used. In the IHFSCC calculations we start from a closed shell ($NpO_2^{3+}$ or $NpO_2Cl_4^{-}$) species and obtain the excitation spectrum of the target system ($NpO_2^{2+}$ or $NpO_2Cl_4^{2-}$) from the (0*h*,1*p*) sector. We have used a slightly smaller active space than employed previously for $NpO_2^{2+}$,[53] consisting of about the same number of virtual spinors (75 and 76, respectively) for both species in the correlated calculations. Of these, 6 are kept in the $P_m$ space and 19 in the $P_1$ space. The number of occupied spinors included for $NpO_2^{2+}$ and $NpO_2Cl_4^{2-}$ are 12 and 28, respectively. We note that due to the presence of the frozen environment, the symmetry in the wavefunction calculations is reduced to $C_1$, and the IHFSCC calculations are, therefore, significantly more expensive than the original $NpO_2^{2+}$ calculations that could exploit the $D_{\infty h}$ symmetry.

In order to automate both the initial DFT-in-DFT calculations and the interoperation of the different quantum chemical packages, we made use of the recently developed PyAdf scripting framework.[97]

## IV. Results and discussion

### A Solvatochromic shifts of acetone in water

The determination of solvatochromic shifts of acetone in water is an ideal benchmark for our method because it has been extensively studied with a range of WFT methodologies, taking into account solvent effects in different ways (see ref. 89 and references therein). Furthermore, Neugebauer *et al.* have previously performed a systematic study[26] on this system using DFT-in-DFT frozen-density embedding.

We have employed the set of geometries obtained in ref. 26 from Car–Parrinello molecular dynamics (CP-MD) simulations of both the gas-phase and the solution. From these simulations, 300 and 220 snapshots were retained for the calculations in the gas-phase and in solution, respectively (Fig. 1). The excitation energies of the n → π transition are then calculated as a weighted average over these snapshots, using the oscillator strengths as weight factors.

We note that our reference DFT-in-DFT calculations differ slightly from the values obtained by Neugebauer and co-workers as we could apply a larger basis set than was feasible with the previous implementation of DFT-in-DFT that was used in ref. 26. Irrespective of these small differences, one of the findings of Neugebauer *et al.* that is very relevant for this work has to do with the construction of the density for the frozen subsystem: they observed nearly identical solvent shifts (differences of about 0.01 eV or less) when the embedding potential was constructed from an approximate density, made up by the superposition of densities obtained for isolated water molecules, compared to when densities from calculations including all waters at once were used to obtain the embedding potential. For this system, different functionals used to produce the density of these unperturbed solvent molecules were proven to yield similar results, which is why we chose the simple LDA functional for this purpose.

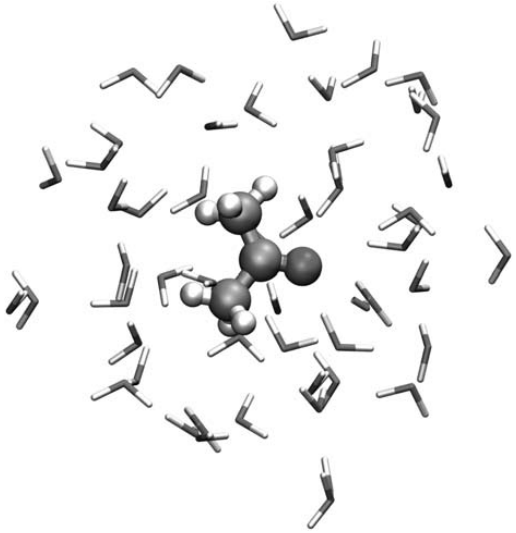


**Fig. 1** CP-MD snapshot for the simulation of acetone in water.

The calculated excitation energies are summarized in Table 1. Inspecting first the CC2 results, one observes a very good agreement with experiment for the solvent shift. This indicates that the embedding potential generated in the underlying DFT-in-DFT calculations does indeed provide a realistic representation for the solute–solvent interaction. Moreover, it also indirectly validates our initial assumptions, that is, that the ground state density of the active subsystem calculated with DFT is nearly identical to the corresponding wavefunction-based density and that the direct contribution of the environment to the response is negligible. To verify the latter assumption we also carried out additional DFT-in-DFT calculations in which the embedding contributions to the kernel were switched off.[28] These calculations indicate that this contribution amounts to 1 meV for the lowest n → π excitation energy that we consider in this work. Other excitations show somewhat larger shifts, but these are generally at least one order of magnitude smaller than the corresponding solvatochromic shift. The state for which the neglect of the response contribution has the largest effect (about 0.03 eV) is the fifth singlet excited state, located at about 8 eV.

The quality of the TDDFT results with the DZP/TZ2P combination appears to be not as good as for the CC2 calculations with as most significant difference an underestimation of the solvent shift. Furthermore, one sees significant basis set effects on the TDDFT results if a crude description for the solvent (DZ) is used, as done previously,[26] with a fortuitous cancellation of errors in the shift for the DZ calculation. One should thereby note, however, that for CC2 both the gas-phase and solution excitation energies are underestimated in comparison to the corresponding experimental values, putting the TDDFT results closer to experiment. Since roughly identical discrepancies are found for results in the

**Table 1** TD-DFT and CC2 n → π excitation energies ($\langle\omega_i\rangle$, in eV) for acetone in gas-phase and solution, calculated as oscillator strength-weighted averages over the CP-MD snapshots, together with the corresponding solvatochromic shifts. For comparison, the CCSD results of Aidas and coworkers[89] as well as experimental results are also shown

| | Basis sets | | $\langle\omega_i\rangle$ | | |
|---|---|---|---|---|---|
| Method | Acetone | Water | Gas phase | Solution | Shift |
| TD-DFT | TZP | DZ | 4.464 | 4.667 | 0.203 |
| | TZ2P | DZP | 4.471 | 4.636 | 0.165 |
| CC2 | TZ2P/aug-cc-pVDZ | DZP | 4.350 | 4.546 | 0.196 |
| CCSD (ref. 89) | | | 4.491 | 4.686 | 0.195 |
| Exp. (ref. 106–108) | | | 4.48–4.49 | 4.68–4.69 | 0.19–0.21 |



gas-phase and solution, this cannot be attributed to inaccuracies in the embedding procedure.

A possible explanation for the underestimated CC2 excitation energies is the sensitivity to the chosen structures of the acetone molecule. The results of Aidas and coworkers[89] show a significant dependence of the excitation energies on the C=O bond distance, where larger distances lower the excitation energies. To test this, we performed a series of test calculations at the equilibrium structures that were obtained using DFT-BLYP (the functional that was used in the CP-MD simulations), DFT-B3LYP, and CCSD. The equilibrium structures and corresponding CC2 excitation energies are summarized in Table 2. The results show that DFT-BLYP overestimates the C=O bond distance, which leads to an underestimation of the CC2 excitation energy. Since the CP-MD structures that we used were based on a simulation performed with the BLYP functional this explains part of the error in the comparison to the experimental data. Since the error is due to a bond length that is found consistently too long in both the gas-phase and in solution structures, only the absolute values but not the solvatochromic shift are affected.

### B *f–f* spectrum of $Cs_2U(Np)O_2Cl_4$

As a second, more challenging application, we apply WFT-in-DFT embedding to the calculation of the *f–f* spectrum of the neptunyl cation ($NpO_2^{2+}$). The spectrum of the isolated neptunyl cation was investigated previously in our group.[53] The available experimental neptunyl spectra were taken from impurities in host crystals such as $Cs_2UO_2Cl_4$ or $CsUO_2(NO_3)_3$.[98,99] To be able to compare directly to these data one thus needs to model the host crystal as well. There has been one such attempt before by Matsika and Pitzer,[100] who performed a calculation on $Cs_2U(Np)O_2Cl_4$ using a cluster model. In their treatment, a central $NpO_2Cl_4^{2-}$ unit is treated using SO-CI, with the six nearest-neighbor caesium atoms described by all-electron model potentials and all other species up to 25 Å away from the central actinyl unit as point charges.

**Table 2** CC2 n → π excitation energies ($\omega_{n\rightarrow\pi}$, in eV) for acetone in the gas-phase at equilibrium geometries obtained with different methods. All calculations have been performed using the aug-cc-pVDZ basis set

| Geometry | $r_{CO}$/Å | $\omega_{n\rightarrow\pi}$/eV |
|---|---|---|
| BLYP | 1.2302 | 4.442 |
| B3LYP | 1.2175 | 4.535 |
| CCSD | 1.2217 | 4.505 |

To include the effect of the crystal environment in our relativistic IHFSCC calculations,[53] we employ a similar partitioning as used in recent applications of DFT-in-DFT frozen-density embedding for transition metal- and lanthanide-containing solids.[30,31] In our case we devise a cluster model, where the active subsystem is treated with relativistic IHFSCC theory and DFT is used to model the nearby ions that may overlap with the impurity (Fig. 2). The Madelung potential arising from the rest of the crystal was evaluated using formal charges placed at the positions given by the X-ray structure of the pure crystal.[101] We thereby utilized[102] the program ENV[103,104] to determine the extent of the intermediate region encapsulating the central active subsystem as well as an array of surrounding point charges that describe the Madelung potential for the crystal. This intermediate region always comprises 20 $UO_2Cl_4^{2-}$ and 90 $Cs^+$ ions. In the central unit we have replaced the uranium by neptunium, and adjusted the Np–O and Np–Cl bond lengths to those from the X-ray structure for $NpO_2Cl_4^{2-}$.[105] We explored two different possibilities for the central unit: (a) one where it was split into $NpO_2^{2+}$ and $Cl_4^{4-}$, and only $NpO_2^{2+}$ was treated using IHFSCC, while the ligands were taken into account with different degrees of sophistication (from simple point charges to densities fully relaxed in freeze-and-thaw cycles); and (b) one where the entire $NpO_2Cl_4^{2-}$ is calculated with IHFSCC.

Our results are summarized in Table 3, which displays the symmetry classification and excitation energies for the different *f–f* states following the ordering found in the experimental

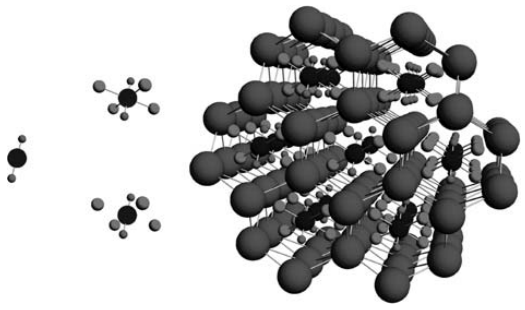


**Fig. 2** Steps in the WFT-in-DFT embedding calculation on the *f–f* spectrum of $Cs_2U(Np)O_2Cl_4$. From left to right, (a) the initial $NpO_2^{2+}$ ion (b) is surrounded by four chlorides (top: chlorides considered explicitly in the WFT calculations; bottom: ligand group $Cl_4^{4-}$ is taken as part of the environment and included *via* the embedding potential) to make up the central unit in the cluster model; and (c) this unit is embedded into a larger environment. The larger model is then embedded in an array of point charges (not shown).

**Table 3** *f–f* excited state energies (in $cm^{-1}$) for $NpO_2^{2+}$ and $NpO_2Cl_4^{2-}$, obtained with and without the inclusion of environment effects *via* WFT-in-DFT embedding. All calculations were performed using the X-ray structure reported in ref. 101 (where $r_{NpO}$ = 1.775 and $r_{NpCl}$ = 2.653 Å, respectively) unless otherwise noted. For comparison, the results of Matsika and Pitzer[100] for the their embedded cluster calculations are shown, together with the experimental values due to Denning and coworkers[98]

| | | Excited electronic states | | | | |
|---|---|---|---|---|---|---|
| Central unit | Surrounding | I | II | III | IV | V |
| $NpO_2^{2+}$ | — | $\Delta_{3/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ |
| IHFSCC[a] | — | 3221 | 8565 | 7225 | 30 877 | 34 947 |
| IHFSCC | — | 4297 | 9661 | 7229 | 29 021 | 32 379 |
| $NpO_2^{2+}$ $L_4^{-4}$ | | $\Delta_{3/2u}$ + $\Phi_{5/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ |
| IHFSCC point charges | — | 2093 | 8032 | 7828 | 23 326 | 26 321 |
| IHFSCC DFT (frozen)[b] | — | 2243 | 8150 | 7677 | 23 323 | 26 433 |
| IHFSCC DFT (relaxed)[c] | — | 1034 | 7307 | 8029 | 20 390 | 23 303 |
| $NpO_2Cl_4^{2-}$ | | $\Delta_{3/2u}$ + $\Phi_{5/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ |
| IHFSCC | — | 886 | 7679 | 9262 | 20 018 | 22 445 |
| IHFSCC | DFT[d] | 1156 | 7738 | 9137 | 20 857 | 26 305 |
| $NpO_2Cl_4^{2-}$ (ref. 100) | | $\Phi$ + $\Delta$ | $\Delta$ + $\Phi$ | $\Phi$ | $\Pi$ | $\Pi$ |
| SO-CI | AIMP[e] | 1663 | 5775 | 8463 | 18 367 | 20 575 |
| Exp. (ref. 98) | | $\Delta_{3/2u}$ + $\Phi_{5/2u}$ | $\Delta_{5/2u}$ | $\Phi_{7/2u}$ | $\Pi_{1/2u}$ | $\Pi_{3/2u}$ |
| | | 900–1050 | 6880 | 7890 | 17 241 | 20 081 |

[a] Using the gas-phase geometry with $r_{NpO}$ = 1.675 Å (from the calculations of ref. 53). [b] Frozen density for the ligand group ($L_4^{-4}$) constructed as superposition of the densities of four $Cl^-$ ions. [c] Frozen density for the ligand group ($L_4^{-4}$) obtained for $Cl_4^{-4}$ after 10 freeze-and-thaw cycles. [d] $[Cs_{90}(UO_2Cl_4)_{20}]^{+50}$ in the presence of point charges that counterbalance its and the active subsystem's charge. [e] Model potential obtained for the six-nearest caesium atoms plus a point charge array.

assignment of these transitions. Considering first the case where only the bare $NpO_2^{2+}$ (with the Np–O bond length from the X-ray geometry[105]) unit is chosen as central unit, we see that these gas-phase results tend to strongly overestimate the excitation energies in comparison to experiment (about 3000 $cm^{-1}$ for states **I** and **II** and about 12 000 $cm^{-1}$ for states **IV** and **V**) with the exception of state **III**, for which gas-phase results underestimate the excitation energy by only 650 $cm^{-1}$. Compared to the energies at the optimized equilibrium geometry for the gas-phase (where $r_{Np-O}$ = 1.675 Å), the values at the X-ray geometry are about 1000 $cm^{-1}$ higher for **I** and **II**, and lower for **IV** and **V** by roughly the same amount. The excitation energy for state **III** is, however, almost not affected by the change in distance. This is easily explained by the fact that this is the pure $\Phi_{7/2u}$ state, that is fully localized on the Np atom and forms the higher spin–orbit component of the $\Phi_{5/2u}$ ground state. The other states have some covalent character that makes them more susceptible to changes in the distance to the oxo-groups.

Before considering the full cluster model, we first consider the equatorial ligands that will have the strongest effect on the excitation energies. The ligand field breaks the linear symmetry of the actinyl and induces mixing between the lowest spin–orbit components of the $\Delta$ and $\Phi$ states, while charge-transfer from the formally negative chloride ions to the neptunium center of the dication weakens the axial bonds of the neptunyl. In order to assess the relative importance of these effects, we present in Table 3 calculations with varying degrees of sophistication in the treatment of this ligand group.

In the simplest case the $NpO_2^{2+}$ species is surrounded by four point charges at the position of the chlorides, each with a negative unit charge. This has a very strong effect on four of the levels, with downward shifts of 2204 $cm^{-1}$ for state **I**, 1629 $cm^{-1}$ for **II**, 5695 $cm^{-1}$ for **IV** and 8626 $cm^{-1}$ for **V**, respectively. State **III** is shifted upwards by 599 $cm^{-1}$ by the presence of the point charges. From the composition of these states in the FSCC wave function, we see that there is a considerable mixing between the $\Delta_{3/2u}$ and $\Phi_{5/2u}$ states in the ground and first excited states, while the others remain pure states. Compared to the experimental results, all excitation energies are improved significantly by this simple electrostatic model. We see that apart from **III** the calculated excitation energies are too high by about 1000 $cm^{-1}$ for **I** and **II** and about 5000 $cm^{-1}$ for **IV** and **V**. State **III** is very close to the experimental value, being only 62 $cm^{-1}$ lower.

More sophisticated models for the ligands where these are represented by an embedding potential, are considered next. The first approach is to have a frozen density constructed as a superposition of atomic densities for the chloride ions. From our results it can be seen that this simplest FDE model hardly changes the point charge picture—in fact, there is even a slightly worse agreement with experiment for all states but **IV**, by about 100–150 $cm^{-1}$. The problem comes from the fact that the charge density for the $Cl^-$ ligands is by construction spherical so that the significant deformation and charge transfer towards the actinyl ion should be accounted for in the calculation of the active system. This is not possible as one can in FDE describe only the flow of charge towards the frozen system and not that from the frozen system. To remedy this problem we have performed 10 freeze-and-thaw cycles to allow for such relaxation. After this procedure a completely different picture emerges. Excitation energies for states with $\Delta$ and/or $\Phi$ character agree within 100–450 $cm^{-1}$ with the experimental data, whereas the error for the $\Pi$ states is reduced to about 3200 $cm^{-1}$. This leads to a much better overall agreement with experiment, with now a correct ordering of the states **II** and **III**. The only exception is the **III** state that is moved to somewhat higher energies by the relaxation of the density.

To make a precise assessment of the quality of the embedding it is, however, better to not compare it directly to the experimental data but compare the subsystem treatment of the ligands in the IHFSCC calculation to a supermolecular calculation of $NpO_2Cl_4^{2-}$. This case can be seen as an exact reference for the above embedding calculations, since in both cases no outer environment was included. Moving from the embedding calculation where the $Cl^-$ densities are relaxed, to a supermolecular calculation has the overall result of a slight increase of the excitation energies of all states: state **I** is shifted by about 150 $cm^{-1}$, while shifts of about 370 $cm^{-1}$ are seen for **II** and **IV**. For **III** and **V** the shifts lie in the 860–1200 $cm^{-1}$ range. This indicates that in this case, where the actinide–chloride bond is known to have significant ionic character, the WFT-in-DFT embedding scheme reproduces the main features of the fully wavefunction-based description.

With the results of the calculations including the ligands, we can summarize the general aspects of modeling the actinyl–ligand interaction: adding point charges on the equatorial plane introduces mixing of the $\Phi_{5/2u}$ and $\Delta_{3/2u}$ states, but underestimates the relative destabilization of the $f^{\phi}$ orbitals relative to the $f^{\delta}$ or $f^{\pi}$ ones. Placing spherical finite volume $Cl^-$ ions in place of the point charges does not change this picture. Allowing the charge density of the ligands to polarize gives the correct picture with the FDE freeze-and-thaw description underestimating slightly the relative destabilization of the $f^{\phi}$ orbitals. This is probably due to incomplete description of the orbital interactions between the chloride ligands and the metal orbitals in the density-only embedding description.

The inclusion of the rest of the crystal environment as an embedding potential to the calculation on $NpO_2Cl_4^{2-}$, brings only modest changes to the excitation energies, usually by no more than 100–200 $cm^{-1}$ for the states with $\Delta$ and $\Phi$ character, a strong evidence that the bulk of the environment effects come from the equatorial ligands rather than from electrostatic interaction with the structural units further away. It is interesting to note that the environment seems to stabilize the states with dominant $\Phi$ character and destabilize states with dominant $\Delta$ character. This may be the effect of the 12 $Cs^+$ ions adjacent to the central unit that will polarize the equatorial ligands, thereby drawing charge away from the $f^{\phi}$ orbitals. The $\Pi$ states experience differential environmental effects (about 800 $cm^{-1}$ for **IV** but about 4000 $cm^{-1}$ for **V**).

We conclude this section by noting that the calculated vertical excitation energies in this proof-of-principle application should of course not be blindly compared to the adiabatic excitation energies that are measured in experiment. The deviation between the two will be largest for states that are structurally different from the ground-state, in particular the $\Pi$ states, where the excitation moves the electron from a non-bonding to a partially antibonding orbital. In these cases (states **IV** and **V**) the adiabatic excitation energies are likely to be significantly lower. In comparison to the results of Matsika and Pitzer, who did consider adiabatic effects in their calculations, our embedded cluster calculations are able to better approach the experimental results for the lowest excited states (**I** and **II**) but agree less well from **III** onwards as should be expected. Finally, another possible source of errors relative to the experimental data comes from the coupled cluster treatment for which the active space was chosen to be somewhat smaller than previously used in our benchmark gas phase calculations. The latter calculations indicate that the effect of increasing electron correlation in the calculations, could also improve the agreement between the WFT-in-DFT embedding calculations and the experimental results as correlation appears to decrease the excitation energies for most states.

## V. Conclusions

We have presented here a simple scheme to incorporate the effect of a frozen environment treated using DFT in the wavefunction-based calculation of excitation energies. In contrast to previously described WFT-in-DFT schemes,[50,52] we do not use the electron density of the active subsystem to update the embedding potential. Instead, we assume that the ground state density obtained with DFT is identical to the density that will arise from a correlated *ab initio* treatment. In cases were TDDFT cannot be applied, such as charge-transfer excitations or open-shell systems with close-lying excited states, this assumption is often justified.

We have applied this WFT-in-DFT embedding scheme in two proof-of-concept applications, the calculation of the solvatochromic shift of acetone in water and the spectrum of $NpO_2^{2+}$ embedded in a $Cs_2UO_2Cl_4$ crystal. For acetone in water, we show that the embedding potential is able to correctly describe the effect of the environment on the $n \rightarrow \pi$ excitation energies both in TDDFT and in CC2 calculations. The efficiency of the WFT-in-DFT embedding scheme makes it possible to perform CC2 calculations for acetone surrounded by a solvent shell of water for 220 different snapshots obtained from an MD simulation, thus making it possible to accurately include the effects of the dynamics on the solvatochromic shift.

For $NpO_2^{2+}$ embedded in a $Cs_2UO_2Cl_4$ crystal, we find that our WFT-in-DFT embedding scheme is able to incorporate the effect of the crystal environment in IHFSCC calculations of the electronically excited states. In particular, the embedding scheme is able to closely reproduce the spectrum of $NpO_2Cl_4^{2-}$ calculated within a fully wavefunction-based treatment, provided we allow the density of the chloride ligands to be polarized. While such an agreement depends on the localized nature of the transitions under consideration, it introduces an economical yet highly accurate way to compute the *f–f* spectra of actinyl ions in complex environments because they can be treated in a 3-atom wave function model.

In both applications, which consider two very different environments, we see that the WFT-in-DFT FDE method is capable of accurately representing the environment. The remaining discrepancies mainly originate from intrinsic errors in the description of the subsystems (such as incomplete basis sets, the degree of electron correlation recovered by the wavefunction-based methods) while deficiencies of the embedding procedure itself are small. This indicates that our initial assumption that the ground-state density of the embedded system is described accurately by DFT is indeed valid for the investigated systems. This makes the WFT-in-DFT approach an interesting and cost-effective solution for applications where DFT is known to yield accurate densities but TDDFT

fails (such as charge-transfer excitations or open-shell systems).

Further work is necessary in situations where DFT is not able to accurately describe the ground state density of the active part, or in cases in which the excitation studied does significantly change the density overlap between the embedded system and its surrounding environment. We are currently working on a generalization of the implementation that allows proper treatment of such cases as well.

## Note added in proof

Very recently, an inconsistency in the ADF implementation of the PW91k functional for the spin-restricted case was reported in ref. 109. Although the effect on our results turned out to be insignificant, all data in this paper were corrected accordingly.

## Acknowledgements

The authors thank The Netherlands Organization for Scientific Research (NWO) for financial support *via* the TOP and the Vici programs and acknowledge computer time provided by the Dutch National Computing Facilities (NCF). We would further like to express our gratitude to Dr Manuel Louwerse (VU Amsterdam) and Dr Johannes Neugebauer (ETH Zurich) for kindly providing us with the geometric data for acetone in the gas-phase and solution, as well as to Dr Florent Réal (PhLAM, Université de Lille 1/CNRS) for generating the embedded cluster coordinates and the Madelung potential for $Cs_2UO_2Cl_4$.

## References

1 F. Furche and D. Rappoport, in *Computational Photochemistry*, ed. M. Olivucci, Computational and Theoretical Chemistry, Elsevier, Amsterdam, 2005, ch. 3, vol. 16.
2 A. Dreuw and M. Head-Gordon, *Chem. Rev.*, 2005, **105**, 4009–4037.
3 B. Ordejón, M. Karbowiak, L. Seijo and Z. Barandiarán, *J. Chem. Phys.*, 2006, **125**, 074511.
4 M. E. Casida, in *Recent Advances in Density-Functional Methods*, ed. D. P. Chong, World Scientific, Singapore, 1995, pp. 155–192.
5 R. van Leeuwen, *Int. J. Mod. Phys. B*, 2001, **15**, 1969.
6 C. Hättig and A. Köhn, *J. Chem. Phys.*, 2002, **117**, 6939–6951.
7 D. Kats, T. Korona and M. Schütz, *J. Chem. Phys.*, 2006, **125**, 104106.
8 D. Kats, T. Korona and M. Schütz, *J. Chem. Phys.*, 2007, **127**, 064107.
9 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227–249.
10 P. Sherwood, in *Modern Methods and Algorithms of Quantum Chemistry*, ed. J. Grotendorst, NIC Series, , John von Neumann Institute for Computing, Jülich, 2000, vol. 1, pp. 257–277.
11 J. Gao, in *Reviews in Computational Chemistry*, ed. K. B. Lipkowitz and D. B. Boyd, VCH, New York, 1995, vol. 7, pp. 119–185.
12 A. Warshel and Z. T. Chu, *J. Phys. Chem. B*, 2001, **105**, 9857–9871.
13 R. Rajamani and J. Gao, *J. Comput. Chem.*, 2001, **23**, 96–105.
14 M. A. L. Marques, X. López, D. Varsano, A. Castro and A. Rubio, *Phys. Rev. Lett.*, 2003, **90**, 258101.
15 A. Sergi, M. Grüning, M. Ferrario and F. Buda, *J. Phys. Chem. B*, 2001, **105**, 4386–4391.
16 U. F. Röhrig, I. Frank, J. Hutter, A. Laio, J. VandeVondele and U. Rothlisberger, *ChemPhysChem*, 2003, **4**, 1177–1182.
17 S. Humbel, S. Sieber and K. Morokuma, *J. Chem. Phys.*, 1996, **105**, 1959–1967.
18 M. Svensson, S. Humbel, R. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem.*, 1996, **100**, 19357–19363.
19 L. Seijo and Z. Barandiaran, *J. Chem. Phys.*, 2004, **121**, 6698–6709.
20 L. Seijo, Z. Barandiaran and J. M. Soler, *Theor. Chem. Acc.*, 2007, **118**, 541–547.
21 Z. Barandiaran and L. Seijo, *J. Chem. Phys.*, 1988, **89**, 5739–5746.
22 L. Seijo, Z. Barandiaran and E. Harguindey, *J. Chem. Phys.*, 2001, **114**, 118–129.
23 B. Ordejon, L. Seijo and Z. Barandiaran, *J. Chem. Phys.*, 2005, **123**, 204502.
24 B. Swerts, C. L. F. Chibotaru, R. Lindh, L. Seijo, Z. Barandiaran, S. Clima, K. Pierloot and F. A. Hendrickx, *J. Chem. Theory Comput.*, 2008, **4**(4), 586–594.
25 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050–8053.
26 J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski, *J. Chem. Phys.*, 2005, **122**, 094115.
27 J. Neugebauer, Ch. R. Jacob, T. A. Wesolowski and E. J. Baerends, *J. Phys. Chem. A*, 2005, **109**, 7805–7814.
28 Ch. R. Jacob, J. Neugebauer, L. Jensen and L. Visscher, *Phys. Chem. Chem. Phys.*, 2006, **8**, 2349–2359.
29 J. Neugebauer and E. J. Baerends, *J. Phys. Chem. A*, 2006, **110**, 8786–8796.
30 J. M. García-Lastra, T. Wesolowski, M. T. Barriuso, J. A. Aramburu and M. Moreno, *J. Phys.: Condens. Matter*, 2006, **18**, 1519–1534.
31 M. Zbiri, C. A. Daul and T. A. Wesolowski, *J. Chem. Theory Comput.*, 2006, **2**(4), 1106–1111.
32 J. Neugebauer, *J. Chem. Phys.*, 2007, **126**, 134116.
33 J. Neugebauer, *J. Phys. Chem. B*, 2008, **112**, 2207–2217.
34 D. J. Tozer, *J. Chem. Phys.*, 2003, **119**, 12697–12699.
35 M. E. Casida, F. Gutierrez, J. Guan, F.-X. Gadea, D. Salahub and J.-P. Daudey, *J. Chem. Phys.*, 2000, **113**, 7062–7071.
36 A. Dreuw, J. L. Weisman and M. Head-Gordon, *J. Chem. Phys.*, 2003, **119**, 2943–2946.
37 J. Neugebauer, O. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2006, **124**, 214102.
38 O. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2004, **121**, 655–660.
39 M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim and M. Elstner, *J. Chem. Phys. B*, 2005, **109**, 3606–3615.
40 A. Dreuw, *ChemPhysChem*, 2006, **7**, 2259–2274.
41 D. Rappoport and F. Furche, *J. Am. Chem. Soc.*, 2004, **126**, 1277–1284.
42 A. Köhn and C. Hättig, *J. Am. Chem. Soc.*, 2004, **126**, 7399–7410.
43 C. Hättig, A. Hellweg and A. Köhn, *J. Am. Chem. Soc.*, 2006, **128**, 15672–15682.
44 M. E. Casida, *J. Chem. Phys.*, 2005, **122**, 054111.
45 N. T. Maitra, F. Zhang, R. J. Cave and K. Burke, *J. Chem. Phys.*, 2004, **120**, 5932–5937.
46 S. A. Cotton, *Annu. Rep. Prog. Chem., Sect. A*, 2006, **102**, 308.
47 M. D. Ward, *Annu. Rep. Prog. Chem., Sect. A*, 2006, **102**, 584.
48 N. Govind, Y. A. Wang, A. J. R. da Silva and E. A. Carter, *Chem. Phys. Lett.*, 1998, **295**, 129–134.
49 N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 1999, **110**, 7677–7688.
50 P. Huang and E. A. Carter, *J. Chem. Phys.*, 2006, **125**, 084102.
51 T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, *Phys. Rev. Lett.*, 2001, **86**, 5954–5957.
52 T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 2002, **116**, 42–54.
53 I. Infante, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2006, **125**, 074301.
54 T. A. Wesolowski, *Phys. Rev. Lett.*, 2002, **88**, 209701.
55 T. Klüner, N. Govind, Y. A. Wang and E. A. Carter, *Phys. Rev. Lett.*, 2002, **88**, 209702.
56 T. A. Wesolowski, *Phys. Rev. A*, 2008, **77**, 012504.
57 T. A. Wesolowski, *J. Chem. Phys.*, 1997, **106**, 8516–8526.
58 T. A. Wesolowski, P.-Y. Morgantini and J. Weber, *J. Chem. Phys.*, 2002, **116**, 6411–6421.
59 R. Kevorkyants, M. Dulak and T. A. Wesolowski, *J. Chem. Phys.*, 2006, **124**, 024104.

60 J. Neugebauer, M. J. Louwerse, P. Belanzoni, T. A. Wesolowski and E. J. Baerends, *J. Chem. Phys.*, 2005, **123**, 114101.
61 R. E. Bulo, Ch. R. Jacob and L. Visscher, *J. Phys. Chem. A*, 2008, **112**, 2640–2647.
62 P. Cortona, *Phys. Rev. B*, 1992, **46**, 2008–2014.
63 P. Cortona, A. Villafiorita Monteleone and P. Becker, *Int. J. Quantum Chem.*, 1995, **56**, 831–837.
64 P. Cortona and A. Villafiorita Monteleone, *J. Phys.: Condens. Matter*, 1996, **8**, 8983–8994.
65 W. N. Mei, L. L. Boyer, M. J. Mehl, M. M. Ossowski and H. T. Stokes, *Phys. Rev. B*, 2000, **61**, 11425–11431.
66 M. M. Ossowski, L. L. Boyer, M. J. Mehl and H. T. Stokes, *Phys. Rev. B*, 2002, **66**, 224302.
67 Ch. R. Jacob and L. Visscher, *J. Chem. Phys.*, 2008, **128**, 155102.
68 T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.*, 1996, **248**, 71–76.
69 M. Dulak and T. A. Wesolowski, *J. Chem. Theory Comput.*, 2006, **2**, 1538–1543.
70 M. E. Casida and T. A. Wesolowski, *Int. J. Quantum Chem.*, 2004, **96**, 577–588.
71 ADF, Amsterdam density functional program (development version), Theoretical Chemistry, Vrije Universiteit Amsterdam, URL: http://www.scm.com, 2008.
72 G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.
73 Ch. R. Jacob, J. Neugebauer and L. Visscher, *J. Comput. Chem.*, 2008, **29**, 1011–1018.
74 A. S. P. Gomes, F. Real, Ch. R. Jacob, V. Vallet and L. Visscher, 2008, in preparation.
75 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200–1211.
76 E. van Lenthe and E. J. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142–1156.
77 P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344–1352.
78 E. van Lenthe, J. G. Snijders and E. J. Baerends, *J. Chem. Phys.*, 1996, **105**, 6505–6516.
79 A. Lembarki and H. Chermette, *Phys. Rev. A*, 1994, **50**, 5328–5331.
80 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
81 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, **46**, 6671–6687.
82 DALTON, a molecular electronic structure program, Release 2.0 (2005), see http://www.kjemi.uio.no/software/dalton/dalton.html, 2005.
83 DIRAC, a relativistic *ab initio* electronic structure program, Release DIRAC04.0 written by H. J. Aa. Jensen, T. Saue, L. Visscher with contributions from V. Bakken, E. Eliav, T. Enevoldsen, T. Fleig, O. Fossgaard, T. Helgaker, J. Laerdahl, C. V. Larsen, P. Norman, J. Olsen, M. Pernpointner, J. K. Pedersen, K. Ruud, P. Salek, J. N. P. van Stralen, J. Thyssen, O. Visser and T. Winther, URL: http://dirac.chem.sdu.dk, 2004.
84 G. te Velde and E. J. Baerends, *J. Comput. Phys.*, 1992, **99**, 84–98.
85 H. Koch, P. Jorgensen and J. Olsen, *Chem. Phys. Lett.*, 1995, **244**, 75–82.
86 O. Christiansen, H. Koch, P. Halkier, Asger Jorgensen, T. Helgaker and A. S. de Méras, *J. Chem. Phys.*, 1996, **105**, 6921.
87 K. Hald, C. Hättig and P. Jorgensen, *J. Chem. Phys.*, 2000, **113**, 7765.
88 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
89 K. Aidas, J. Kongsted, A. Osted, K. Mikkelsen and O. Christiansen, *J. Phys. Chem. A*, 2005, **109**, 8001–8010.
90 M. Iliaš and T. Saue, *J. Chem. Phys.*, 2007, **126**, 064102.
91 Program AMFI, B. Schimmelpfennig, Stockholm University, Sweden.
92 B. A. Hess, C. M. Marian, U. Wahlgren and O. Gropen, *Chem. Phys. Lett.*, 1996, **251**, 365–371.
93 K. G. Dyall, *Theor. Chem. Acc.*, 2006, **90**, 491–500.
94 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2000, **113**, 9905–9910.
95 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 6862–6865.
96 L. Visscher, E. Eliav and U. Kaldor, *J. Chem. Phys.*, 2001, **115**(21), 9720–9726.
97 Ch. R. Jacob, R. E. Bulo and L. Visscher, 2008, in preparation.
98 R. G. Denning, J. O. W. Norris and D. Brown, *Mol. Phys.*, 1982, **46**, 287–323.
99 R. G. Denning, J. O. W. Norris and D. Brown, *Mol. Phys.*, 1982, **46**, 325–364.
100 S. Matsika and R. M. Pitzer, *J. Phys. Chem. A*, 2001, **105**, 637–645.
101 D. J. Watkin, R. G. Denning and K. Prout, *Acta Crystallogr., Sect. C*, 1991, **47**, 2517–2519.
102 F. Real, 2007, private communication.
103 A. Gelle and M.-B. Lepetit, *J. Chem. Phys.*, 2008, in press.
104 A. Gelle and M.-B. Lepetit, Fast calculation of the electrostatic potential in ionic crystals by direct summation method, arXiv:cond-mat.str-el/0711.2888, 2007.
105 M. P. Wilkerson, C. A. Arrington, J. M. Berg and B. L. Scott, *J. Alloys Compd.*, 2007, **444**, 634–639.
106 C. W. Porter and C. Iddings, *J. Am. Chem. Soc.*, 1926, **48**, 40–44.
107 N. S. Bayliss and E. G. McRae, *J. Phys. Chem.*, 1954, **58**, 1006–1011.
108 N. S. Bayliss and G. Wills-Johnson, *Spectrochim. Acta, Part A*, 1968, **24**, 551.
109 S. Fux, K. Kiewisch, Ch. R. Jacob, J. Neugebauer and M. Reiher, *Chem. Phys. Lett.*, submitted.

## D.2 Paper IX

# Software News and Updates
# PyADF — A Scripting Framework for Multiscale Quantum Chemistry

**CHRISTOPH R. JACOB,[1] S. MAYA BEYHAN,[2] ROSA E. BULO,[2] ANDRÉ SEVERO PEREIRA GOMES,[3] ANDREAS W. GÖTZ,[4] KARIN KIEWISCH,[2] JETZE SIKKEMA,[2] LUCAS VISSCHER[2]**

[1]*Center for Functional Nanostructures, Karlsruhe Institute of Technology (KIT), Wolfgang-Gaede-Str. 1a, 76131 Karlsruhe, Germany*

[2]*Amsterdam Center for Multiscale Modeling (ACMM), VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*

[3]*Laboratoire PhLAM, Université de Lille 1, CNRS UMR 8523, Bat P5, F-59655 Villeneuve d'Ascq Cedex, France*

[4]*San Diego Supercomputer Center, University of California San Diego, 9500 Gilman Drive MC0505, La Jolla, California 92093-0505*

*Received 18 January 2011; Revised 7 March 2011; Accepted 20 March 2011*
*DOI 10.1002/jcc.21810*
*Published online 3 May 2011 in Wiley Online Library (wileyonlinelibrary.com).*

**Abstract:** Applications of quantum chemistry have evolved from single or a few calculations to more complicated workflows, in which a series of interrelated computational tasks is performed. In particular multiscale simulations, which combine different levels of accuracy, typically require a large number of individual calculations that depend on each other. Consequently, there is a need to automate such workflows. For this purpose we have developed PyADF, a scripting framework for quantum chemistry. PyADF handles all steps necessary in a typical workflow in quantum chemistry and is easily extensible due to its object-oriented implementation in the Python programming language. We give an overview of the capabilities of PyADF and illustrate its usefulness in quantum-chemical multiscale simulations with a number of examples taken from recent applications.

 J Comput Chem 32: 2328–2338, 2011

**Key words:** multiscale; scripting; workflow; embedding

## Introduction

In modern applications of quantum-chemical program packages, one usually needs to perform series of calculations. For instance, for the same molecule a single calculation is in general not sufficient, but calculations using different theoretical methods, different basis sets, different technical settings, and in many cases also using different program packages are necessary. Furthermore, such series of calculations are mostly needed not only for a single molecular structure, but for many different ones.[1,2]

In addition, these calculations are commonly part of a collection of tasks that are interrelated, i.e., the results of one calculation serve as input for a subsequent one.[3,4] A basic example of such a workflow would be performing geometry optimizations for a series of molecules using density-functional theory (DFT) with a smaller basis set, followed by single-point energy calculations using a more accurate wave-function theory (WFT) method and/or a calculation of molecular properties. Finally, the appropriate results have to be extracted from the output files and have to be processed to obtain the sought-after quantities (e.g., total energy differences).

Clearly, there is a need to automate these workflows. Many computational chemists have developed their own solutions for this purpose, mostly shell scripts or small programs (see, e.g., ref. 5). However, these usually address only certain steps, such as generating input files or extracting results from output files. Some quantum-chemical program packages take a further step by providing a scripting interface. For instance, NWChem[6] offers a Python interface[7] for executing simple workflows, Molpro[8] provides scripting facilities in its input files and produces XML output for an easier post-processing of results, ADF[9,10] comes with tools for the automatic generation of input files and for easily extracting the

***Correspondence to:*** Ch. R. Jacob; e-mail: christoph.jacob@kit.edu

Contract/grant sponsor: Deutsche Forschungsgemeinschaft (DFG) (Center for Functional Nanostructures), The Netherlands Organization for Scientific Research (NWO) (Veni and Vici grant), Ministère chargé de la Recherche, Région Nord-Pas de Calais, Fond Européen de Développement Economique des Régions (FEDER), German Academic Exchange Service (DAAD) (Postdoc-Programme)

results of calculations,[11] and the MOLECONTROL[12] add-on to TURBOMOLE[13,14] allows for the execution of series of calculations and simple workflows.

However, as workflows become more and more involved, ad hoc shell-scripting solutions reach their limitations. Furthermore, the existing scripting interfaces are usually specific to one quantum-chemical program package and can often only handle parts of the required workflows. Therefore, more flexible and general solutions will be useful.

In the area of chemo- and bio-informatics, where very large datasets are generated and processed, general-purpose workflow engines that allow the user to organize and schedule different tasks (usually using a graphical user interface) are very common.[15,16] Such workflow engines have also been adapted to handle computational chemistry problems, in particular in the context of grid computing.[17–20] However, these solutions are mostly either tailored to executing a single program package (see, e.g., ref. 21), or very general and thereby present significant obstacles for their initial use and for the extension to tasks from quantum chemistry.

In the past years, work in our groups has focussed on "quantum-chemical multiscale simulations," which combine different levels of quantum-chemical descriptions. These are based on the frozen-density embedding (FDE) scheme initially proposed by Wesolowski and Warshel[22] (following earlier work by Senatore and Subbaswamy[23] and by Cortona[24]) and its extension to WFT-in-DFT embedding, first proposed by Carter and coworkers.[25–28]

Particularly for such quantum-chemical multiscale simulations one encounters very complex workflows that are beyond the capabilities of standard tools. Typically, they involve hundreds of individual calculations, and the results of a subset of these calculations are needed as input for following steps. For instance, applications of the FDE scheme to calculate solvent effects on molecular properties[29–31] require the construction of an approximate solvent electron density, followed by an FDE calculation of the molecular property of interest for the embedded solute molecule. This needs to be done for a large number of snapshots taken from a molecular dynamics simulation. Similarly, a subsystem-DFT treatment of proteins[32] requires calculations for all subsystems (e.g., the individual amino acids), each embedded in an environment constructed from all other subsystems. For WFT-in-DFT embedding calculations (see, e.g., ref. 33), it is necessary to combine different quantum-chemical program packages and to pass the embedding potential and/or the frozen environment electron density between these programs.

To automate these rather complicated workflows, we have developed PYADF, a scripting framework for quantum chemistry. It handles all the steps required in typical workflows of quantum chemistry: generation of input files, execution of the different programs, error handling, as well as extraction and post-processing of the results and offers a very flexible and extensible framework for combining these different steps. PYADF is written in the Python programming language[34] (for an introduction to Python from the perspective of computational sciences, see, e.g., refs. 35 and 36).

Even though the functionality available in PYADF currently focusses on quantum-chemical multiscale simulations, it is in no way limited to this type of calculations. Instead, it provides a general framework for processing workflows in quantum chemistry, which can easily be extended to additional computational tasks. Despite its name — which indicates its historical origin as an extension to the ADF package — PYADF is not specific to a single program, but works with a number of different quantum chemistry codes. In this article, we give an overview of the PYADF scripting framework, and illustrate its usefulness by discussing a number of examples.

This work is organized as follows. First, we outline the design of PYADF and give an overview of its most important features. This is followed by a demonstration of the capabilities of PYADF for automating workflows commonly encountered in quantum chemistry and an overview of various applications of PYADF in quantum-chemical multiscale simulations. Finally, we summarize and give an outlook on planned and ongoing developments.

## PYADF — Design and Overview

The driving idea behind PYADF is the definition of a framework that will provide mechanisms for both controlling the execution of different computational tasks and for managing the communication between these tasks. This should be achieved in such a way that users are provided simple, yet powerful ways to define their computational workflows.

To this end, PYADF makes use of object-oriented programming techniques in the high-level programming language Python. The central paradigm of object-oriented programming is the definition of classes, i.e., objects that contain both data (also know as "attributes") and actions (the so-called "methods"). The definition of different classes allows us to group together the different aspects involved in one step of the workflow into a single entity. This way, as many details as possible are hidden from the user, who only needs to know how to use these classes on a higher level. Such an object-oriented design has further advantages. It is possible to establish a hierarchical relation between classes (known as "inheritance"), so that one can utilize existing classes to construct new ones. The classes provided by PYADF can then easily be extended by the user. This makes it, for instance, rather straightforward to incorporate scientific codes from third parties.

The input files to PYADF are, in effect, simply Python scripts, so the full power of the language and its numerous extensions is immediately available to the user. To illustrate how we utilize the classes provided by PYADF to arrive at a very simple, high-level definition of a basic workflow, a sample input file is given in Figure 1.

This input defines an elementary workflow: The molecular coordinates are read from a file, a single point calculation is performed with ADF, and finally the magnitude of the dipole moment is extracted from the resulting output and printed.

In the first line, a molecule object (i.e., an instance of the `molecule` class) is created. In the simplest form, which is used here, it is initialized by reading the molecular coordinates from

```
mol = molecule('h2o.xyz')

job = adfsinglepointjob(mol, basis='TZVP')

results = job.run()

print results.get_dipole_magnitude()
```

**Figure 1.** A minimal PYADF input file.

an xyz-file. Internally, PyADF uses the OpenBabel library[37–39] for storing the molecular coordinates so that any file format supported by OpenBabel can be understood. The `molecule` class of PyADF has a number of methods for manipulating molecular coordinates (e.g., joining different molecules, splitting a large system into its molecular fragments, adding hydrogen atoms to protein structures, and aligning molecules). Most of these methods also rely on functionality provided by OpenBabel.

In the second line, a job object is initialized. PyADF provides a number of different job classes for different types of calculations. Here, the class `adfsinglepointjob`, which represents a single point DFT calculation with the ADF program, is used. There are other job classes for other types of calculations (e.g., `adfgeometryjob` for geometry optimizations, `adfnmrjob` for the calculation of NMR chemical shifts, and `adfcpljob` for calculating spin–spin coupling constants with ADF, or `daltonsinglepointjob` and `diracsinglepointjob` for single point calculations with the Dalton[40] and Dirac[41] programs, respectively). These job classes are part of a class hierarchy, i.e., they are related by inheritance. A complete list of the types of calculations currently available in PyADF can be found in the PyADF documentation (available at `http://www.pyadf.org`).

Creating a job object only defines the type of calculation and allows one to specify the technical settings (such as basis set and exchange–correlation functional) that should be used in this calculation. It does not execute the calculation itself. This is achieved by calling the job's `run` method, as shown in the third line of the script considered here. This method will generate the necessary input files, call the appropriate executable(s), and save the output file(s) produced by the program (both standard output and binary restart files, depending on the program).

The job's `run` method returns a results object. This results object is an instance of a results class corresponding to the type of the job. Using it, the results of the calculation can be accessed. For instance, on the fourth line of the considered input file, the magnitude of the dipole moment is extracted with the `get_dipole_magnitude` method. There is a hierarchy of such results classes that is analogous to the job class hierarchy. Each of them provides methods for accessing the different quantities that have been calculated. Which quantities are available obviously depends on the type of the calculation. A full list of predefined computational tasks and of the associated classes can be found in the PyADF documentation.

Internally, PyADF implements a file manager, which stores the files produced by each of the calculations. The results objects then use this file manager to access these files. If a certain quantity is requested, the corresponding method of the results object knows how to extract the quantity from these files — either by using a regular expression to extract it from the output file or by reading it from binary restart files. For the convenience of the user, the output of PyADF clearly identifies the related files for each job, so that these files can easily be inspected in case of a problem.

PyADF also provides extensive restart facilities. Because Python is an interpreted language, errors in the PyADF input file will be detected only at runtime. Furthermore, it is of course possible that one of the programs called by PyADF encounters an error condition during its execution. As a consequence, it can happen that a PyADF run is aborted after a large number of calculations have been performed. In this case, PyADF generates an archive of its results that can be imported when re-running with a corrected input file. In such a restarted run, the calculations that have already been completed earlier will not be executed again. This is achieved by generating a checksum of the input file(s) for each calculation. Before a calculation is actually executed, PyADF checks whether a calculation with the same checksum has already been performed. If so, a results object for this previous calculation is returned, using the files from the previous run.

## Automating Common Workflows in Quantum Chemistry with PyADF

### *A Simple Example: Calculation of NMR Shieldings*

To demonstrate how the building blocks of PyADF described in the previous section (i.e., the `molecule` class as well as job and result classes) can be used to construct quantum-chemical workflows, we first consider a simple example: a geometry optimization followed by a calculation of NMR shieldings. A flowchart of this workflow is shown in Figure 2a, which also indicates which information has to be exchanged between the different tasks.

First, the molecular coordinates are read from a file. Starting from this initial structure, a geometry optimization is then performed using a small basis set. This is followed by a single point calculation for the optimized geometry, employing a larger basis set. Subsequently, a calculation of the NMR shieldings is performed. This NMR calculation requires the results (most importantly, the MO coefficients) of the previous single point calculation. Finally, the calculated NMR shieldings are printed.

The input file in Figure 2b shows how such a workflow can be realized with PyADF. First, a molecule object is initialized by reading the molecular coordinates from an xyz-file and a list with the numbers of the nuclei for which the NMR shieldings should be calculated is created (lines 1–5). After that, the settings for the subsequent ADF calculations are initialized, specifying the exchange–correlation functional and the numerical integration accuracy (lines 7–10). Then, an `adfgeometryjob` instance is created and the job is run (lines 12–14). The `run` method returns a results object, from which the optimized molecular coordinates (in form of another molecule object) are obtained with the `get_molecule` method (line 17). This new molecule object is then used to initialize and run an `adfsinglepointjob`, for which a higher integration accuracy and a larger basis set are used (lines 19–22). The new results object is then used to initialize and run an `adfnmrjob` (lines 24–26). Finally, the `get_all_shieldings` method of the NMR job's results object is used to extract the calculated shieldings, which can then be printed (lines 28–31).

Since all the features of the Python language are available in PyADF input files, it is easy to extend the simple workflow outlined above. Among the many possibilities is the application of this workflow to a large number of molecules. This can be achieved by a loop over all xyz-files in a given directory, allowing the calculations to be performed for each of these molecules. If needed, the numbers of the nuclei for which the NMR shieldings have to be calculated could be determined for each of these molecules individually. To this end, one can, for instance, use the functionality provided by OpenBabel to identify the nuclei of interest using a SMARTS pattern.[42]

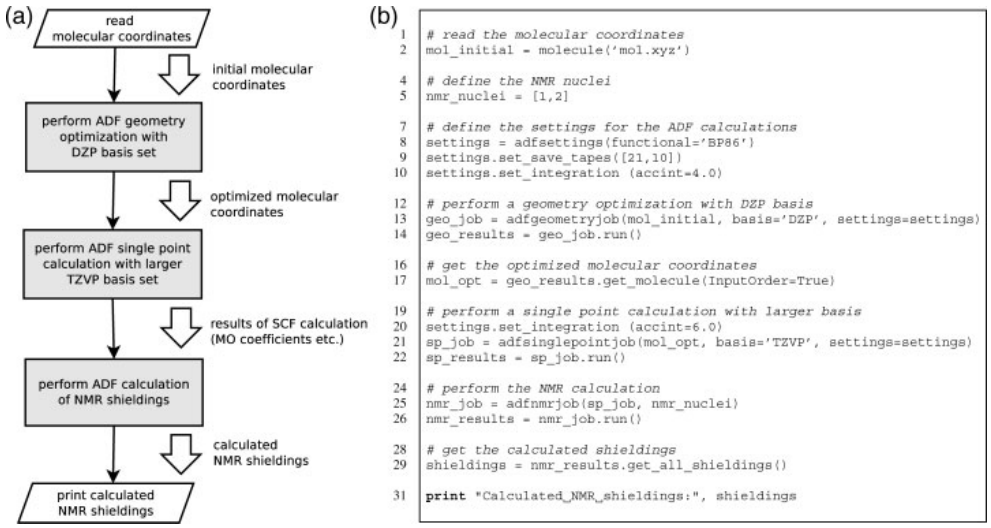



**Figure 2.** (a) Flowchart of a simple quantum-chemical workflow. The steps in gray boxes stand for the calculations performed with an external program. The large arrows indicate results that are passed between the different tasks. (b) A PyADF input file for realizing this workflow.

Furthermore, a calculation of the NMR shieldings for a reference compound could be added, so that the calculated shieldings can be converted to chemical shifts inside PyADF before they are printed.

### *Running Calculations for Large Test Sets of Molecules*

In recent years, work in our groups has focussed on quantum-chemical multiscale simulations, and in particular on the frozen-density embedding (FDE) scheme.[22] In this DFT-based scheme, the total electron density $\rho_{\mathrm{tot}}(\boldsymbol{r})$ is divided into the densities of $N$ subsystems $\rho^{(n)}(\boldsymbol{r})$ $(n = 1, \ldots, N)$, with

$$\rho_{\mathrm{tot}}(\boldsymbol{r}) = \sum_{n=1}^{N} \rho^{(n)}(\boldsymbol{r}). \tag{1}$$

Given an (approximate) density for all other subsystems, the density in one active subsystem can be determined from a set of KS-like equations, in which the effect of the frozen environment density enters through an effective embedding potential (see, e.g., refs. 22,43 and 44 for details). By iteratively updating each of the subsystem densities in so-called freeze-and-thaw cycles,[45] one obtains a subsystem-DFT scheme[24] that can be used as an efficient alternative to conventional Kohn–Sham (KS) DFT calculations.

However, even though the frozen-density embedding potential is in principle exact (in the sense that it should lead to the same total electron density as a KS-DFT calculation on the full system), additional approximations are required for the kinetic-energy component of the embedding potentials. Likewise, when calculating energies, approximations have to be introduced for the nonadditive kinetic energy, i.e., an approximate kinetic-energy density functional has to be used. This naturally raises the question how accurate these approximations are. To assess the quality of the available approximations (for an overview, see, e.g., ref. 46) one possibility is to compare the interaction energies between two fragments obtained from a subsystem-DFT calculation to those obtained from a supermolecular KS-DFT calculation.

Using this strategy, three of us have recently presented a comparison of various kinetic-energy density functionals within such subsystem-DFT calculations for a large test set of intermolecular complexes and transition metal coordination compounds.[47] The workflow applied in this study is illustrated in Figure 3. It consists of a loop over all systems in the test set. For each of these (usually bimolecular) systems, the coordinates of the two subsystems are read and calculations for the supermolecule as well as for the isolated fragments are performed. This is followed by a subsystem-DFT calculation (i.e., the densities of both subsystems are updated iteratively), which uses the densities of the isolated fragments as initial guess. Finally, the reference value for the interaction energy is calculated as $E_{\mathrm{int}}^{(\mathrm{ref})} = E_{\mathrm{supermol}} - E_1 - E_2$ and the subsystem-DFT interaction energy is calculated as $E_{\mathrm{int}}^{(\mathrm{FDE})} = E_{\mathrm{FDE}} - E_1 - E_2$ (with corrections for the basis set superposition error where appropriate). After the loop over the test set is complete, a statistical analysis of the errors in the interaction energies can be performed (e.g., calculation of the root-mean square deviation).

Such a workflow can be easily realized in PyADF: Using the functionality of the Python language, a loop over the test set can be performed. For each complex, single point calculations for both the supermolecule and the individual fragments can be executed with the help of the `adfsinglepointjob` class. Frozen-density embedding and subsystem-DFT jobs are handled by the class `adffragmentsjob`. Such jobs require a list of fragments (i.e., molecule objects and possibly the results objects of previous calculations) and for each of these fragments, it can then be chosen

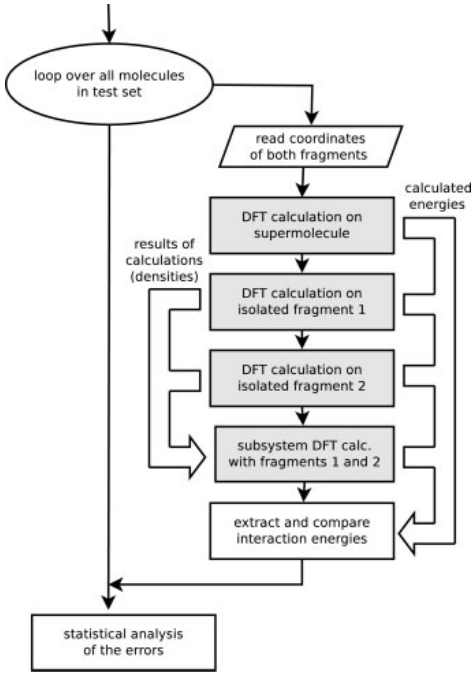



**Figure 3.** Workflow encountered for testing the accuracy of kinetic-energy functionals within subsystem DFT. The steps in gray boxes stand for the calculations performed with an external program. The large arrows indicate results that are passed between the different tasks.

whether it is active, frozen, or a frozen fragment that is iteratively updated in freeze-and-thaw cycles (see also the description of the flexible FDE implementation in the ADF package[44]). In the case considered here, two fragments (one active and one that is updated in freeze-and-thaw cycles) are used. Ultimately, the required energies can be extracted from the corresponding results objects, and their further processing can be done using the functionality of Python.

Since workflows similar to the one described here are quite common (another example would be the assessment of different exchange–correlation functionals for test sets of molecules), PYADF provides a convenience class `datasetjob` for such applications.

#### *Postprocessing of Results and Plotting*

Besides comparing energies, the quality of any quantum-chemical calculation can also be assessed by comparing the computed electron density with that of a reference calculation. PYADF provides the general functionalities for such a comparison. As an example, we will again consider subsystem-DFT calculations. With the exact kinetic-energy functional, a subsystem-DFT calculation would yield the same electron density as a supermolecular KS-DFT calculation on the full system.[43,45] Therefore, the difference between the density from the supermolecular KS-DFT calculation and the subsystem-DFT density obtained with an approximate kinetic-energy functional can be used to judge the quality of these approximations.

Performing the subsystem-DFT calculation for a test set of intermolecular complexes can be performed using PYADF as described in the previous section. However, in contrast to an assessment of the energy (a single real number), comparisons of the electron density (a function of three variables) require additional functionality for the postprocessing of the results. It is necessary to obtain the values of the different densities on a grid and compare them to each other, either by plotting the difference density or by quantifying the deviation by integrating the absolute value of the difference density.[48]

An excerpt from an input file showing how these postprocessing steps can be performed using PYADF is depicted in Figure 4. The results objects of both the supermolecular KS-DFT calculation and the subsystem-DFT calculation have a `get_density` method, which returns a "density object" (lines 6 and 9, respectively). This density can, for instance, be written in a cube file with the `get_cubfile` method (lines 7 and 10), which can in turn be used to prepare an isosurface plot. Example isosurface plots of the densities from a supermolecular KS-DFT calculation and from a subsystem-DFT calculation are shown in Figure 4 for the example of the coordination complex formed from $BH_3$ and $NH_3$. Despite the fact that this is a case for which the available GGA-type kinetic-energy functionals[50] have been shown to fail,[51,52] the two isosurface plots can hardly be distinguished. Therefore, it is more instructive to look at the difference density. In PYADF, the difference density can be obtained by simply subtracting the two density objects (line 12), resulting in a new density object that again can be written in a cube file (line 13).

The `get_density` method allows one to choose on which grid the density is needed. For plotting, one typically needs an evenly spaced grid, as it is selected on line 4. Such a grid is, however, not suitable for performing an accurate numerical integration. For this purpose, the more precise numerical integration grid employed by ADF can also be chosen (line 17), to calculate the difference density on this grid in a completely analogous way (lines 19–22). In this representation, it is then possible to calculate the numerical integral over the absolute value of the difference density accurately (line 24).

In a similar way, molecular orbitals, localized orbitals, as well as the Kohn–Sham potential and its individual components can be handled. For instance, the recent work on accurate frozen-density embedding potentials presented in ref. 52 used PYADF extensively for manipulating and plotting the different components of the potentials.

### Multiscale Simulations with PYADF

#### *Solvent Effects on Molecular Properties*

An important multiscale application of the FDE scheme is the calculation of solvent effects on molecular properties (e.g., electronic excitation energies,[29,30] ESR hyperfine coupling constants,[53] or NMR shieldings[31]). Such calculations employ the sequential molecular dynamics followed by quantum-mechanics calculations (S-MD/QM) strategy.[54] This strategy is illustrated for the calculation of the NMR shielding of acetonitrile in water in Figure 5, together with a simplifed version of the PYADF input file used for such calculations in ref. 31.

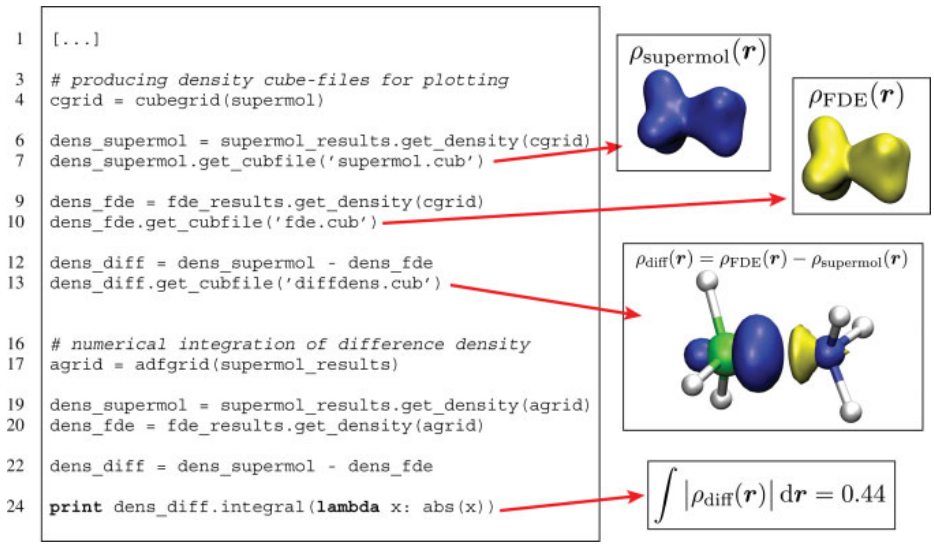



**Figure 4.** A PyADF input file for analyzing the densities from FDE calculations along with isosurface plots of the (difference) densities for $BH_3NH_3$. Isosurface plots have been prepared with VMD.[49]

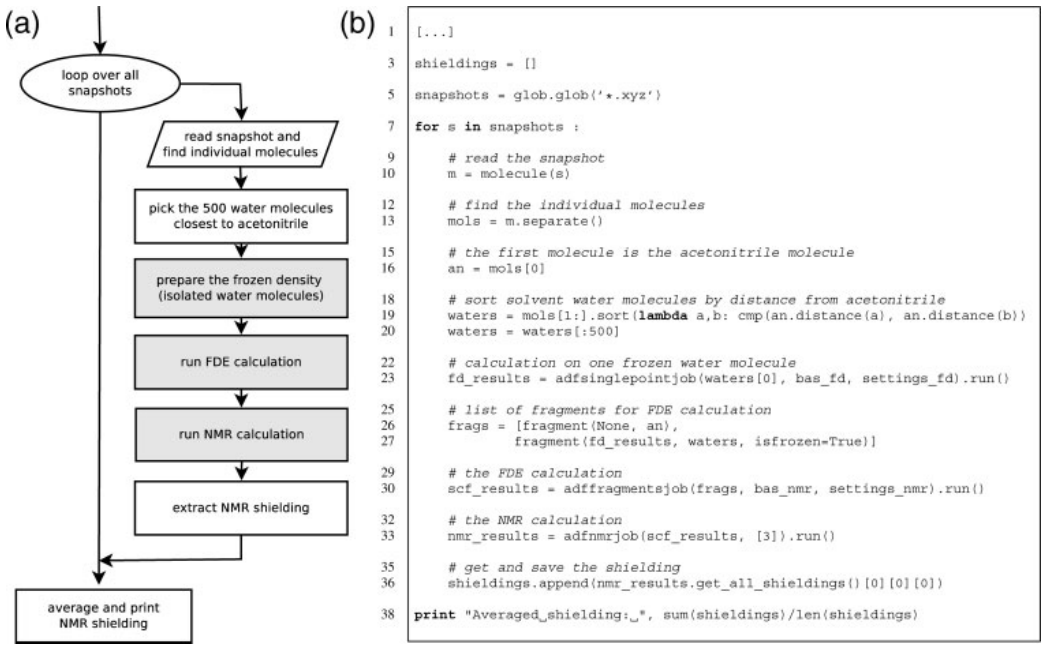



**Figure 5.** (a) Flowchart of the workflow used for calculating the solvent effect on the NMR shieldings using FDE. The steps in gray boxes stand for the calculations performed with an external program. (b) A PyADF input file for realizing this workflow.

First, using periodic boundary conditions, a classical molecular dynamics simulation of an acetonitrile molecule in water is performed. From this simulation, a sufficiently large number of snapshots are extracted. For each of these snapshots, a calculation of the NMR shielding is then performed for the solute molecule surrounded by the 500 nearest solvent molecules. Since a supermolecular calculation of molecular properties would not be feasible for such large systems, the FDE scheme is used in this step. The calculation is performed for the solute molecule as the active subsystem embedded in a frozen environment. For this solvent environment, an approximate electron density obtained by adding the densities calculated for isolated water molecules is used. Finally, the NMR shieldings calculated for each snapshot are averaged.

In the loop over all snapshots, all xyz-files from the current directory are used (line 5–7, see Fig. 5b). These xyz-files of the snapshots have to be generated using the molecular dynamics program. For each snapshot, the `separate` method of the `molecule` class is used to divide the snapshot into individual molecules, which are returned as a list `mols` (line 13). In this example, the first molecule in the list is the acetonitrile molecule (line 16), but if this is not the case it would be easy to add code for identifying it in this list. The remaining water molecules are then sorted by their distance to the acetonitrile molecule (line 19) and the 500 closest ones are chosen (line 20). In lines 25–27, the list of fragments for the FDE calculation is set up: The acetonitrile is the active fragment, and the list of the 500 water solvent molecules is used to create frozen fragments, each with the density of an isolated water molecule. In the present case, the classical molecular dynamics simulation used a water model with a fixed geometry. Therefore, a single frozen density calculated for one isolated water molecule (line 23) can be used for all 500 solvent molecules. If the geometries of the water molecules differ, it is possible to introduce a loop over all frozen solvent molecules, in which a specific frozen density is calculated for each of them. Similarly, it would also be possible to update the densities of some of the solvent molecules in freeze-and-thaw iterations by choosing the appropriate options for these fragments. In the remainder of the input, the FDE calculation and the NMR shielding calculation are performed, and after the loop over all snapshots is finished the calculated shieldings are averaged.

### *Subsystem-DFT for Proteins*

The currently available kinetic-energy functionals are not suitable to apply the FDE scheme to subsystems connected by covalent bonds.[51,52] However, this problem can be circumvented by using a more general partitioning, as it was initially proposed in the molecular fractionation with conjugate caps (MFCC) method.[55–57] This partitioning is shown in Figure 6a. Instead of partitioning a covalent bond directly, caps are introduced to terminate the dangling bonds between the fragments. The auxiliary molecule obtained by joining the caps is then subtracted again. Recently, two of us showed how such a partitioning can be applied within the FDE scheme and demonstrated that this 3-partition FDE (3-FDE) scheme presents an efficient method for a subsystem treatment of proteins.[32]

In ref. 32 the 3-FDE scheme was applied to the protein ubiquitin, containing 76 amino acids, which was chosen because it is small enough for a supermolecular calculation on the full protein to be possible for comparison. However, already for such a small protein a scripting framework such as PyADF is invaluable. First, PyADF's molecule class provides methods for partitioning the protein, introducing the caps, and generating the corresponding cap fragments. For ubiquitin, the protein is partitioned into its 76 amino acids, each terminated by caps, resulting in the creation of 75 cap fragments (see Fig. 6). This partitioning can be performed automatically by PyADF. Besides cutting the protein at the peptide bonds, as was required for ubiquitin, an extension that allows for partitioning by cutting through disulfide bridges[58] is also available. In addition, more general partitionings using larger fragments (e.g., around an active center) are being developed.[59]

Once a partitioning is established, calculations on each of the isolated fragments have to be performed to obtain an initial density. This corresponds to the MFCC scheme and is handled in PyADF by the class `adfmfccjob`. The initial guess is then improved by 3-FDE calculations on each of the subsystems. These calculations are repeated iteratively until the freeze-and-thaw cycles are converged. All these individual calculations are automatically executed by PyADF's `adf3fdejob`. Finally, PyADF is indispensable for post-processing of the final results, e.g., for adding (and subtracting) the densities of all the individual subsystems.

### *WFT-in-DFT Embedding*

The multiscale simulations discussed so far are solely based on DFT (DFT-in-DFT embedding). However, in many cases DFT with the currently available exchange–correlation functionals is not accurate enough. One prominent example is the failure of (adiabatic) time-dependent (TD) DFT for charge-transfer excitations.[60–64] Therefore, there is considerable interest in schemes for embedding a wave-function based *ab initio* description in an environment treated with DFT (WFT-in-DFT embedding). This allows one to systematically improve the accuracy by employing a hierarchy of accurate WFT treatments for the subsystem of interest.

The FDE scheme can be extended to WFT-in-DFT embedding by using the FDE embedding potential (which has been derived in a DFT context) as an additional local one-electron potential in the WFT calculation.[25–28] It can be shown that using such an embedding potentials is formally exact (i.e., it is exact if the exact exchange–correlation and kinetic-energy functionals are used and the limit of an exact WFT description is reached).[65,66]

In ref. 33, three of us proposed a simplified scheme for the calculation of local excitation energies with WFT-in-DFT embedding. In many cases in which TD-DFT fails (such as in the case of charge-transfer excitations), the density obtained for the ground-state is still rather accurate. Therefore, it is reasonable to determine the embedding potential in a DFT-in-DFT embedding calculation — possibly using freeze-and-thaw cycles for (parts of) the environment — and to import the resulting DFT-in-DFT embedding potential in the WFT calculations. Such a scheme was applied in ref. 33 to investigate local excitations of a neptunyl ion ($NpO_2^{2+}$) embedded as a defect in a $Cs_2UO_2Cl_4$ crystal.

The workflow corresponding to this simplified scheme is shown in Figure 7a. After an initial frozen density is determined (typically requiring several calculations on the fragments constituting the environment), a DFT-in-DFT embedding calculation is performed. The environment density (or part of it) is then updated iteratively in

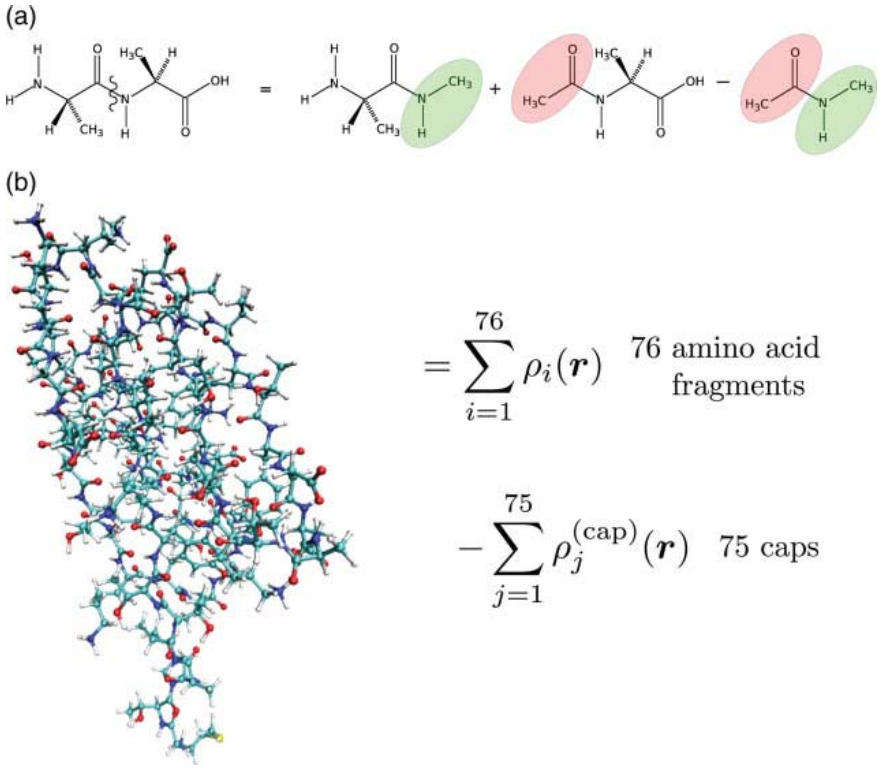



**Figure 6.** (a) Partitioning employed in subsystem-DFT calculations for proteins illustrated for dialanine. (b) For the calculations on ubiquitin in ref. 32, the protein is partitioned into its 76 amino acids and 75 cap fragments.

freeze-and-thaw cycles. After these freeze-and-thaw cycles are converged, the values of the embedding potential at the points of the numerical integration grid are read from ADF's binary results files and exported to a text file. This embedding potential as well as the coordinates and weights of the numerical integration grid itself are then imported in the WFT code. As soon as a WFT job is requested and is passed the results object of a DFT-in-DFT embedding calculation, PYADF internally takes care of creating the necessary files (input files, imported potential, etc.) and of passing them to the WFT code. Currently, DIRAC10 as well as locally modified versions of DALTON and NWCHEM support such an import of an embedding potential.

The drawback of the simplified scheme of ref. 33 is that the DFT ground-state density is used to determine the embedding potential, not the possibly more accurate ground-state density from the WFT calculation. In those cases where DFT fails to provide an acceptable ground-state density, a more complete WFT-in-DFT embedding scheme is needed. The workflow of such a scheme is presented in Figure 7b.

After a first DFT-in-DFT embedding calculation on the subsystem of interest is performed, the embedding potential is again exported and used in the WFT calculation. The values of the density, its gradient and Laplacian, and of the Coulomb potential generated by the density are then exported on the numerical integration grid. Such an export is currently supported by a development version of DIRAC for HF, DFT, and MP2 calculations. PYADF then takes care of converting the XML files written by DIRAC to the binary format that can be imported by ADF. Then, the environment density is updated in ADF, using the WFT density for the active subsystem and with this new environment density, a new embedding potential is exported. This is repeated iteratively, until these freeze-and-thaw iterations are converged. For such WFT-in-DFT calculations, PYADF provides the class `wftindftjob`, that implements all the steps in the workflow of Figure 7b.

## Conclusions and Outlook

PYADF is a scripting framework for quantum chemistry that can be used for automating quantum-chemical workflows. This is achieved by providing job classes, which can be used to set up and execute different types of quantum-chemical calculations, and results classes that are returned when running a job and which can be used to extract the results of the calculations. In addition, PYADF offers powerful

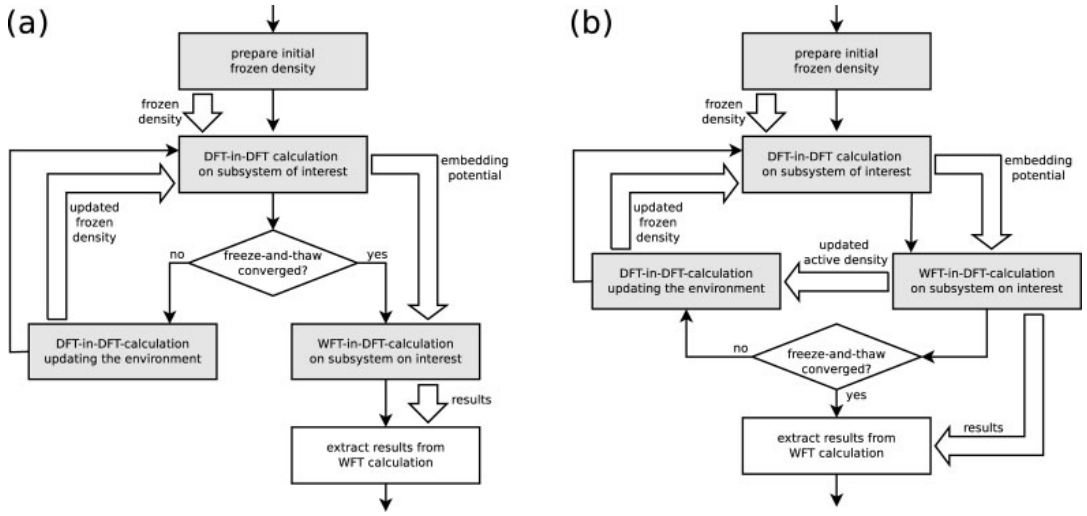



**Figure 7.** The workflow in WFT-in-DFT embedding calculations. The steps in gray boxes stand for the calculations performed with an external program. The large arrows indicate results that are passed between the different tasks. (a) The simplified scheme of ref. 33, in which the embedding potential from a DFT-in-DFT embedding calculation is imported in the WFT calculation. (b) A freeze-and-thaw WFT-in-DFT embedding setup, in which the embedding potential is updated iteratively, using the density from the WFT calculation.

features for reading and manipulating molecular coordinates and for the post-processing of the results (e.g., for handling orbitals, densities, and potentials available on a numerical grid). Since the full power of the Python language is available in PyADF input files, these building blocks can be combined into complicated workflows.

These features are particularly useful in multiscale simulations. We have illustrated this using examples of applications of PyADF from the recent research in our groups, such as the assessment of kinetic-energy functionals in the FDE scheme, the explicit treatment of solvent effects on molecular properties, subsystem-DFT calculations for proteins, and WFT-in-DFT embedding calculations. Other recent and ongoing work in our groups also relies on PyADF. For instance, the code for performing adaptive QM/MM molecular dynamics simulations developed by Bulo et al.[67,68] uses PyADF for executing the QM calculations.

PyADF is under active development. Currently, it supports ADF, DALTON, DIRAC, and NWCHEM calculations, even though the types of calculations available for each program package differ. We are working on extending these existing interfaces to support additional types of calculations. In this context, we also plan to add a unified interface to similar types of calculations performed with different program packages by providing an additional layer of job classes (e.g., `singlepointjob`, which will then execute either `adfsinglepointjob`, `daltonsinglepointjob`, or `diracsinglepointjob`). Furthermore, interfaces to additional programs, most importantly TURBOMOLE (using the MOLECONTROL environment), are being developed.

Another important topic we are considering is parallelization. Currently, PyADF executes all external tasks sequentially, even though each task can use multiple processors if the program packages support parallelization. A parallel version of PyADF to allow for coarse-grain parallelization of potentially concurrent tasks in a workflow is currently being developed. To achieve this, external tasks will not be executed immediately when a job's `run` method is called, but instead be gathered into a "job queue." Only when the results of one job are requested, the accumulated tasks will be executed in parallel. This way, an "automatic parallelization" can be achieved, in which PyADF input files are still written in a sequential way, and the problem of determining the dependencies between the different jobs and of executing the external tasks in parallel is handled behind the scenes by the scripting framework.

PyADF version 1.0 is available free of charge at `http://www.pyadf.org` under the GNU General Public License (GPL). On the website, one can also find an extensive documentation and examples of input files, including those for the examples discussed here.

## Acknowledgments

The authors thank Drew A. McCormack (Amsterdam) for many inspiring discussions.

## References

1. Cramer, Ch. J. Essentials of Computational Chemistry, Wiley: New York, 2002.

2. Heine, T.; Joswig, J.-O.; Gelessus, A. Computational Chemistry Workbook: Learning Through Examples; Wiley-VCH: Weinheim, 2009.
3. Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. J Chem Phys 2004, 120, 4129.
4. Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. J Chem Phys 2007, 126, 084108.
5. van Zeist, W.-J.; Fonseca Guerra, C.; Bickelhaupt, F. M. J Comput Chem 2008, 29, 312.
6. Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. Comput Phys Commun 2010, 181, 1477.
7. Environmental Molecular Sciences Laboratory (EMSL) at Pacific Northwest National Laboratory (PNNL), NWCHEM 6.0 Python interface, 2010. Available at http://www.nwchem-sw.org/index.php/Python, Accessed on 6 March 2010.
8. Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Schütz, M.; Celani, P.; Knizia, G.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklaß, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. MOLPRO, version 2010.1, a package of ab initio programs. Available at: http://www.molpro.net, 2010.
9. Baerends, E. J.; Ziegler, T.; Autschbach, J.; Bashford, D.; Bérces, A.; Bickelhaupt, F. M.; Bo, C.; Boerrigter, P. M.; Cavallo, L.; Chong, D. P.; Deng, L.; Dickson, R. M.; Ellis, D. E.; van Faassen, M.; Fan, L.; Fischer, T. H.; Fonseca Guerra, C.; Ghysels, A.; Giammona, A.; van Gisbergen, S. J. A.; Götz, A. W.; Groeneveld, J. A.; Gritsenko, O. V.; Grüning, M.; Gusarov, S.; Harris, F. E.; van den Hoek, P.; Jacob, C. R.; Jacobsen, H.; Jensen, L.; Kaminski, J. W.; van Kessel, G.; Kootstra, F.; Kovalenko, A.; Krykunov, M. V.; van Lenthe, E.; McCormack, D. A.; Michalak, A.; Mitoraj, M.; Neugebauer, J.; Nicu, V. P.; Noodleman, L.; Osinga, V. P.; Patchkovskii, S.; Philipsen, P. H. T.; Post, D.; Pye, C. C.; Ravenek, W.; Rodríguez, J. I.; Ros, P.; Schipper, P. R. T.; Schreckenbach, G.; Seldenthuis, J. S.; Seth, M.; Snijders, J. G.; Solà, M.; Swart, M.; Swerhone, D.; te Velde, G.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T. A.; van Wezenbeek, E. M.; Wiesenekker, G.; Wolff, S. K.; Woo, T. K.; Yakovlev, A. L. ADF, Amsterdam density functional program, 2010. Available at: http://www.scm.com, Accessed on 6 March 2010.
10. te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. J Comput Chem 2001, 22, 931.
11. Scientific Computing and Modelling, ADFPREP and ADFREPORT, 2008. Available at: http://www.scm.com/Doc/ Doc2010.01/ ADF/Utilities, Accessed on 6 March 2010.
12. Doll, C.; Schäfer, A.; Sittel, F., MOLECONTROL, Available at: www.turbomole.com, Accessed on 6 March 2010, 2010.
13. Ahlrichs, R.; et al. TURBOMOLE. Available at: http://www.turbomole.com, Accessed on 6 March 2010.
14. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, Ch. Chem Phys Lett 1989, 162, 165.
15. Kuhn, T.; Willighagen, E.; Zielesny, A.; Steinbeck, Ch. BMC Bioinf 2010, 11, 159.
16. Tiwari, A.; Sekhar, A. K. T. Comput Biol Chem 2007, 31, 305.
17. Gomes, A.; Merzky, A.; Visscher, L. In Proceedings of the Computational Science – ICCS 2006, Pt. 3, Lecture Notes in Computer Science; Springer: Berlin, 2006; pp. 97–104.
18. Sudholt, W.; Altintas, I.; Baldridge, K. In Proceedings of the Computational Science – ICCS 2006, Pt. 3, Alexandrov, V. N.; VanAlbada, G. D.; Sloot, P. M. A.; Dongarra, J., Eds.; Lecture Notes in Computer Science; Springer-Verlag: Berlin, 2006; pp. 69–76.
19. Koehler, M.; Ruckenbauer, M.; Janciak, I.; Benkner, S.; Lischka, H.; Gansterer, W. N. In Proceedings of the Computational Science and its Applicatiosn - ICCSA 2010, pt. 4, Taniar, D.; Gervasi, O.; Murgante, B.; Pardede, E.; Apduhan, B. O., Eds.; Lecture Notes in Computer Science; Springer-Verlag: Berlin, 2010; pp. 13–28.
20. Wang, J.; Korambath, P.; Kim, S.; Johnson, S.; Jin, K.; Crawl, D.; Altintas, I.; Smallen, S.; Labate, B.; Houk, K. N. In Proceedings of the ICCS 2010 - International Confercence on Computational Science, Procedia Computer Science; Elsevier: Amsterdam, 2010; pp. 1169–1178.
21. Sanna, N.; Castrignano, T.; De Meo, P. D.; Carrabino, D.; Grandi, A.; Morelli, G.; Caruso, P.; Barone, V. Theor Chem Acc 2007, 117, 1145.
22. Wesolowski, T. A.; Warshel, A. J Phys Chem 1993, 97, 8050.
23. Senatore, G.; Subbaswamy, K. R. Phys Rev B 1986, 34, 5754.
24. Cortona, P. Phys Rev B 1992, 46, 2008.
25. Govind, N.; Wang, Y. A.; da Silva, A. J. R.; Carter, E. A. Chem Phys Lett 1998, 295, 129.
26. Govind, N.; Wang, Y. A.; Carter, E. A. J Chem Phys 1999, 110, 7677.
27. Klüner, T.; Govind, N.; Wang, Y. A.; Carter, E. A. Phys Rev Lett 2001, 86, 5954.
28. Klüner, T.; Govind, N.; Wang, Y. A.; Carter, E. A. J Chem Phys 2002, 116, 42.
29. Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesolowski, T. A. J Chem Phys 2005, 122, 094115.
30. Neugebauer, J.; Jacob, Ch. R.; Wesolowski, T. A.; Baerends, E. J. J Phys Chem A 2005, 109, 7805.
31. Bulo, R. E.; Jacob, Ch. R.; Visscher, L. J Phys Chem A 2008, 112, 2640.
32. Jacob, Ch. R.; Visscher, L. J Chem Phys 2008, 128, 155102.
33. Gomes, A. S. P.; Jacob, Ch. R.; Visscher, L. Phys Chem Chem Phys 2008, 10, 5353.
34. van Rossum, G.; et al., PYTHON. Available at: http://www.python.org, Accessed on 6 March 2010.
35. McCormack, D. A. Scientific Scripting with Python, 1st ed., lulu.com, Raleigh, NC, USA, 2009. Accessed on 6 March 2010.
36. Langtangen, H. P. Python Scripting for Computational Science, 3rd ed.; Springer, Berlin, 2008.
37. The Open Babel package. Available at: http://openbabel.sourceforge.net/ Accessed on 6 March 2010.
38. Guha, R.; Howard, M. T.; Hutchison, G. R.; Murray-Rust, P.; Rzepa, H.; Steinbeck, Ch.; Wegner, J.; Willighagen, E. L. J Chem Inf Model 2006, 46, 991.
39. O'Boyle, N.; Morley, C.; Hutchison, G. Chem Cent J 2008, 2, 5.
40. DALTON, a molecular electronic structure program, Release 2.0. Available at: http://www.kjemi.uio.no/software/dalton/dalton.html, Accessed on 6 March 2010, 2005.
41. Saue, T.; Visscher, L.; Jensen, H. J. Aa. DIRAC, a relativistic ab initio electronic structure program, Release DIRAC10. Available at: http://dirac.chem.vu.nl, Accessed on 6 March 2010, 2010.
42. SMARTS — A Language for Describing Molecular Patterns. Available at: http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html, Accessed on 6 March 2010.
43. Wesolowski, T. A. In Computational Chemistry: Reviews of Current Trends, vol. 10; Leszczynski, J., Ed.; World Scientific: Singapore, 2006; pp. 1–82.
44. Jacob, Ch. R.; Neugebauer, J.; Visscher, L. J Comput Chem 2008, 29, 1011.
45. Wesolowski, T. A.; Weber, J. Chem Phys Lett 1996, 248, 71.
46. Wang, Y. A.; Carter, E. A. In Theoretical Methods in Condensed Phase Chemistry, Schwartz, S. D., Ed.; Kluwer: Dordrecht, 2000; pp. 117–184.
47. Götz, A. W.; Beyhan, S. M.; Visscher, L. J Chem Theory Comput 2009, 5, 3161.
48. Beyhan, S. M.; Götz, A. W.; Jacob, Ch. R.; Visscher, L. J Chem Phys 2010, 132, 044114.

49. Humphrey, W.; Dalke, A.; Schulten, K. J Mol Graphics 1996, 14, 33.
50. Lembarki, A.; Chermette, H. Phys Rev A 1994, 50, 5328.
51. Fux, S.; Kiewisch, K.; Jacob, Ch. R.; Neugebauer, J.; Reiher, M. Chem Phys Lett 2008, 461, 353.
52. Fux, S.; Jacob, Ch. R.; Neugebauer, J.; Visscher, L.; Reiher, M. J Chem Phys 2010, 132, 164101.
53. Neugebauer, J.; Louwerse, M. J.; Belanzoni, P.; Wesolowski, T. A.; Baerends, E. J. J Chem Phys 2005, 123, 114101.
54. Malaspina, T.; Coutinho, K.; Canuto, S. J Chem Phys 2002, 117, 1692.
55. Zhang, D. W.; Zhang, J. Z. H. J Chem Phys 2003, 119, 3599.
56. Gao, A. M.; Zhang, D. W.; Zhang, J. Z. H.; Zhang, Y. Chem Phys Lett 2004, 394, 293.
57. Mei, Y.; Zhang, D. W.; Zhang, J. Z. H. J Phys Chem A 2005, 109, 2.
58. Chen, X. H.; Zhang, D. W.; Zhang, J. Z. H. J Chem Phys 2004, 120, 839.
59. Kiewisch, K.; Jacob, Ch. R.; Visscher, L. to be published, 2011.
60. Tozer, D. J. J Chem Phys 2003, 119, 12697.
61. Dreuw, A.; Weisman, J. L.; Head-Gordon, M. J Chem Phys 2003, 119, 2943.
62. Gritsenko, O.; Baerends, E. J. J Chem Phys 2004, 121, 655.
63. Neugebauer, J.; Gritsenko, O.; Baerends, E. J. J Chem Phys 2006, 124, 214102.
64. Hieringer, W.; Görling, A. Chem Phys Lett 2006, 419, 557.
65. Wesolowski, T. A. Phys Rev A 2008, 77, 012504.
66. Khait, Y. G.; Hoffmann, M. R. J Chem Phys 2010, 133, 044107.
67. Bulo, R. E.; Ensing, B.; Sikkema, J.; Visscher, L. J Chem Theory Comput 2009, 5, 2212.
68. Nielsen, S. O.; Bulo, R. E.; Moore, P. B.; Ensing, B. Phys Chem Chem Phys 2010, 12, 12401.

## D.3 Paper X

THE JOURNAL OF CHEMICAL PHYSICS **136**, 044104 (2012)

# Molecular properties via a subsystem density functional theory formulation: A common framework for electronic embedding

Sebastian Höfener,[1,a)] André Severo Pereira Gomes,[2,b)] and Lucas Visscher[1,c)]
[1]*Amsterdam Center for Multiscale Modeling (ACMM), VU University Amsterdam, Theoretical Chemistry, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*
[2]*Laboratoire PhLAM, CNRS UMR 8523, Université de Lille 1, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France*

(Received 31 August 2011; accepted 19 December 2011; published online 24 January 2012)

In this article, we present a consistent derivation of a density functional theory (DFT) based embedding method which encompasses wave-function theory-in-DFT (WFT-in-DFT) and the DFT-based subsystem formulation of response theory (DFT-in-DFT) by Neugebauer [J. Neugebauer, J. Chem. Phys. **131**, 084104 (2009)] as special cases. This formulation, which is based on the time-averaged quasi-energy formalism, makes use of the variation Lagrangian techniques to allow the use of non-variational (in particular: coupled cluster) wave-function-based methods. We show how, in the time-independent limit, we naturally obtain expressions for the ground-state DFT-in-DFT and WFT-in-DFT embedding via a local potential. We furthermore provide working equations for the special case in which coupled cluster theory is used to obtain the density and excitation energies of the active subsystem. A sample application is given to demonstrate the method. © *2012 American Institute of Physics*. [doi:10.1063/1.3675845]

## I. INTRODUCTION

Non-empirical electronic structure methods, and their implementations in sophisticated computer programs, have become viable tools to study the molecular basis of natural phenomena. By carrying out calculations on quantum chemical models of varying size and complexity, one may, e.g., investigate in detail how interactions between the constituents of a biochemical system determine its function. An attractive feature of this modelling is that one is not restricted to reproduce experimental observations but may also carry out numerical experiments to make predictions about the effect of modifications in a system. These experiments may then be used to aid in tuning the behaviour of artificial or biochemically modified natural systems.

In order to be useful, a given method should be able to provide reliable numerical data with a reasonable computational effort. Methods that have proven to work well in applications on small, isolated, molecules may be difficult to scale up for models of condensed phase systems. This is due to the steep computational scaling with the number of atoms in the system that most methods exhibit. Methods are typically based on density functional theory (DFT) or on post-Hartree-Fock (HF) wave-function (WFT) approaches that have cubic or worse scaling of computational costs with system size. For conventional algorithms, this leads to a limitation in system size of ≃tens of atoms (or ≃hundreds of atoms in the case of HF or DFT) that is only slowly increased by advances in computer technology.

One way to push the limit of applicability of these methods forward is to utilize techniques in which long-range interactions are treated in a simplified and, therefore, more efficient manner. This is facilitated by the density fitting or resolution-of-identity approach and allows for accurate calculations of medium-sized molecules by coupled cluster (CC) techniques.[1–3] While such linear scaling implementations are essential for benchmark and highly accurate studies, they are still too demanding for standard applications. Another complication of such global descriptions is the interpretation of results in terms of qualitative models. This typically requires an additional analysis step in which the wave-function and molecular properties are decomposed into local contributions.

An alternative is to employ a subsystem approach, in which the total system is *a priori* divided into small, chemically meaningful, units that are considered separately. One may thereby easily approximate less important parts of the system by a computationally efficient approximate method such as molecular mechanics (MM). The most popular realization of such a scheme is the two-level QM/MM method,[4–7] but more general methods in which an arbitrary number of computational methods are combined are also in use.[8,9] The flexibility to combine the most suitable methods (including, e.g., specialized implementations) for the different tasks has lead to a multitude of implementations of multilevel approaches. One may thereby distinguish between the so-called embedding approaches, in which the accurate description is intended only for one part of the system, and the true subsystem approaches that build a global property of the system from local properties. In the embedding schemes, one may furthermore distinguish between methods that treat the environment as an unstructured continuum and methods that allow for atomistic detail and include specific interactions with environment. Techniques to calculate molecular

a)Electronic mail: s.hoefener@vu.nl.
b)Electronic mail: andre.gomes@univ-lille1.fr.
c)Author to whom correspondence should be addressed. Electronic mail: l.visscher@vu.nl.

properties by DFT, HF, and CC methods have been successfully combined with both specific (polarizable force-field) and non-specific (dielectric continua) models of the environment by Christiansen, Mikkelsen, Kongsted, and co-workers.[10–14]

While these approaches are very attractive in terms of computational efficiency, they do rely on the chosen parameterization of the environment—which is a drawback, if there is only a limited amount of experimental data available to parameterize the force fields or the continuum description, or when simple parametrizations are difficult due to the nature of the interactions. An alternative is then to resort to ab initio methods in which also the environment is modelled as consisting of a collection of interacting units that are each calculated using an appropriate quantum-mechanical method. Such discrete quantum-mechanical (QM/QM) methods are, however, considerably more expensive than QM/MM approaches and require efficient approximations in the less interesting "environmental" region of the system. A promising method is the so-called frozen-density embedding (FDE) scheme by Wesolowski and Warshel,[15] following an approach originally proposed by Senatore and Subbaswamy[16] and later Cortona[17] for solid-state calculations. In FDE all subsystems and their interactions are described by DFT, with computational savings resulting from the fact that typically only one system of interest is fully optimized. The other subsystems are described using a suitably chosen *frozen* electron density. The method is formally exact if a number of boundary conditions on the initial subsystem densities are fulfilled.[18,19] In practice, the quality of results depends on the employed non-additive parts of the kinetic and exchange-correlation energy functionals and derivatives thereof[20–22] to describe the interaction between the chosen subsystems. With the currently available functionals, one may describe primarily electrostatic and hydrogen bonded interactions rather well,[23–25] whereas coordination or covalent bond still present a major problem.[26,27] While the FDE ansatz has been mostly applied in the embedding regime (one small active system surrounded by a large frozen environment), one may also formulate this model as a special case of a more general subsystem DFT approach.[28,29] One then writes the total density as a sum of subsystem densities

$$\rho_{\text{tot}} = \sum_i \rho_i, \tag{1}$$

that are each optimized separately with the density of the other fragments fixed.

The formulation of response theory within the FDE framework was first proposed by Casida and Wesolowski,[30] but only reached its full potential when Neugebauer[31,32] extended the formalism to a general subsystem DFT response approach and provided an efficient implementation in the ADF program package.[33] His formulation does not only recover important environment contributions[34] on polarizabilities and excitation energies in dimers, but also allows for the coupling of local excitations in a complete model.[35] A growing number of applications shows the promise of subsystem DFT in both the (frozen-density) embedding mode[36–40] as well as in the (fully self-consistent) subsystem mode[41] to describe molecular properties.

Notwithstanding the success of the applications mentioned, DFT-in-DFT embedding approaches will always be constrained by the limitations of the DFT itself. One may encounter cases in which present-day functionals fail to provide a quantitatively correct description of one or more of the subsystems. In such cases, we would like to employ WFT approaches, and progress through one of its well-defined hierarchy of methods[42] to improve and check the reliability of the calculated results.

A very useful feature of the FDE is the fact that the embedding potential that is used to obtain the density of the embedded system is local. This absence of nonlocal projection operators facilitates the integration of DFT- and WFT-based methods in one overall model. In order to include WFT in FDE, we need to consider a subsystem $j$ for which the energy is obtained by optimizing the parameters of a many-electron wave function $\Psi_j$. This system should then interact with the other subsystems only via its density

$$\Psi_j^* \Psi_j \rightarrow \rho_j, \tag{2}$$

as in the DFT-in-DFT case. Such a WFT-in-DFT embedding scheme has been pursued by Carter and co-workers,[43–47] who combined DFT and variational methods such as Hartree-Fock, complete active space self-consistent field (CASSCF), or (multireference) CI. Their approach has so far mainly been used to describe localized excitations in solids or surfaces, e.g., for the calculation of excitation energies of CO adsorbed on a platinum surface.[48,49] Some of us[50] implemented furthermore an approximate scheme to employ non-variational WFT methods (based on coupled cluster theory) for the calculation of the low-lying spectra of solvated acetone and the *f–f* spectra of the Neptunyl ion embedded in a $Cs_2UO_2Cl_4$ crystal.

None of these approaches has so far gone beyond the use of an embedding potential constructed for the ground state in the determination of the excited state energies or response properties such as polarizabilities. For such applications, one needs to consider the change in interaction energy caused by changes in the active system as well as by responses of the environment. A straightforward way of taking those changes into account is to choose a state-specific determination of the embedding potential, as recently proposed by Khait and co-workers.[51] This has a drawback, however, that multiple calculations are required if one is interested in more than one excited states. Problematic is also the inclusion of non-variational methods such as (multireference) coupled cluster in which the wave function and the corresponding density are not explicitly calculated. For such methods, it is convenient to formulate the environment contribution to molecular properties and electronic excitations in terms of response theory, as this provides a natural connection to the techniques used in non-variational WFT methods.

Our goal in this paper is to work out a novel and rigorous FDE response theory framework with which it is possible to calculate molecular properties within a general subsystem formulation—capable of handling both DFT-in-DFT and WFT-in-DFT embedding. We will make use of the time-averaged quasi-energy formalism[52,53] which provides a natural way to treat variational and non-variational electronic structure methods in the same fashion. After providing the

necessary background on the FDE energy expressions, we will start by discussing response theory in a subsystem formulation and show how this reduces to the standard formulation in the case of non-interacting subsystems. Next, we will recast the DFT-in-DFT formalism of Neugebauer[31,32] into the time-averaged quasi-energy formalism and define key quantities for the WFT-in-DFT approach. This case is first considered for Hartree-Fock, and then for the case of non-variational coupled cluster methods. We will briefly discuss possible approximations, related to the extent one wishes to consider the coupling of the different subsystems in the time-dependent treatment. With the working equations available, we finish by addressing the similarities and differences between the formalism discussed here and the QM/MM response theory schemes proposed by Christiansen, Mikkelsen, Kongsted, and co-workers.[10–14]

## II. SUBSYSTEM DFT

For the optimization of a particular density $\rho_{\rm I}$, it is convenient to sum the other densities to a frozen environment density $\rho_{\rm II}$ and rewrite the density partitioning of Eq. (1) as

$$\rho_{\rm tot} = \rho_{\rm I} + \sum_{i\neq {\rm I}} \rho_i = \rho_{\rm I} + \rho_{\rm II}. \tag{3}$$

The total energy of the system, $E_{\rm tot}[\rho_{\rm tot}]$, can then be written as

$$E_{\rm tot}[\rho_{\rm I} + \rho_{\rm II}] = E_{\rm I}[\rho_{\rm I}] + E_{\rm II}[\rho_{\rm II}] + E_{\rm int}[\rho_{\rm I}, \rho_{\rm II}], \tag{4}$$

with the internal energy of each of the subsystems $i$ given as

$$E_i[\rho_i] = \int \rho_i(\mathbf{r}) v^i_{\rm nuc}(\mathbf{r}) {\rm d}\mathbf{r} + \frac{1}{2}\iint \frac{\rho_i(\mathbf{r})\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} {\rm d}\mathbf{r}\,{\rm d}\mathbf{r}' + E_{\rm xc}[\rho_i] + T_s[\rho_i] + E^i_{\rm nuc}, \tag{5}$$

with $v^i_{\rm nuc}$ the nuclear potential due to the set of atoms associated with subsystem $i$ and $E^i_{\rm nuc}$ the nuclear repulsion energy. The interaction energy is similarly given by the expression

$$E_{\rm int}[\rho_{\rm I}, \rho_{\rm II}] = \int \rho_{\rm I}(\mathbf{r}) v^{\rm II}_{\rm nuc}(\mathbf{r}) {\rm d}\mathbf{r} + \int \rho_{\rm II}(\mathbf{r}) v^{\rm I}_{\rm nuc}(\mathbf{r}) {\rm d}\mathbf{r} + E^{\rm I,II}_{\rm nuc} + \iint \frac{\rho_{\rm I}(\mathbf{r})\rho_{\rm II}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} {\rm d}\mathbf{r}\,{\rm d}\mathbf{r}' + E^{\rm nadd}_{\rm xc}[\rho_{\rm I}, \rho_{\rm II}] + T^{\rm nadd}_{\rm s}[\rho_{\rm I}, \rho_{\rm II}], \tag{6}$$

where non-additive contributions are defined as (see, e.g., Ref. 29)

$$X^{\rm nadd}[\rho_{\rm I}, \rho_{\rm II}] = X[\rho_{\rm I} + \rho_{\rm II}] - X[\rho_{\rm I}] - X[\rho_{\rm II}]. \tag{7}$$

All interaction energies are defined solely in terms of the subsystem densities that are either determined by a Kohn-Sham (KS) approach or by optimization of the wave function for an interacting system (WFT approach).[50,54] We note that orbitals of different subsystems always belong to independent subsets that are therefore, in general, non-orthogonal.

Equation (4) is the starting point for the response formulation in which we will first consider different parameterizations of the subsystem densities. We note that the internal energy of the environment does not depend on the active density $\rho_{\rm I}$ so that minimizing the *total* energy of the system with respect to $\rho_{\rm I}$ yields the Euler-Lagrange equation

$$\frac{\delta E_{\rm I}[\rho_{\rm I}]}{\delta\rho_{\rm I}} + \frac{\delta E_{\rm int}[\rho_{\rm I}, \rho_{\rm II}]}{\delta\rho_{\rm I}} = \mu, \tag{8}$$

with the Lagrange multiplier $\mu$ introduced to keep the number of electrons in system I constant. While this constraint can be avoided in the context of DFT-in-DFT embedding as shown recently by Elliot *et al.*,[55,56] in WFT-in-DFT embedding it can only be relaxed in the DFT subsystems,[57] because wave-function-based methods can only provide accurate densities for systems with an integer number of electrons. In our general formulation, the fixed electron number approximation is applied to all subsystems, offering also the possibility of treating all subsystems with WFT. The derivative of the interaction energy functional that carries the inter-system dependence is the embedding potential, that can be decomposed into the Coulomb interactions with the environment (nuclei and frozen electron density) plus derivatives of the non-additive parts of the exchange-correlation and kinetic energy

$$v^{\rm I}_{\rm emb}(\mathbf{r}) = \frac{\delta E_{\rm int}[\rho]}{\delta\rho_{\rm I}(\mathbf{r})} = v^{\rm II}_{\rm nuc}(\mathbf{r}) + \int \frac{\rho_{\rm II}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}{\rm d}\mathbf{r}' + v^{\rm nadd}_{\rm xc}[\rho_{\rm I}, \rho_{\rm II}] + \left.\frac{\delta T^{\rm nadd}_{\rm s}[\rho]}{\delta\rho(\mathbf{r})}\right|_{\rho_{\rm I}}. \tag{9}$$

Regardless of the chosen density parameterization, and methods for evaluating the subsystem energy, $E^{\rm nadd}_{\rm xc}$ and $T^{\rm nadd}_s$ are always calculated using a density functional. In this article, we will not discuss details of these density functionals (and their derivatives); benchmarks of various kinetic energy functionals for use in FDE are well available.[23–25,58] We note that improved functionals can nowadays be easily implemented via automatic differentiation techniques.[59]

The conventional way to obtain the density of a subsystem $i$ is to construct a non-interacting reference system and employ the Kohn-Sham equation for a constrained electron density (KSCED).[15] In this equation,

$$F^{\rm KS}\phi_{p_i}(\mathbf{r}) = \left[-\frac{1}{2}\Delta + v^i_{\rm nuc}(\mathbf{r}) + \int \frac{\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}{\rm d}\mathbf{r}' + v_{\rm xc}[\rho_i](\mathbf{r}) + v^i_{\rm emb}(\mathbf{r})\right]\phi_{p_i}(\mathbf{r}) = \varepsilon^{\rm KS}_{p_i}\phi_{p_i}(\mathbf{r}), \tag{10}$$

the local embedding potential of Eq. (9) is seen to represent the environment. The subsystem energy is then calculated according to Eq. (5) if desired.

The alternative way to obtain the density is to employ WFT and consider a constrained minimization of the total energy of the system as a function of the free parameters in the wave function used to model the electrons contained in system $i$. In this minimization, the contribution from the derivative of the interaction energy is identical to the DFT expression given above in Eq. (9), but the terms coming from the subsystem energy itself depend on the chosen wave-function model and parameterization thereof. If we take the simplest wave-function model, the single-determinant (SD) Hartree-Fock wave

function, we find the Hartree-Fock analog of the KSCED equation,

$$F^{\mathrm{HF}}\phi_{p_i}(\mathbf{r}) = \left[-\frac{1}{2}\Delta + v^i_{\mathrm{nuc}}(\mathbf{r}) + \int \frac{\rho_i(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}' + v^i_{\mathrm{emb}}(\mathbf{r})\right] \times \phi_{p_i}(\mathbf{r}) - \int \frac{\gamma_i(\mathbf{r}',\mathbf{r})\phi_{p_i}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}' = \varepsilon^{\mathrm{HF}}_{p_i}\phi_{p_i}(\mathbf{r}). \quad (11)$$

The resulting orbitals may be used to evaluate the subsystem Hartree-Fock energy. Note that this energy should not contain the interaction energy contribution, even though for both Hartree-Fock and Kohn-Sham, the definition of canonical orbitals includes the embedding potential,

$$F_{p_\mathrm{I}q_\mathrm{I}} \rightarrow F_{p_\mathrm{I}q_\mathrm{I}} + \langle p_\mathrm{I}|v^{\mathrm{I}}_{\mathrm{emb}}|q_\mathrm{I}\rangle = \delta_{p_\mathrm{I}q_\mathrm{I}}\varepsilon_{p_\mathrm{I}}. \quad (12)$$

The density-only expression of Eq. (6) is applicable for all methods and the total energy is given according to Eq. (4), independent of the precise method used to determine $E_i$ and $\rho_i$. Since one cannot straightforwardly calculate the interaction energy contribution as an expectation value of the embedding potential (due to the partially nonlinear dependence of the energy on the density[53,60,61]), the subsystem DFT scheme differs from most other embedding approaches (see Sec. VI). Another point that should be mentioned at this stage is that we always assume that the wave-function method is capable of providing the exact subsystem density and energy. This is only rigorously true for a full configuration interaction method in a complete basis, but compensating for missing electron correlation contributions in WFT by adding a correlation functional[62,63] is a notoriously difficult problem that we will not attempt to solve in this work.

## III. QUASI-ENERGY RESPONSE THEORY

In our derivation of subsystem response theory, we follow the work on frequency-dependent response functions of Christiansen *et al.*[52] which is restricted in its time-averaged formulation to time-periodic perturbations. For the sake of completeness and to introduce the notation, we repeat the most important definitions and equations. For a general discussion concerning the applicability of the quasi-energy formalism and DFT, see, e.g., Ref. 53. Let $\hat{H}^t$ be a general, time-dependent Hamiltonian,

$$\hat{H}^t = \hat{H} + \hat{V}^t, \quad (13)$$

where $\hat{V}^t$ is the time-dependent perturbation operator which is given as

$$\hat{V}^t = \sum_{k=-N}^{N} \exp(-i\omega_k t)\sum_x \epsilon_x(\omega_k)\hat{X} \quad (14)$$

with $\epsilon_x(\omega_k)$ parameters that denote the strength of the perturbing fields. The linear response function is defined via the time evolution of the expectation value of an operator $\hat{X}$,

$$\langle X\rangle(t) = \langle X\rangle_0 + \sum_{k_1}\exp(-i\omega_{k_1}t)\sum_y \langle\langle X;Y\rangle\rangle_{\omega_{k_1}}\epsilon_y(\omega_{k_1}) + \cdots. \quad (15)$$

The quasi-energy $Q(t)$ and its time-average $\{Q(t)\}_T$ are defined as

$$Q(t) = \langle\tilde{O}|\left(\hat{H}^t - i\frac{\partial}{\partial t}\right)|\tilde{O}\rangle, \quad (16)$$

$$\{Q(t)\}_T = \frac{1}{T}\int_{-T/2}^{T/2} Q(t)\,\mathrm{d}t, \quad (17)$$

respectively, where $T$ is the period of the perturbation in Eq. (14) and the tilde denotes the phase-isolated form of the wave function. For variational methods, such as Hartree-Fock or DFT, the calculation of the linear response function proceeds directly from the quasi-energy itself. In this treatment, the linear response function is obtained as the second derivative of the time-averaged quasi-energy. For non-variational wave functions, first a Lagrangian,

$$L(\lambda,\dot{\lambda},\bar{\lambda}) = Q(\lambda,\dot{\lambda}) + \bar{\lambda}\,e(\lambda,\dot{\lambda}), \quad (18)$$

with appropriate constraints $e(\lambda,\dot{\lambda})$ and Lagrange multipliers $\bar{\lambda}$ needs to be introduced before proceeding to derive the response functions.

Lagrangian-based formulations can be extended to incorporate environment effects,[64] and will be central to our development. For the subsystem treatment, we define a total quasi-energy Lagrangian consisting of the quasi-energy expressions of the subsystems as well as their interaction,

$$L_{\mathrm{tot}}(t) = \sum_i L_i(t) + Q_{\mathrm{int}}(t). \quad (19)$$

While Eq. (19) allows an arbitrary number of subsystems, it is more convenient to again restrict the derivation to the case of two subsystems. Thus, in the following, subsystem I represents the "active" subsystem of interest, whereas subsystem II consists of the sum of all other subsystems and represents the "environment." The total quasi-energy expression in Eq. (19) then reduces to

$$L_{\mathrm{tot}}(\lambda,\dot{\lambda},\bar{\lambda};\lambda_{\mathrm{II}},\dot{\lambda}_{\mathrm{II}},\bar{\lambda}_{\mathrm{II}}) = [Q(\lambda,\dot{\lambda}) + \bar{\lambda}e(\lambda,\dot{\lambda})] + Q_{\mathrm{int}}(\lambda,\dot{\lambda},\bar{\lambda};\lambda_{\mathrm{II}},\dot{\lambda}_{\mathrm{II}},\bar{\lambda}_{\mathrm{II}}) + [Q_{\mathrm{II}}(\lambda_{\mathrm{II}},\dot{\lambda}_{\mathrm{II}}) + \bar{\lambda}_{\mathrm{II}}e_{\mathrm{II}}(\lambda_{\mathrm{II}},\dot{\lambda}_{\mathrm{II}})], \quad (20)$$

where we make the dependence of $L_{\mathrm{tot}}(t)$ on the wave-function parameters and constraints explicit. Note that all inter-system dependencies are contained in the interaction energy. In the following, we will assume that the environment is optimized with DFT, which is variational, so that we can omit the multipliers $\bar{\lambda}_{\mathrm{II}}$.

The time-dependent Lagrangian can then be expanded in orders of the perturbation,

$$L(t) = L^{(0)} + L^{(1)}(t) + L^{(2)}(t) + \cdots, \quad (21)$$

TABLE I. Vectors and matrices for general response functions (see Ref. 52).

| Quantity | Derivative expr. | Quantity | Derivative expr. | Quantity | Derivative expr. |
|---|---|---|---|---|---|
| $\boldsymbol{\eta}^Y$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\lambda^{(1)}(\omega_X)\partial\epsilon_Y(\omega_Y)}$ | | | ${}^{\mathrm{II}}\boldsymbol{\eta}^Y$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\lambda_{\mathrm{II}}^{(1)}(\omega_X)\partial\epsilon_Y(\omega_Y)}$ |
| $\boldsymbol{\xi}^Y$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\bar{\lambda}^{(1)}(\omega_X)\partial\epsilon_Y(\omega_Y)}$ | | | | |
| $\mathbf{J}(\omega_Y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\bar{\lambda}^{(1)}(\omega_X)\partial\bar{\lambda}^{(1)}(\omega_Y)}$ | | | | |
| $\mathbf{A}(\omega_Y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\lambda^{(1)}(\omega_X)\partial\bar{\lambda}^{(1)}(\omega_Y)}$ | ${}^{\mathrm{I,II}}\mathbf{A}(\omega_Y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\bar{\lambda}^{(1)}(\omega_X)\partial\lambda_{\mathrm{II}}^{(1)}(\omega_Y)}$ | | |
| $\mathbf{F}(\omega_Y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\lambda^{(1)}(\omega_X)\partial\lambda^{(1)}(\omega_Y)}$ | ${}^{\mathrm{I,II}}\mathbf{F}(\omega_Y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\lambda^{(1)}(\omega_X)\partial\lambda_{\mathrm{II}}^{(1)}(\omega_Y)}$ | ${}^{\mathrm{II,II}}\mathbf{F}(\omega_Y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\lambda_{\mathrm{II}}^{(1)}(\omega_X)\partial\lambda_{\mathrm{II}}^{(1)}(\omega_Y)}$ |
| $\boldsymbol{\eta}$ | $\frac{\partial\{Q^{(0)}\}_T}{\partial\lambda^{(0)}}$ | | | ${}^{\mathrm{II}}\boldsymbol{\eta}$ | $\frac{\partial\{Q_{\mathrm{II}}^{(0)}\}_T}{\partial\lambda^{(0)}}$ |

and Fourier transformed to the frequency domain. We may similarly expand the parameters in terms of the perturbation strength, obtaining, e.g., the first-order expression,

$$\lambda^{(1)}(\omega_{k_1}) = \sum_x \epsilon_x(\omega_{k_1})\,\lambda^X(\omega_{k_1}). \tag{22}$$

Response functions are obtained as derivatives of the time-averaged quasi-energy Lagrangian of $n$th order $\{L^{(n)}\}_T$ with respect to the field-strength variables, e.g., for second order,

$$\langle\langle X;Y\rangle\rangle_{\omega_{k_1}} = \frac{\mathrm{d}^2\{L^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})}, \quad \text{where} \quad \omega_{k_1} = -\omega_0. \tag{23}$$

Using the abbreviations given in Table I and the fact that the contributions from second-order parameters, such as $\lambda^{(2)}(t)$, $\bar{\lambda}^{(2)}(t)$, and $\lambda_{\mathrm{II}}^{(2)}(t)$, are zero due to the $2n+1$ rule, the second derivative becomes

$$\frac{\mathrm{d}^2\{L_{\mathrm{tot}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})} = \frac{\mathrm{d}^2\{L_{\mathrm{I}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})} + \frac{\mathrm{d}^2\{Q_{\mathrm{II}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})} + \frac{\mathrm{d}^2\{Q_{\mathrm{int}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})}, \tag{24}$$

where the first two terms are obtained similar to Eq. (3.28) of Ref. 52,

$$\begin{aligned}\frac{\mathrm{d}^2\{L_{\mathrm{I}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})} &= P(X(\omega_0),Y(\omega_{k_1}))\\ &\times\left\{\left[\boldsymbol{\eta}^X + \frac{1}{2}\mathbf{F}\,\boldsymbol{\lambda}^X(\omega_0)\right]\boldsymbol{\lambda}^Y(\omega_{k_1})\right.\\ &+\bar{\boldsymbol{\lambda}}^X(\omega_0)\left[\frac{1}{2}\mathbf{J}\,\bar{\boldsymbol{\lambda}}^Y(\omega_{k_1}) + \boldsymbol{\xi}^Y\right.\\ &\left.\left.+\mathbf{A}\,\boldsymbol{\lambda}^Y(\omega_{k_1})\right]\right\},\end{aligned} \tag{25}$$

$$\begin{aligned}\frac{\mathrm{d}^2\{Q_{\mathrm{II}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})} &= P\big(X(\omega_0),Y(\omega_{k_1})\big)\\ &\times\left\{\left[{}^{\mathrm{II}}\boldsymbol{\eta}^X + \frac{1}{2}{}^{\mathrm{II,II}}\mathbf{F}\,\boldsymbol{\lambda}_{\mathrm{II}}^X(\omega_0)\right]\boldsymbol{\lambda}_{\mathrm{II}}^Y\big(\omega_{k_1}\big)\right\},\end{aligned} \tag{26}$$

while the interaction term reads

$$\begin{aligned}\frac{\mathrm{d}^2\{Q_{\mathrm{int}}^{(2)}\}_T}{\mathrm{d}\epsilon_x(\omega_0)\mathrm{d}\epsilon_y(\omega_{k_1})} &= P\big(X(\omega_0),Y(\omega_{k_1})\big)\\ &\times\left\{\bar{\boldsymbol{\lambda}}^X(\omega_0)\,\mathbf{A}^{\mathrm{int}}\,\boldsymbol{\lambda}^Y(\omega_{k_1})\right.\\ &+\frac{1}{2}\left[\mathbf{F}^{\mathrm{int}}\,\boldsymbol{\lambda}^X(\omega_0)\boldsymbol{\lambda}^Y(\omega_{k_1})\right.\\ &+\mathbf{J}^{\mathrm{int}}\,\bar{\boldsymbol{\lambda}}^X(\omega_0)\bar{\boldsymbol{\lambda}}^Y\big(\omega_{k_1}\big)]\\ &+{}^{\mathrm{I,II}}\mathbf{F}^{\mathrm{int}}\,\boldsymbol{\lambda}^X(\omega_0)\boldsymbol{\lambda}_{\mathrm{II}}^Y\big(\omega_{k_1}\big)\\ &\left.\left.+\bar{\boldsymbol{\lambda}}^X(\omega_0)\;{}^{\mathrm{I,II}}\mathbf{A}^{\mathrm{int}}\;\boldsymbol{\lambda}_{\mathrm{II}}^Y\big(\omega_{k_1}\big)\right\}.\right.\end{aligned} \tag{27}$$

$P(X(\omega_0),Y(\omega_{k_1}))$ ensures symmetry with respect to the interchange of $X$ and $Y$ and associated frequencies $\omega_0$ and $\omega_{k_1}$. Throughout the paper, we use supermatrix notation where vectors and matrices are multiplied in order—a notation useful especially for higher order response properties.

For the present purpose, we need only the first-order perturbed quantities. These are obtained by requiring stationarity of the Lagrangian with respect to variations in first-order multipliers $\bar{\boldsymbol{\lambda}}^X$, first-order amplitudes $\boldsymbol{\lambda}^X$, and first-order parameters of the environment $\boldsymbol{\lambda}_{\mathrm{II}}^X$, yielding a set of coupled linear response equations for frequency $\omega_y$,

$$\left(\begin{array}{c}\frac{0}{0}\\ \hline 0\end{array}\right) = \left(\begin{array}{c}\frac{\boldsymbol{\xi}^Y}{\boldsymbol{\eta}^Y}\\ \hline {}^{\mathrm{II}}\boldsymbol{\eta}^Y\end{array}\right) + \left(\begin{array}{c|c|c}\mathbf{A}(\omega_y) & \mathbf{J} & {}^{\mathrm{I,II}}\mathbf{A}\\ \hline \mathbf{F} & \mathbf{A}(-\omega_y) & {}^{\mathrm{I,II}}\mathbf{F}\\ \hline {}^{\mathrm{I,II}}\mathbf{F} & {}^{\mathrm{I,II}}\mathbf{A} & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y)\end{array}\right) \times \left(\begin{array}{c}\boldsymbol{\lambda}^Y(\omega_y)\\ \hline \bar{\boldsymbol{\lambda}}^Y(\omega_y)\\ \hline \boldsymbol{\lambda}_{\mathrm{II}}^Y(\omega_y)\end{array}\right). \tag{28}$$

It is convenient to separate contributions from the subsystems Lagrangian and their interaction explicitly by decomposing

the matrix above as

$$\left(\begin{array}{c|c|c} \mathbf{A}(\omega_y) & 0 & 0 \\ \hline \mathbf{F} & \mathbf{A}(-\omega_y) & 0 \\ \hline 0 & 0 & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y) \end{array}\right) + \left(\begin{array}{c|c|c} \mathbf{A}_{\mathrm{int}} & \mathbf{J}_{\mathrm{int}} & {}^{\mathrm{I,II}}\mathbf{A}_{\mathrm{int}} \\ \hline \mathbf{F}_{\mathrm{int}} & \mathbf{A}_{\mathrm{int}} & {}^{\mathrm{I,II}}\mathbf{F}_{\mathrm{int}} \\ \hline {}^{\mathrm{I,II}}\mathbf{F}_{\mathrm{int}} & {}^{\mathrm{I,II}}\mathbf{A}_{\mathrm{int}} & {}^{\mathrm{II,II}}\mathbf{F}_{\mathrm{int}} \end{array}\right). \tag{29}$$

The $\mathbf{J}$-term drops out when the $2n+2$ rule based on decoupled response equations for $\boldsymbol{\lambda}^X$ and $\bar{\boldsymbol{\lambda}}^X$ can be applied[10,52] but contains a non-zero contribution due to the interaction. The response equations (Eq. (28)) allow to write the expression for the total response function as

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = \frac{1}{2}C^{\pm\omega}\,P(X(\omega_x),Y(\omega_y))\times\left\{\frac{1}{2}\mathbf{J}\,\bar{\boldsymbol{\lambda}}^X(\omega_x)\bar{\boldsymbol{\lambda}}^Y(\omega_y) + \left[\begin{pmatrix}\boldsymbol{\eta}^X \\ {}^{\mathrm{II}}\eta^X\end{pmatrix}^T + \frac{1}{2}\begin{pmatrix}\boldsymbol{\lambda}^X(\omega_x) \\ \boldsymbol{\lambda}^X_{\mathrm{II}}(\omega_x)\end{pmatrix}^T \times \left(\begin{array}{c|c} \mathbf{F}(\omega_y) & {}^{\mathrm{I,II}}\mathbf{F}(\omega_y) \\ \hline {}^{\mathrm{I,II}}\mathbf{F}(\omega_y) & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y)\end{array}\right)\right]\begin{pmatrix}\boldsymbol{\lambda}^Y(\omega_y) \\ \boldsymbol{\lambda}^Y_{\mathrm{II}}(\omega_y)\end{pmatrix}\right\}. \tag{30}$$

The operator $C^{\pm\omega}$ enforces symmetrization with respect to simultaneous complex conjugation and inversion of the sign of the frequencies.[52] Alternatively, we can rewrite Eq. (30) in terms of the individual ("uncoupled") subsystems as well as the coupling contribution,

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = \langle\langle X;Y\rangle\rangle^{\mathrm{unc,I}}_{\omega_y} + \langle\langle X;Y\rangle\rangle^{\mathrm{unc,II}}_{\omega_y} + \langle\langle X;Y\rangle\rangle^{\mathrm{I,II}}_{\omega_y}. \tag{31}$$

The individual subsystem contributions can be extracted from Eq. (30) and are given as[52,53,65]

$$\langle\langle X;Y\rangle\rangle^{\mathrm{unc,I}}_{\omega_y} = \frac{1}{2}C^{\pm\omega}P(X(\omega_x),Y(\omega_y)) \times\left\{\frac{1}{2}\mathbf{J}\,\bar{\boldsymbol{\lambda}}^X(\omega_x)\bar{\boldsymbol{\lambda}}^Y(\omega_y) + \left[\boldsymbol{\eta}^X + \frac{1}{2}\mathbf{F}\,\boldsymbol{\lambda}^X(\omega_x)\right]\boldsymbol{\lambda}^Y(\omega_y)\right\}, \tag{32}$$

$$\langle\langle X;Y\rangle\rangle^{\mathrm{unc,II}}_{\omega_y} = \frac{1}{2}C^{\pm\omega}P(X(\omega_x),Y(\omega_y)) \times\left[{}^{\mathrm{II}}\boldsymbol{\eta}^X + \frac{1}{2}\,{}^{\mathrm{II,II}}\mathbf{F}\,\boldsymbol{\lambda}^X_{\mathrm{II}}(\omega_x)\right]\boldsymbol{\lambda}^Y_{\mathrm{II}}(\omega_y). \tag{33}$$

If only these are included, the result will be denoted "uncoupled"—implying that the response of the interaction energy is included in the "intra-subsystem" blocks in Eq. (28) but that the "inter-subsystem" blocks are neglected. The full, i.e., coupled, result includes also the inter-subsystem response function

$$\langle\langle X;Y\rangle\rangle^{\mathrm{I,II}}_{\omega_y} = C^{\pm\omega}P(X(\omega_x),Y(\omega_y))\{{}^{\mathrm{I,II}}\mathbf{F}_{\mathrm{int}}\boldsymbol{\lambda}^X(\omega_x)\boldsymbol{\lambda}^Y_{\mathrm{II}}(\omega_y)\}. \tag{34}$$

In the limit of non-interacting subsystems, all interaction contributions vanish and the total response function reduces to the sum of the isolated subsystems.

For variational wave functions, we can remove contributions of Lagrangian multipliers and the environment in Eq. (28), and the linear response function then takes the form[52]

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = -\boldsymbol{\eta}^X\,(\mathbf{F}(\omega_y))^{-1}\,\boldsymbol{\eta}^Y. \tag{35}$$

The excitation energies are finally calculated from the poles of Eq. (35), whereas properties are obtained from the evaluation of the linear response function at a given frequency $\omega_y$. Due to the computational cost, however, instead of calculating the inverse in Eq. (35), typically the linear set of equations,

$$\mathbf{F}(\omega_y)\,\boldsymbol{\lambda}^Y(\omega_y) = -\boldsymbol{\eta}^Y, \tag{36}$$

is solved,[53,66] from which the linear response function and thus properties such as, frequency-dependent dipole-dipole polarizabilities, are calculated as

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = \boldsymbol{\eta}^X\cdot\boldsymbol{\lambda}^Y(\omega_y). \tag{37}$$

With separate coupling contributions, it is possible to devise computational strategies adapted to the properties of interest.[35] For instance, in cases where the coupling is important, such as excitation energies or polarizabilities, these can be approximated, and in the cases where these are less important, such as local excitations or NMR chemical shifts, these can be dropped. Furthermore, it offers the possibility to operate mostly with the quantities for the isolated subsystems, thus avoiding the formation and handling of matrices and vectors with the dimension of the supermolecular basis.[31,32] In Secs. IV and V, we will derive explicit working expressions for the components of the quasi-energy Lagrangian, as well as the different matrices and vectors needed.

## IV. REVISITING DFT-IN-DFT RESPONSE THEORY

In this section, we discuss the DFT-in-DFT FDE response theory using the quasi-energy formalism. In order to reformulate the subsystem Kohn-Sham theory in the formalism of second quantization, we follow closely the notation used by Saue and Helgaker.[67] We start by introducing the parameterization of a closed-shell Kohn-Sham determinant of one subsystem in terms of an unitary exponential orbital-rotation operator,

$$|0_i\rangle = \exp(\hat{\kappa}_i)|\tilde{0}_i\rangle \quad \text{with}\ \hat{\kappa}_i = \sum_{p_i>q_i}(\kappa_{p_iq_i}E_{p_iq_i} - \kappa^*_{p_iq_i}E_{q_ip_i}). \tag{38}$$

In the framework of FDE, the subsystem orbitals are to be considered as two independent sets of auxiliary quantities that serve to provide the exact density and its responses. This implies that admixture of orbitals from a subsystem $i$ into the orbitals of a different subsystem $j$ is to be excluded, and the

orbital rotations fulfil the condition

$$\kappa_{p_i q_j} = \delta_{ij}\kappa_{p_i q_i}. \tag{39}$$

The density of a given subsystem can be written as

$$\begin{aligned}\rho_i(\mathbf{r}, \boldsymbol{\kappa}_i) &= \langle 0_i | \exp(-\hat{\kappa}_1)\hat{\rho}_i(\mathbf{r})\exp(\hat{\kappa}_1)|0_i\rangle \\ &= \sum_{p_i q_i} \rho_{p_i q_i}(\mathbf{r}) D_{p_i q_i}(\boldsymbol{\kappa}_i),\end{aligned} \tag{40}$$

where $\hat{\rho}(\mathbf{r})$ is the density operator, given in second quantization as

$$\hat{\rho}_i(\mathbf{r}) = \sum_{p_i q_i} \rho_{p_i q_i}(\mathbf{r})\, E_{p_i q_i}, \tag{41}$$

$$\rho_{p_i q_i}(\mathbf{r}) = \phi^*_{p_i}(\mathbf{r})\phi_{q_i}(\mathbf{r}). \tag{42}$$

In order to calculate the density response, the density matrix $\mathbf{D}(\boldsymbol{\kappa}_i)$ is expanded in orders of orbital rotation parameters using the Baker-Campell-Hausdorff (BCH) formula,

$$\begin{aligned}D_{p_i q_i}(\boldsymbol{\kappa}_i) &= \langle 0_i | E_{p_i q_i} | 0_i\rangle + \langle 0_i | [\hat{\kappa}_i, E_{p_i q_i}] | 0_i\rangle \\ &\quad + \frac{1}{2!}\langle 0_i | [\hat{\kappa}_i, [\hat{\kappa}_i, E_{p_i q_i}]] | 0_i\rangle + \mathcal{O}\big(\boldsymbol{\kappa}_i^3\big).\end{aligned} \tag{43}$$

For convenience, we also introduce the (response) density matrices $\boldsymbol{\Gamma}$ for a single-determinant exponential parameterization,

$$\Gamma^{\mathrm{SD}}_{x_i y_i; p_i q_i} = \left(\frac{\partial D^{\mathrm{SD}\,(1)}_{x_i y_i}}{\partial \kappa^{(1)}_{p_i q_i}}\right) = \langle 0_i | [E_{p_i q_i}, E_{x_i y_i}] | 0_i\rangle, \tag{44}$$

$$\Gamma^{\mathrm{SD}}_{x_i y_i; p_i q_i, r_i s_i} = \left(\frac{\partial^2 D^{\mathrm{SD}\,(2)}_{x_i y_i}}{\partial \kappa^{(1)}_{p_i q_i}\partial \kappa^{(1)}_{r_i s_i}}\right) = \frac{1}{2}\langle 0_i | [E_{p_i q_i}, [E_{r_i s_i}, E_{x_i y_i}]] | 0_i\rangle. \tag{45}$$

Note that the number of indices indicates the order of the derivatives.

### A. DFT response theory for isolated subsystems

We first consider the DFT response theory using the quasi-energy formalism and second quantization[53,65] for a single isolated subsystem. Starting point is the DFT quasi-energy expression,

$$\begin{aligned}Q[\rho](t, \epsilon) &= T_s[\rho] + V^t[\rho] + V_{\mathrm{nuc}}[\rho] + J[\rho] \\ &\quad + E_{\mathrm{xc}}[\rho(\mathbf{r}, t)] - \left\{\langle \tilde{0} | i\frac{\partial}{\partial t} | \tilde{0}\rangle\right\}_T,\end{aligned} \tag{46}$$

where $J[\rho]$ denotes the Coulomb contribution. The energy expression is obtained as the (time-independent) zeroth-order quasi-energy,

$$Q^{(0)}[\rho^{(0)}] = T_s[\rho^{(0)}] + V_{\mathrm{nuc}}[\rho^{(0)}] + J[\rho^{(0)}] + E_{\mathrm{xc}}[\rho^{(0)}]. \tag{47}$$

In order to calculate the linear response, we adopt the adiabatic approximation, assuming that the time dependence of the exchange-correlation potential may be fully described through the time evolution of the density. For a single subsystem, only expressions for $\mathbf{F}$ and $\boldsymbol{\eta}^Y$ in terms of the single-determinant Kohn-Sham ansatz are needed,[52,53,68]

$$F_{pq,rs}(\omega_y) = \frac{\partial^2\{Q^{(2)}\}_T}{\partial \kappa^{(1)}_{pq}(\omega_x)\partial \kappa^{(1)}_{rs}(\omega_y)} = E^{[2]}_{\mathrm{DFT};pq,rs} - \omega_y S^{[2]}_{pq,rs}, \tag{48}$$

$$\eta^Y_{pq} = \frac{\partial^2\{Q^{(2)}\}_T}{\partial \kappa^{(1)}_{pq}(\omega_x)\partial \epsilon_Y(\omega_y)} = \langle 0 | [\hat{Y}, E_{pq}] | 0\rangle, \tag{49}$$

and the Hessian $\mathbf{E}^{[2]}_{\mathrm{DFT}}$ for the closed-shell case reads

$$\mathbf{E}^{[2]}_{\mathrm{DFT}} = \begin{pmatrix}\mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^*\end{pmatrix}, \tag{50}$$

$$A_{ia,jb} = \delta_{ij}\delta_{ab}\big(\varepsilon^{\mathrm{KS}}_a - \varepsilon^{\mathrm{KS}}_i\big) + 2(ia|bj) + (ia|w_{\mathrm{xc}}|bj), \tag{51}$$

$$B_{ia,jb} = 2(ia|jb) + (ia|w_{\mathrm{xc}}|jb). \tag{52}$$

Throughout the article, round brackets denote charge-cloud notation for two-electron integrals.

### B. DFT-in-DFT response theory

To the contributions from the isolated subsystems considered above, we may next add the contributions from the interaction term $Q_{\mathrm{int}}$. Starting point is the expansion

$$\begin{aligned}Q_{\mathrm{int}} &= E^{[0]}_{\mathrm{int}} + \mathbf{E}^{[1]}_{\mathrm{int;I}}\boldsymbol{\kappa}_{\mathrm{I}} + \mathbf{E}^{[1]}_{\mathrm{int;II}}\boldsymbol{\kappa}_{\mathrm{II}} \\ &\quad + \frac{1}{2!}\big(\mathbf{E}^{[2]}_{\mathrm{int;I,I}}\boldsymbol{\kappa}_{\mathrm{I}}\boldsymbol{\kappa}_{\mathrm{I}} + \mathbf{E}^{[2]}_{\mathrm{int;II,II}}\boldsymbol{\kappa}_{\mathrm{II}}\boldsymbol{\kappa}_{\mathrm{II}} + \mathbf{E}^{[2]}_{\mathrm{int;II,I}}\boldsymbol{\kappa}_{\mathrm{II}}\boldsymbol{\kappa}_{\mathrm{I}} \\ &\quad + \mathbf{E}^{[2]}_{\mathrm{int;I,II}}\boldsymbol{\kappa}_{\mathrm{I}}\boldsymbol{\kappa}_{\mathrm{II}}\big) + \mathcal{O}(\boldsymbol{\kappa}^3).\end{aligned} \tag{53}$$

In order to evaluate the derivatives of a functional $E[\rho(\boldsymbol{\kappa})]$, the functional chain rule,[60]

$$\left(\frac{\partial E[\rho(\mathbf{r}, \boldsymbol{\kappa})]}{\partial \kappa_{pq}}\right) = \int \left(\frac{\delta E[\rho]}{\delta \rho(\mathbf{r}')}\right)\left(\frac{\partial \rho(\mathbf{r}', \boldsymbol{\kappa})}{\partial \kappa_{pq}}\right) \mathrm{d}\mathbf{r}', \tag{54}$$

is employed. We furthermore use the short-hand notation for the first [Eq. (9)] and second functional derivative, respectively,

$$v^i_{\mathrm{emb}}(\mathbf{r}') = \frac{\delta E_{\mathrm{int}}[\rho]}{\delta \rho_i(\mathbf{r}')}, \quad w^{ij}_{\mathrm{emb}}(\mathbf{r}', \mathbf{r}'') = \frac{\delta^2 E_{\mathrm{int}}[\rho]}{\delta \rho_i(\mathbf{r}')\delta \rho_j(\mathbf{r}'')}. \tag{55}$$

As this interaction between system I and the sum of the other systems can be considered an embedding of system I, we will denote the potential and the kernel arising from the interaction term as the "embedding potential" and "embedding kernel," respectively. The first-order contribution of the interaction term can be formulated as

$$\sum_{r_i s_i} E^{[1]}_{\mathrm{int};r_i s_i}\kappa_{r_i s_i} = \sum_{r_i s_i}\sum_{x_i y_i} \Gamma^{\mathrm{SD}}_{x_i y_i; r_i s_i}\langle x_i | v^i_{\mathrm{emb}} | y_i\rangle \kappa_{r_i s_i}. \tag{56}$$

Utilizing the definition of the density matrix in Eq. (44), the non-zero elements for subsystem I are simply

$$E^{[1]}_{\mathrm{int};i_{\mathrm{I}} a_{\mathrm{I}}} = \langle i_{\mathrm{I}} | v^{\mathrm{I}}_{\mathrm{emb}} | a_{\mathrm{I}}\rangle. \tag{57}$$

For the diagonal ($i = j$) second-order term

$$E^{[2]}_{\mathrm{int};p_iq_i,r_is_i} = \iint \left(\frac{\delta^2 E_{\mathrm{int}}}{\delta\rho(\mathbf{r}')\rho(\mathbf{r}'')}\right)\left(\frac{\partial\rho^{(1)}(\mathbf{r}')}{\partial\kappa^{(1)}_{p_iq_i}}\right) \times \left(\frac{\partial\rho^{(1)}(\mathbf{r}'')}{\partial\kappa^{(1)}_{r_is_i}}\right)\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}'' + \int\left(\frac{\delta E_{\mathrm{int}}}{\delta\rho(\mathbf{r}')}\right)\left(\frac{\partial^2\rho^{(2)}(\mathbf{r}')}{\partial\kappa^{(1)}_{p_iq_i}\partial\kappa^{(1)}_{r_is_i}}\right)\mathrm{d}\mathbf{r}', \quad (58)$$

the second derivative needs to be evaluated, which yields

$$E^{[2]}_{\mathrm{int};p_iq_i,r_is_i} = \sum_{x_iy_i,t_iu_i}\Gamma^{\mathrm{SD}}_{x_iy_i;p_iq_i}\,\Gamma^{\mathrm{SD}}_{t_iu_i;r_is_i}\,(x_iy_i|w^{ii}_{\mathrm{emb}}|t_iu_i) + \sum_{x_iy_i}\Gamma^{\mathrm{SD}}_{x_iy_i;p_iq_i,r_is_i}\,\langle x_i|v^{i}_{\mathrm{emb}}|y_i\rangle. \quad (59)$$

After evaluation of the density matrices, this gives the non-vanishing elements for subsystem I,

$$E^{[2]}_{\mathrm{int};i_1a_1,b_1j_1} = (i_1a_1|w^{\mathrm{I,I}}_{\mathrm{emb}}|b_1j_1) + \delta_{i_1j_1}\langle a_1|v^{\mathrm{I}}_{\mathrm{emb}}|b_1\rangle - \delta_{a_1b_1}\langle i_1|v^{\mathrm{I}}_{\mathrm{emb}}|j_1\rangle, \quad (60)$$

$$E^{[2]}_{\mathrm{int};i_1a_1,j_1b_1} = (i_1a_1|w^{\mathrm{I,I}}_{\mathrm{emb}}|j_1b_1). \quad (61)$$

The mixed second derivative leads to a coupling between the sub-blocks of the Hessian for which only the kernel contributions ($w_{\mathrm{emb}}$) survive,

$$E^{[2]}_{\mathrm{int};p_iq_i,r_js_j} = \iint w^{ij}_{\mathrm{emb}}(\mathbf{r}',\mathbf{r}'')\left(\frac{\partial\rho^{(1)}_i(\mathbf{r}')}{\partial\kappa^{(1)}_{p_iq_i}}\right)\left(\frac{\partial\rho^{(1)}_j(\mathbf{r}'')}{\partial\kappa^{(1)}_{r_js_j}}\right)\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}'' = \sum_{x_iy_i,t_ju_j}\Gamma^{\mathrm{SD}}_{x_iy_i;p_iq_i}\,\Gamma^{\mathrm{SD}}_{t_ju_j;r_js_j}\,(x_iy_i|w^{ij}_{\mathrm{emb}}|t_ju_j). \quad (62)$$

The full embedding kernel contribution can be expressed as

$$w^{ij}_{\mathrm{emb}}(\mathbf{r}',\mathbf{r}'') = w_{\mathrm{xck}}(\mathbf{r}',\mathbf{r}'') - \delta_{ij}\left[\left.\frac{\delta^2 T_{\mathrm{s}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_i} + \left.\frac{\delta^2 E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_i}\right] + (1-\delta_{ij})\frac{1}{|\mathbf{r}'-\mathbf{r}''|}, \quad (63)$$

with the delta function indicating that the Coulomb term is only evaluated for the inter-subsystem interaction ($i \neq j$). For convenience, we introduce auxiliary kernel contributions to specify the kinetic energy and exchange-correlation terms in the embedding kernel,

$$w_{\mathrm{xck}}(\mathbf{r}',\mathbf{r}'') = \left.\frac{\delta^2 E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_{\mathrm{tot}}} + \left.\frac{\delta^2 T_{\mathrm{s}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_{\mathrm{tot}}}, \quad (64)$$

$$w^{i}_{\mathrm{xck}}(\mathbf{r}',\mathbf{r}'') = w_{\mathrm{xck}}(\mathbf{r}',\mathbf{r}'') - \left[\left.\frac{\delta^2 T_{\mathrm{s}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_i} + \left.\frac{\delta^2 E_{\mathrm{xc}}}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_i}\right]. \quad (65)$$

### 1. Linear response function

The general response equations in Eq. (28) reduce in the DFT-in-DFT case to

$$\begin{pmatrix}\boldsymbol{\lambda}^{Y}(\omega_y)\\ \boldsymbol{\lambda}^{Y}_{\mathrm{II}}(\omega_y)\end{pmatrix} = -\left(\begin{array}{c|c}\mathbf{F}(\omega_y) & {}^{\mathrm{I,II}}\mathbf{F}\\ \hline {}^{\mathrm{I,II}}\mathbf{F} & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y)\end{array}\right)^{-1}\begin{pmatrix}\boldsymbol{\eta}^{Y}\\ {}_{\mathrm{II}}\boldsymbol{\eta}^{Y}\end{pmatrix}. \quad (66)$$

Substituting this result into Eq. (30), we obtain

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = -\begin{pmatrix}\boldsymbol{\eta}^{X}\\ {}_{\mathrm{II}}\boldsymbol{\eta}^{X}\end{pmatrix}^{T}\left(\begin{array}{c|c}\mathbf{F}(\omega_y) & {}^{\mathrm{I,II}}\mathbf{F}\\ \hline {}^{\mathrm{I,II}}\mathbf{F} & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y)\end{array}\right)^{-1}\begin{pmatrix}\boldsymbol{\eta}^{Y}\\ {}_{\mathrm{II}}\boldsymbol{\eta}^{Y}\end{pmatrix}. \quad (67)$$

The full matrix $\mathbf{F}$ collects the different sub-matrices

$$\left(\begin{array}{c|c}\mathbf{F}(\omega_y) & {}^{\mathrm{II,I}}\mathbf{F}\\ \hline {}^{\mathrm{I,II}}\mathbf{F} & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y)\end{array}\right) = \left(\begin{array}{c|c}{}^{\mathrm{I,I}}\mathbf{E}^{[2]} & {}^{\mathrm{II,I}}\mathbf{E}^{[2]}\\ \hline {}^{\mathrm{I,II}}\mathbf{E}^{[2]} & {}^{\mathrm{II,II}}\mathbf{E}^{[2]}\end{array}\right) - \omega_y\left(\begin{array}{c|c}{}^{\mathrm{I,I}}\mathbf{m} & 0\\ \hline 0 & {}^{\mathrm{II,II}}\mathbf{m}\end{array}\right), \quad (68)$$

where $\mathbf{m}$ stands for the metric containing 1 and $-1$ on the diagonal. With the expressions above, we obtain the Hessian contributions to the matrices ${}^{i,j}\mathbf{F}^{\mathrm{int}}$ as

$$\mathbf{E}^{[2]}_{\mathrm{int};p_iq_i,r_js_j} = \begin{pmatrix}{}^{i,j}\mathbf{A}^{\mathrm{int}} & {}^{i,j}\mathbf{B}^{\mathrm{int}}\\ {}^{i,j}\mathbf{B}^{\mathrm{int}^*} & {}^{i,j}\mathbf{A}^{\mathrm{int}^*}\end{pmatrix}, \quad (69)$$

where the diagonal *kernel contributions* to the elements of Eq. (69) are given (for real orbitals) by

$$(i_1a_1|w^{\mathrm{I,I}}_{\mathrm{emb,I}}|j_1b_1) \rightarrow {}^{\mathrm{I,I}}A^{\mathrm{int}}_{i_1a_1,j_1b_1},\ {}^{\mathrm{I,I}}B^{\mathrm{int}}_{i_1a_1,j_1b_1}. \quad (70)$$

Adding this contribution to those from Eqs. (48) and (50), we obtain the complete expressions for the supermatrix $\mathbf{F}$ for each of the subsystems, e.g., for subsystem I,

$${}^{\mathrm{I,I}}A_{i_1a_1,j_1b_1} = \delta_{i_1j_1}\delta_{a_1b_1}\left(\varepsilon^{\mathrm{KS}}_{a_1} - \varepsilon^{\mathrm{KS}}_{i_1}\right) + 2(i_1a_1|b_1j_1) + (i_1a_1|w_{\mathrm{xc}}|b_1j_1) + (i_1a_1|w^{\mathrm{I}}_{\mathrm{xck}}|b_1j_1), \quad (71)$$

$${}^{\mathrm{I,I}}B_{i_1a_1,j_1b_1} = 2(i_1a_1|j_1b_1) + (i_1a_1|w_{\mathrm{xc}}|j_1b_1) + (i_1a_1|w^{\mathrm{I}}_{\mathrm{xck}}|j_1b_1). \quad (72)$$

These diagonal blocks are coupled by the pure interaction block

$${}^{\mathrm{I,II}}A^{\mathrm{int}}_{i_1a_1,j_{\mathrm{II}}b_{\mathrm{II}}} = {}^{\mathrm{I,II}}B^{\mathrm{int}}_{i_1a_1,j_{\mathrm{II}}b_{\mathrm{II}}} = 2(i_1a_1|j_{\mathrm{II}}b_{\mathrm{II}})_{\mathrm{int}} + (i_1a_1|w_{\mathrm{xck}}|j_{\mathrm{II}}b_{\mathrm{II}}). \quad (73)$$

Note that we use the subscript "int" on the right-hand side to emphasize that this Coulomb term arises due to the interaction energy expression in Eq. (63) which goes back to Eq. (6).

Note also that the orbital energies include both the effective and the embedding potential of the subsystem. For DFT-in-DFT, it is therefore convenient to add the non-additive exchange-correlation contribution (contained in the embedding potential) to the subsystem exchange-correlation contribution (which then becomes exchange-correlation potential

for the total density). This gives as final equation the simple expression,

$$ {}^{\mathrm{I,I}}A_{i_1a_1,j_1b_1} = \delta_{i_1j_1}\delta_{a_1b_1}\left(\varepsilon^{\mathrm{KS}}_{a_1} - \varepsilon^{\mathrm{KS}}_{i_1}\right) + 2(i_1a_1|b_1j_1) + (i_1a_1|\check{w}^{\mathrm{I}}_{\mathrm{xck}}|b_1j_1), \tag{74} $$

$$ {}^{\mathrm{I,I}}B_{i_1a_1,j_1b_1} = 2(i_1a_1|j_1b_1) + (i_1a_1|\check{w}^{\mathrm{I}}_{\mathrm{xck}}|j_1b_1), \tag{75} $$

where the contribution $\check{w}^{\mathrm{I}}_{\mathrm{xck}}$ is defined as

$$ \check{w}^{\mathrm{I}}_{\mathrm{xck}}(\mathbf{r}', \mathbf{r}'') = w_{\mathrm{xck}}(\mathbf{r}', \mathbf{r}'') - \left.\frac{\delta^2 T_{\mathrm{s}}[\rho]}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r}'')}\right|_{\rho_{\mathrm{I}}}. \tag{76} $$

Up to this point, the subsystem approach enabled the extraction of explicit interaction contributions, but the dimensionality of the problem remains the same compared to a supermolecular treatment. Only in case of uncoupled excitations it is evident that both the response equations and the linear response function become decoupled, leading to smaller dimensions and, thus, significantly reduced computational costs.

As shown by Neugebauer, it is possible to avoid matrices and vectors with supermolecular dimensions in the coupled treatment.[31,32] In a first step, the lowest excitation energies for the different subsystems are calculated. Subsequently, a truncated eigenvalue equation is set up with reduced dimensions for which a reduced number of coupling elements are calculated. Therewith, the frozen-density approach significantly speeds up the calculation of molecular properties, while retaining the accuracy close to a supermolecular calculation.

## V. WFT-IN-DFT RESPONSE THEORY

As discussed in the Introduction, wave-function-based methods present another valid way to obtain the electron density. Following the typical hierarchy in the wave-function ansatz, we start with the variational, single-determinant, Hartree-Fock method. Hartree-Fock should thereby be considered as an approximation and the first step towards coupled-cluster theory.

### A. HF-in-DFT

#### 1. Density parameterization

Analogously to the DFT-in-DFT embedding, we introduce an independent exponential parameterization for each subsystem,

$$ \exp(-\hat{\kappa}_{\mathrm{I}})|0_{\mathrm{HF}}\rangle, \quad (-\hat{\kappa}_{\mathrm{II}})|0_{\mathrm{DFT}}\rangle, \tag{77} $$

so that the total electron density remains the sum of both subsystems,

$$ \rho_{\mathrm{tot}}(\mathbf{r}, t) \overset{\mathrm{FDE}}{=} \rho_{\mathrm{WFT}}(\mathbf{r}, \boldsymbol{\kappa}_{\mathrm{I}}, t) + \rho_{\mathrm{DFT}}(\mathbf{r}, \boldsymbol{\kappa}_{\mathrm{II}}, t). \tag{78} $$

With all variational parameters expressed in terms of orbital rotations, the full quasi-energy expression reads

$$ Q(t) = \langle 0|\exp[-\hat{\kappa}(t)]\left(\hat{H} + \hat{V}^t - i\frac{\partial}{\partial t}\right)\exp[\hat{\kappa}(t)]|0\rangle + Q_{\mathrm{int}}[\rho_{\mathrm{HF}}(t), \rho_{\mathrm{DFT}}(t)] + Q[\rho_{\mathrm{DFT}}(t)]. \tag{79} $$

#### 2. Linear response

The working equations of linear response theory in the HF-in-DFT case are very similar to the DFT-in-DFT case, with differences due to the exact exchange (see, e.g., Ref. 69) appearing only in the diagonal subsystem blocks. Because the treatment of the interaction energy remains identical to DFT, the matrix $\mathbf{F}_{\mathrm{int}}$ can again be partitioned into $\mathbf{A}$ and $\mathbf{B}$ sub-blocks with the expressions for the matrix elements of those sub-blocks being the same as those in Eqs. (70) and (73) (with of course the Hartree-Fock density replacing the DFT density in subsystem I).

The HF-in-DFT derivation thus yields the following one-electron and two-electron contributions for the matrix $\mathbf{A}^{\mathrm{HF}}$,

$$ A^{\mathrm{HF}}_{i_1a_1,j_1b_1} = \delta_{i_1j_1}F^{\mathrm{HF,I}}_{a_1b_1} - \delta_{a_1b_1}F^{\mathrm{HF,I}}_{i_1j_1} + [i_1a_1||b_1j_1] + \delta_{i_1j_1}\langle a_1|v^{\mathrm{I}}_{\mathrm{emb}}|b_1\rangle - \delta_{a_1b_1}\langle i_1|v^{\mathrm{I}}_{\mathrm{emb}}|j_1\rangle + (i_1a_1|w^{\mathrm{I,I}}_{\mathrm{emb}}|b_1j_1), \tag{80} $$

where square brackets are defined as anti-symmetrized spin-free two-electron integrals,

$$ [pq||rs] = 2(pq|rs) - (ps|rq). \tag{81} $$

Collecting all contributions that belong to the Fock matrix, the expressions simplify to

$$ A^{\mathrm{HF}}_{i_1a_1,j_1b_1} = \delta_{i_1j_1}\delta_{a_1b_1}\left(\varepsilon^{\mathrm{HF}}_{a_1} - \varepsilon^{\mathrm{HF}}_{i_1}\right) + [i_1a_1||b_1j_1] + (i_1a_1|w^{\mathrm{I}}_{\mathrm{xck}}|b_1j_1), \tag{82} $$

$$ B^{\mathrm{HF}}_{i_1a_1,j_1b_1} = [i_1a_1||j_1b_1] + (i_1a_1|w^{\mathrm{I}}_{\mathrm{xck}}|j_1b_1), \tag{83} $$

where we have used the diagonal form of the Fock matrix (including the embedding potential) and the short-hand notation adapted from Eq. (65). Note that $\check{w}^{i}_{\mathrm{xck}}$ used in Eq. (74) and $w^{i}_{\mathrm{xck}}$ used in Eq. (82) differ by the presence of the second term in Eq. (65), which is based on the fact that the Hartree-Fock part has no exchange-correlation contribution it could cancel with.

A simple approximation in this WFT-in-DFT approach would be to remove the matrix $\mathbf{B}$ from the environment and coupling blocks. Introducing the 3-component acronym subsystemI-coupling-subsystemII to specify a particular approximation in the coupling block, the TDHF-TDA-TDA model results in the following form of the supermatrix $\mathbf{F}$:

$$ \left(\begin{array}{c|c} \mathbf{F}(\omega_y) & {}^{\mathrm{II,I}}\mathbf{F} \\ \hline {}^{\mathrm{I,II}}\mathbf{F} & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y) \end{array}\right) = \left(\begin{array}{cc|cc} {}^{\mathrm{I,I}}\mathbf{A}^{\mathrm{HF}} & {}^{\mathrm{I,I}}\mathbf{B}^{\mathrm{HF}} & {}^{\mathrm{II,I}}\mathbf{A} & 0 \\ {}^{\mathrm{I,I}}\mathbf{B}^{\mathrm{HF}} & {}^{\mathrm{I,I}}\mathbf{A}^{\mathrm{HF}} & 0 & {}^{\mathrm{II,I}}\mathbf{A} \\ \hline {}^{\mathrm{I,II}}\mathbf{A} & 0 & {}^{\mathrm{II,II}}\mathbf{A}^{\mathrm{KS}} & 0 \\ 0 & {}^{\mathrm{I,II}}\mathbf{A} & 0 & {}^{\mathrm{II,II}}\mathbf{A}^{\mathrm{KS}} \end{array}\right) - \omega_y\left(\begin{array}{cc|c} {}^{\mathrm{I,I}}\mathbf{m} & & 0 \\ & & \\ \hline 0 & & {}^{\mathrm{II,II}}\mathbf{m} \end{array}\right), \tag{84} $$

where we have used superscripts HF and KS to denote the way in which the density is generated. Neglecting also the ${}^{\mathrm{I,I}}\mathbf{B}^{\mathrm{HF}}$ blocks takes us to a CIS-TDA-TDA model. Such simplifications may be interesting when large environments are to be considered.

### B. Coupled-cluster quasi-energy response theory for an isolated system

We now derive the equations for CC-in-DFT frozen-density embedding, as an example, in which we need Lagrangian multipliers to treat a nonvariational wave function. We start by briefly summarizing conventional coupled-cluster response theory to introduce the Lagrangian technique. For a more detailed discussion, see, e.g., Refs. 52, 65, and 70. In the conventional formalism, orbital rotations are not treated explicitly but enter implicitly via the single-excitation amplitudes.[52,70,71]

We note that $t$ denotes the time, whereas $\mathbf{t}$ denote the coupled-cluster amplitudes, which are included in the cluster operator $\hat{T}$, and $\bar{\mathbf{t}}$ Lagrangian multipliers. Since the coupled-cluster energy is not variationally optimized, the time-dependent coupled cluster quasi-energy,

$$Q(\mathbf{t};t) = \langle \mathrm{HF}|\hat{H}^t \exp(\hat{T}(t))|\mathrm{HF}\rangle, \tag{85}$$

is combined with the time-dependent cluster amplitude equations

$$0 = \langle \nu| \exp(-\hat{T}(t)) \left(\hat{H}^t - i\frac{\partial}{\partial t}\right) \exp(\hat{T}(t))|\mathrm{HF}\rangle = \Omega_\nu(\mathbf{t};t) \tag{86}$$

to give a quasi-energy Lagrangian,

$$L_{\mathrm{CC}}(\mathbf{t},\bar{\mathbf{t}};t) = Q(\mathbf{t};t) + \sum_\nu \bar{t}_\nu \Omega_\nu(\mathbf{t};t). \tag{87}$$

In the following, the projection manifold is often not specified to keep the derivation general.

In the presence of a (quasi-) periodic perturbation, the time evolution of the system is completely determined by the condition that the time average of the quasi-energy Lagrangian is stationary with respect to variations of the cluster amplitudes and the Lagrangian multipliers. Requiring stationarity of the Lagrangian with respect to the coupled-cluster amplitudes, equations for the zeroth-order Lagrangian multipliers $\bar{\mathbf{t}}^{(0)}$ are obtained (see also Ref. 12). Including both singles and doubles excitations in the cluster operator yields the CCSD model. Computationally cheaper is the approximated coupled-cluster singles and doubles model, denoted as CC2.[42,72] The CC2 energy and amplitude equations read in the similarity-transformed formulation using the specific projection manifold,[73]

$$E_{\Delta\mathrm{CC}} = \sum_{ia,jb} \left(t_i^a t_j^b + t_{ij}^{ab}\right)[ia||jb], \tag{88}$$

$$\Omega_{ai} = \tilde{F}_{ai} + \sum_{kc} \left(2t_{ik}^{ac} - t_{ki}^{ac}\right)\tilde{F}_{kc} + \sum_{cdk} \left(2t_{ik}^{cd} - t_{ki}^{cd}\right)(kd|\tilde{a}c) - \sum_{dkl} \left(2t_{kl}^{ad} - t_{lk}^{ad}\right)(ld|k\tilde{i}), \tag{89}$$

$$\Omega_{ai,jb} = \sum_c \left(t_{ij}^{cb} F_{ac} + t_{ij}^{ac} F_{bc}\right) - \sum_k \left(t_{kj}^{ab} F_{ik} + t_{ik}^{ab} F_{kj}\right) + (\tilde{i}\tilde{a}|\tilde{j}\tilde{b}). \tag{90}$$

The tilde indicates quantities calculated from $T_1$-transformed molecular orbitals. The CC2 equations are useful to provide an example of CC-in-DFT embedding and can be further approximated to provide a CCS treatment. Note that in order to do this, we do not assume canonical orbitals, because we will in the following consider cases in which the embedding potential is updated relative to the one used in the Hartree-Fock stage of the calculation (in order to be consistent with the coupled cluster density rather than with an input HF or DFT density).

#### 1. Linear response

As for the DFT, we refer to the original references for the details of the unembedded CC2 derivation.[52] Adapting Eq. (32) to the coupled-cluster case and applying the $2n + 2$ rule, the linear response function becomes

$$\langle\langle X;Y\rangle\rangle_{\omega_y}^{\mathrm{CC}} = \frac{1}{2} C^{\pm\omega} P(X(\omega_x), Y(\omega_y)) \times \left[\boldsymbol{\eta}^X + \frac{1}{2}\mathbf{F}\,\mathbf{t}^X(\omega_x)\right]\mathbf{t}^Y(\omega_y), \tag{91}$$

and the solution of the linear response equations yields

$$\mathbf{t}^Y(\omega_y) = -{}^{t\bar{t}}\mathbf{A}(\omega_y)^{-1}\boldsymbol{\xi}^Y, \tag{92}$$

$$\bar{\mathbf{t}}^Y(\omega_y) = -{}^{t\bar{t}}\mathbf{A}(-\omega_y)^{-1}(\boldsymbol{\eta}^Y - {}^{tt}\mathbf{F}\,\mathbf{t}^Y(\omega_y)). \tag{93}$$

Moreover, since $\mathbf{t}^X(\omega_x) = \mathbf{t}^X(-\omega_y) = -{}^{t\bar{t}}\mathbf{A}(-\omega_y)^{-1}\boldsymbol{\xi}^X$, it can be seen that the response function has poles at frequencies corresponding to the eigenvalues of the coupled cluster Jacobian ${}^{t\bar{t}}\mathbf{A}$ (see Table II),

$${}^{t\bar{t}}\mathbf{A}\,\mathbf{R}^f = \omega_f\,\mathbf{R}^f. \tag{94}$$

### C. CC-in-DFT

In order to derive working equations for the CC-in-DFT, the expansion of the interaction term has to be carried out in orders of both the coupled-cluster amplitudes and Lagrangian multipliers,

$$Q_{\mathrm{int}} = E_{\mathrm{int}}^{[0]}[\rho_{\mathrm{CC}}(\mathbf{t},\bar{\mathbf{t}}), \rho_{\mathrm{DFT}}(\boldsymbol{\kappa}_{\mathrm{II}})] + \mathbf{E}_{\mathrm{int};t}^{[1]}\mathbf{t} + \mathbf{E}_{\mathrm{int};\bar{t}}^{[1]}\bar{\mathbf{t}} + \mathbf{E}_{\mathrm{int};\kappa_{\mathrm{II}}}^{[1]}\boldsymbol{\kappa}_{\mathrm{II}} + \cdots. \tag{95}$$

Similar to the SD cases, the amplitudes and multipliers have not yet been expanded in the different orders of the perturbation.

#### 1. The coupled-cluster electron density

The coupled-cluster electron density $\rho_{\mathrm{CC}}(\mathbf{t},\bar{\mathbf{t}})$ is now needed, which can be calculated as an expectation value. One then uses

$$\langle\Lambda(t)| = \langle \mathrm{HF}| + \sum_\nu \bar{t}_\nu(t)\langle\nu|\exp(-\hat{T}(t)) \tag{96}$$

as the bra state and the normal coupled-cluster wave function as the ket state, so that the norm of such a bra-ket is conserved during time evolution,[74]

$$1 = \langle\Lambda(t)|\mathrm{CC}(t)\rangle. \tag{97}$$

TABLE II. Coupled-cluster response quantities.

| Quantity | | Vacuum contr. | Uncoup. emb. contr.[a] | cf. Ref. |
|---|---|---|---|---|
| $\mathbf{t}^{(0)}$ | $0 \stackrel{!}{=} \frac{\partial\{L^{(0)}\}_T}{\partial \bar{\mathbf{t}}_\nu^{(0)}}$ | $0 = \Omega_\nu$ | $+E^{[1]}_{\text{int,WFT};\bar{t};\nu}$ | 11, 12 |
| $\bar{\mathbf{t}}^{(0)}$ | $0 \stackrel{!}{=} \frac{\partial\{L^{(0)}\}_T}{\partial \mathbf{t}_\mu^{(0)}}$ | $0 = \sum_\nu \bar{t}_\nu^{(0)} \frac{\partial \Omega_\nu}{\partial t_\mu^{(0)}} + {}^t\eta_\mu$ | $+E^{[1]}_{\text{int,WFT};t;\mu}$ | 11, 12 |
| ${}^t\boldsymbol{\eta}$ | $\frac{\partial E_{\text{CC}}}{\partial t_\mu^{(0)}}$ | $\langle \text{HF}\vert \exp(-\hat{T}^{(0)})\,[\hat{H}, \tau_\mu]\,\exp(\hat{T}^{(0)})\vert \text{HF}\rangle$ | | 52 |
| $\boldsymbol{\eta}^Y_\mu$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial \epsilon_Y(\omega_y)\partial t_\mu^{(1)}(\omega_y)}$ | $\langle \Lambda\vert[\hat{Y}, \tau_\mu]\vert \text{CC}\rangle$ | | 11, 13, 52 |
| $\mathbf{t}^Y(\omega_y)$ | $0 \stackrel{!}{=} \frac{\partial\{L^{(0)}\}_T}{\partial \bar{\mathbf{t}}_\nu^{(1)}}$ | $0 = ({}^{t\bar{t}}\mathbf{A} - \omega_y\mathbf{1})\mathbf{t}^Y(\omega_y) + \langle \boldsymbol{\nu}\vert\hat{Y}\vert \text{CC}\rangle$ | $+\bar{\mathbf{t}}\,\mathbf{E}^{[2]}_{\text{int,WFT};\bar{t}t}$ | 11, 13[b] |
| $\bar{\mathbf{t}}^Y(\omega_y)$ | | $0 = \bar{\mathbf{t}}^Y(\omega_y)({}^{t\bar{t}}\mathbf{A} + \omega_y\mathbf{1}) + \boldsymbol{\eta}^Y + {}^{tt}\mathbf{F}\,\mathbf{t}^Y(\omega_y)$ | | 11, 13[b] |
| ${}^uF_{\mu\nu}$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial t_\mu^{(1)}(\omega_x)\partial t_\nu^{(1)}(\omega_y)}$ | $\langle \Lambda\vert[[\hat{H}, \tau_\mu], \tau_\nu]\vert \text{CC}\rangle$ | $+E^{[2]}_{\text{int,WFT};tt;\mu\nu}$ | 11, 13 |
| ${}^{t\bar{t}}A_{\mu\nu}(\omega_y)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial t_\mu^{(1)}(\omega_x)\partial \bar{t}_\nu^{(1)}(\omega_y)}$ | $\langle \nu\vert[\hat{H}, \tau_\mu]\vert \text{CC}\rangle$ | $+E^{[2]}_{\text{int,WFT};t\bar{t};\mu\nu}$ | 11, 13 |
| ${}^{\bar{t}\bar{t}}J_{\mu\nu}$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial \bar{t}_\mu^{(1)}(\omega_x)\partial \bar{t}_\nu^{(1)}(\omega_y)}$ | | $E^{[2]}_{\text{int,WFT};\bar{t}\bar{t};\mu\nu}$ | 11, 13 |

[a]Embedding contribution without coupling.
[b]Note the embedding contribution in ${}^{t\bar{t}}\mathbf{A}$ and ${}^{tt}\mathbf{F}$.

The (time-dependent) coupled-cluster electron density can thus be calculated as the expectation value of the electron density operator,

$$\rho(\mathbf{r};t) = \langle \Lambda(t)\vert\hat{\rho}(\mathbf{r})\vert \text{CC}(t)\rangle = \sum_{pq} \phi_p(\mathbf{r})\phi_q(\mathbf{r}) D^\Lambda_{pq}(t), \quad (98)$$

where $\mathbf{D}^\Lambda$ is the one-electron coupled-cluster density matrix.[75] The expansion of $\rho(\mathbf{r};t)$ in orders of perturbation is then carried out by expressing $D^\Lambda_{pq}(t)$ in orders of perturbation,

$$D^{\Lambda(0)}_{pq}(0) = \langle \Lambda\vert E_{pq}\vert \text{CC}\rangle, \quad (99)$$

$$D^{\Lambda(1)}_{pq}(t) = \langle \Lambda\vert[E_{pq}, \hat{T}^{(1)}(t)]\vert \text{CC}\rangle + \sum_\nu \bar{t}^{(1)}_\nu(t)\langle \nu\vert \exp(-\hat{T}^{(0)})E_{pq}\vert \text{CC}\rangle, \quad (100)$$

$$D^{\Lambda(2)}_{pq}(t) = \langle \Lambda\vert[E_{pq}, \hat{T}^{(2)}(t)] + \frac{1}{2}[[E_{pq}, \hat{T}^{(1)}(t)], \hat{T}^{(1)}(t)]\vert \text{CC}\rangle + \sum_\nu \left[\bar{t}^{(1)}_\nu(t)\langle \nu\vert \exp(-\hat{T}^{(0)})[E_{pq}, \hat{T}^{(1)}(t)]\vert \text{CC}\rangle + \bar{t}^{(2)}_\nu(t)\langle \nu\vert \exp(-\hat{T}^{(0)})E_{pq}\vert \text{CC}\rangle\right], \quad (101)$$

and so on. Explicit expressions for $\mathbf{D}^{\Lambda(0)}$ can be found in Ref. 73, for example.

### 2. Energy expression

As an example, we list the CC2-in-DFT ground state contributions. The total energy of the CC-in-DFT approach is calculated from the zeroth-order Lagrangian,

$$L^{(0)} = L^{(0)}_{\text{CC}} + Q^{(0)}_{\text{int}} + Q^{(0)}_{\text{DFT}}, \quad (102)$$

which reads explicitly

$$L^{(0)} = \langle 0_{\text{HF}}\vert\hat{H}_{\text{I}}\vert 0_{\text{HF}}\rangle + E^{[0]}_{\Delta\text{CC}} + E^{[0]}_{\text{int}} + E^{[0]}_{\text{DFT}}. \quad (103)$$

Similar to HF-in-DFT, the Hartree-Fock energy contribution denotes the expectation value of the Hartree-Fock wave function over the Hamiltonian of subsystem I *without* any explicit embedding contributions, but with orbitals obtained using the embedding potential in the Hartree-Fock equations. The contributions to the amplitude equations are obtained as derivatives of the Lagrangian with respect to the multipliers,

$$0 = \frac{\partial L^{(0)}}{\partial \bar{t}^{(0)}_\nu} = \Omega_\nu + \frac{\partial Q^{(0)}_{\text{int}}}{\partial \bar{t}^{(0)}_\nu} = \Omega_\nu + E^{[1]}_{\text{int};\bar{t};\nu}, \quad (104)$$

compare also Table II. This yields additional embedding contributions to the normal quantities (see also the discussion in Ref. 76),

$$\Omega_{ia} \leftarrow \langle \tilde{a}\vert v^{\text{I}}_{\text{emb}}\vert \tilde{i}\rangle + \sum_{kc}\left(2t^{ac}_{ik} - t^{ac}_{ki}\right)\langle k\vert v^{\text{I}}_{\text{emb}}\vert c\rangle, \quad (105)$$

$$\Omega_{ia,jb} \leftarrow \sum_c \left(t^{cb}_{ij}\langle a\vert v^{\text{I}}_{\text{emb}}\vert c\rangle + t^{ac}_{ij}\langle b\vert v^{\text{I}}_{\text{emb}}\vert c\rangle\right) - \sum_k \left(t^{ab}_{kj}\langle i\vert v^{\text{I}}_{\text{emb}}\vert k\rangle + t^{ab}_{ik}\langle k\vert v^{\text{I}}_{\text{emb}}\vert j\rangle\right). \quad (106)$$

Note that $v^{\text{I}}_{\text{emb}}$ itself is not calculated from $\hat{T}_1$-transformed orbitals and there is no tilde on the second contribution to the singles amplitude equations. This is in agreement with the conventional coupled cluster in which effectively only the two-electron contribution of the latter term is $\hat{T}_1$-transformed. Equations (103)–(106) correspond to the treatment in Ref. 50 in which the embedding potential was included in the Fock matrix elements and no update of the density and the embedding potential was carried out.

There are different strategies possible to achieve full self-consistency for amplitudes and multipliers in case of CC-in-DFT. The simplest is to start from a converged DFT guess for both subsystem densities,[50] and correct for differences between the calculated CC density and the DFT density (note that this difference only arises in approximate theory, in exact theory both densities would be identical) by carrying out the following procedure:

1. Determine the HF and CC parameters and the Lagrange multipliers for the WFT subsystem with a fixed embedding potential.
2. Calculate the coupled-cluster density and correct the embedding potential for difference between the actual and input active density.
3. If not converged, go back to step 1 and update all parameters using the updated potential.

Due to the high computational costs, such a fully converged procedure is probably not worthwhile to pursue, but it may be desirable to correct in case large differences are found between the input (DFT) density and the calculated WFT density. This may both be due to differences in the formalism as well as differences in the basis set that is applied in both calculations (in case the DFT density is obtained using a different program).

### 3. Linear response

In order to calculate the response contributions arising from the interaction energy, $Q_{\rm int}$ in Eq. (95), we need to expand and sort with respect to orders of the perturbation. Relevant for determining $\{^{2n+1}Q^{(2)}_{\rm int}\}_T$ is, e.g., the second derivative with respect to amplitudes and multipliers,

$$E^{[2]}_{{\rm int};t\bar{t};\mu,\nu} = \iint w^{\rm I,I}_{\rm emb}(\mathbf{r}',\mathbf{r}'')\left(\frac{\partial\rho^{(1)}(\mathbf{r}')}{\partial t^{(1)}_\mu}\right)\left(\frac{\partial\rho^{(1)}(\mathbf{r}'')}{\partial \bar{t}^{(1)}_\nu}\right)\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}'' + \int v^{\rm I}_{\rm emb}(\mathbf{r}')\left(\frac{\partial^2\rho^{(2)}(\mathbf{r}')}{\partial t^{(1)}_\mu \partial \bar{t}^{(1)}_\nu}\right)\mathrm{d}\mathbf{r}' \tag{107}$$

$$= \sum_{x_{\rm I}y_{\rm I},s_{\rm I}u_{\rm I}} {}^{t}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu}\,{}^{\bar{t}}\Gamma^{\rm CC}_{s_{\rm I}u_{\rm I};\nu}(x_{\rm I}y_{\rm I}|w^{\rm I,I}_{\rm emb}|s_{\rm I}u_{\rm I}) + \sum_{x_{\rm I}y_{\rm I}} {}^{t\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu\nu}\langle x_{\rm I}|v^{\rm I}_{\rm emb}|y_{\rm I}\rangle. \tag{108}$$

Using the expansion of the electron density, expressions for the intermediate densities are obtained,

$${}^{t}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu} = \left(\frac{\partial D^{\Lambda\,(1)}_{x_{\rm I}y_{\rm I}}}{\partial t^{(1)}_\mu}\right) = \langle\Lambda|[E_{x_{\rm I}y_{\rm I}},\tau_\mu]|{\rm CC}\rangle, \tag{109}$$

$${}^{\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\nu} = \left(\frac{\partial D^{\Lambda\,(1)}_{x_{\rm I}y_{\rm I}}}{\partial \bar{t}^{(1)}_\nu}\right) = \langle\nu|\exp(-T)E_{x_{\rm I}y_{\rm I}}|{\rm CC}\rangle, \tag{110}$$

$${}^{t\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu\nu} = \left(\frac{\partial^2 D^{\Lambda\,(2)}_{x_{\rm I}y_{\rm I}}}{\partial t^{(1)}_\mu \partial \bar{t}^{(1)}_\nu}\right) = \langle\nu|\exp(-T)[E_{x_{\rm I}y_{\rm I}},\tau_\mu]|{\rm CC}\rangle, \tag{111}$$

where the expressions for $D^{\Lambda\,(n)}_{x_{\rm I}y_{\rm I}}$ arise from Eqs. (98)–(101). Furthermore, the following densities are needed:

$${}^{tt}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu\nu} = \left(\frac{\partial^2 D^{\Lambda\,(2)}_{x_{\rm I}y_{\rm I}}}{\partial t^{(1)}_\mu \partial t^{(1)}_\nu}\right) = \langle\Lambda|[[E_{x_{\rm I}y_{\rm I}},\tau_\mu],\tau_\nu]|{\rm CC}\rangle, \tag{112}$$

$${}^{\bar{t}\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu\nu} = \left(\frac{\partial^2 D^{\Lambda\,(2)}_{x_{\rm I}y_{\rm I}}}{\partial \bar{t}^{(1)}_\mu \partial \bar{t}^{(1)}_\nu}\right) = 0, \tag{113}$$

to express

$${}^{t\bar{t}}A^{\rm int}_{\mu,\nu} = \sum_{x_{\rm I}y_{\rm I},s_{\rm I}u_{\rm I}} {}^{t}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu}\,{}^{\bar{t}}\Gamma^{\rm CC}_{s_{\rm I}u_{\rm I};\nu}(x_{\rm I}y_{\rm I}|w^{\rm I,I}_{\rm emb}|s_{\rm I}u_{\rm I}) + \sum_{x_{\rm I}y_{\rm I}} {}^{t\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu\nu}\langle x_{\rm I}|v^{\rm I}_{\rm emb}|y_{\rm I}\rangle, \tag{114}$$

$${}^{tt}F^{\rm int}_{\mu,\nu} = \sum_{x_{\rm I}y_{\rm I},s_{\rm I}u_{\rm I}} {}^{t}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu}\,{}^{t}\Gamma^{\rm CC}_{s_{\rm I}u_{\rm I};\nu}(x_{\rm I}y_{\rm I}|w^{\rm I,I}_{\rm emb}|s_{\rm I}u_{\rm I}) + \sum_{x_{\rm I}y_{\rm I}} {}^{tt}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu\nu}\langle x_{\rm I}|v^{\rm I}_{\rm emb}|y_{\rm I}\rangle, \tag{115}$$

$${}^{\bar{t}\bar{t}}J^{\rm int}_{\mu,\nu} = \sum_{x_{\rm I}y_{\rm I},s_{\rm I}u_{\rm I}} {}^{\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu}\,{}^{\bar{t}}\Gamma^{\rm CC}_{s_{\rm I}u_{\rm I};\nu}(x_{\rm I}y_{\rm I}|w^{\rm I,I}_{\rm emb}|s_{\rm I}u_{\rm I}). \tag{116}$$

In order to calculate the contributions to the total linear response function, the interaction energy has to be expanded not only in orders of the amplitudes and multipliers, but also in orders of the orbital rotation parameters of the subsystem II. Therefore, for the elements of the coupling matrix, expressions analogous to Eq. (62) are obtained, but now with the appropriate auxiliary coupled-cluster densities $\Gamma^{\rm CC}$ replacing $\Gamma^{\rm SD}$ for subsystem I,

$${}^{\rm I,II}A_{\mu,r_{\rm II}s_{\rm II}} = \sum_{x_{\rm I}y_{\rm I},t_{\rm II}u_{\rm II}} {}^{\bar{t}}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu_{\rm I}}\,\Gamma^{\rm KS}_{t_{\rm II}u_{\rm II};r_{\rm II}s_{\rm II}}(x_{\rm I}y_{\rm I}|w^{\rm I,II}_{\rm emb}|t_{\rm II}u_{\rm II}), \tag{117}$$

$${}^{\rm I,II}F_{\mu,r_{\rm II}s_{\rm II}} = \sum_{x_{\rm I}y_{\rm I},t_{\rm II}u_{\rm II}} {}^{t}\Gamma^{\rm CC}_{x_{\rm I}y_{\rm I};\mu_{\rm I}}\,\Gamma^{\rm KS}_{t_{\rm II}u_{\rm II};r_{\rm II}s_{\rm II}}(x_{\rm I}y_{\rm I}|w^{\rm I,II}_{\rm emb}|t_{\rm II}u_{\rm II}). \tag{118}$$

The superscripts of the embedding kernel indicate that the Coulomb contribution is present for the inter-subsystem contributions.

The approximate uncoupled linear response function is obtained from Eqs. (28) and (30),

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = \frac{1}{2}C^{\pm\omega}\,P(X(\omega_x),Y(\omega_y)) \times \left\{\frac{1}{2}\,{}^{\bar{t}\bar{t}}\mathbf{J}\,\bar{\mathbf{t}}^X(\omega_x)\,\bar{\mathbf{t}}^Y(\omega_y) + \left[\begin{pmatrix}\boldsymbol{\eta}^X\\ {}^{\rm II}\boldsymbol{\eta}^X\end{pmatrix}^T + \frac{1}{2}\begin{pmatrix}\mathbf{t}^X(\omega_x)\\ \boldsymbol{\kappa}^X_{\rm II}(\omega_x)\end{pmatrix}^T\left(\begin{array}{c|c}{}^{tt}\mathbf{F} & 0\\ \hline 0 & {}^{\rm II,II}\mathbf{F}(\omega_y)\end{array}\right)\right] \times \begin{pmatrix}\mathbf{t}^Y(\omega_y)\\ \boldsymbol{\kappa}^Y_{\rm II}(\omega_y)\end{pmatrix}\right\}, \tag{119}$$

and the perturbed parameters are calculated from the decoupled set of linear response equations (cf. Sec. VI),

$$\begin{pmatrix} 0 \\ \hline 0 \\ \hline 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\xi}^Y \\ \hline \boldsymbol{\eta}^Y \\ \hline {}^{\mathrm{II}}\boldsymbol{\eta}^Y \end{pmatrix} + \left(\begin{array}{c|c|c} {}^{t\bar{t}}\mathbf{A} - \omega_y \mathbf{1} & {}^{\bar{t}\bar{t}}\mathbf{J} & 0 \\ \hline {}^{tt}\mathbf{F} & {}^{t\bar{t}}\mathbf{A} + \omega_y \mathbf{1} & 0 \\ \hline 0 & 0 & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y) \end{array}\right) \cdot \begin{pmatrix} \mathbf{t}^Y(\omega_y) \\ \hline \bar{\mathbf{t}}^Y(\omega_y) \\ \hline \boldsymbol{\kappa}^Y_{\mathrm{II}}(\omega_y) \end{pmatrix}. \tag{120}$$

Therefore, the linear response function can trivially be expressed as the sum of the two subsystem contributions.

For coupled response properties, the full response function in Eq. (31) becomes

$$\langle\langle X;Y\rangle\rangle_{\omega_y} = \langle\langle X;Y\rangle\rangle^{\mathrm{CC}}_{\omega_y} + {}^{\mathrm{I,II}}\mathbf{F}\,\mathbf{t}^X(\omega_x)\boldsymbol{\kappa}^Y_{\mathrm{II}}(\omega_y) + \langle\langle X;Y\rangle\rangle^{\mathrm{DFT}}_{\omega_y} \tag{121}$$

and can be calculated after solving for the perturbed amplitudes and multipliers according to Eq. (28),

$$\begin{pmatrix} 0 \\ \hline 0 \\ \hline 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\xi}^Y \\ \hline \boldsymbol{\eta}^Y \\ \hline {}^{\mathrm{II}}\boldsymbol{\eta}^Y \end{pmatrix} + \left(\begin{array}{c|c|c} {}^{t\bar{t}}\mathbf{A} - \omega_y \mathbf{1} & {}^{\bar{t}\bar{t}}\mathbf{J} & {}^{\mathrm{I,II}}\mathbf{A} \\ \hline {}^{tt}\mathbf{F} & {}^{t\bar{t}}\mathbf{A} + \omega_y \mathbf{1} & {}^{\mathrm{I,II}}\mathbf{F} \\ \hline {}^{\mathrm{I,II}}\mathbf{F} & {}^{\mathrm{I,II}}\mathbf{A} & {}^{\mathrm{II,II}}\mathbf{F}(\omega_y) \end{array}\right) \cdot \begin{pmatrix} \mathbf{t}^Y(\omega_y) \\ \hline \bar{\mathbf{t}}^Y(\omega_y) \\ \hline \boldsymbol{\kappa}^Y_{\mathrm{II}}(\omega_y) \end{pmatrix}. \tag{122}$$

In this case, all three parameter responses ($\mathbf{t}^X$, $\bar{\mathbf{t}}^X$, and $\boldsymbol{\kappa}^X$) are coupled. Again, as already discussed for DFT-in-DFT, the computational cost becomes a key issue. Since the coupled-cluster Jacobian is typically already very large, a further increase of the dimension should be avoided. Here, one may first transform to a smaller basis of solutions before considering the coupling between the systems, similar to the strategy employed for the DFT-in-DFT.[31] Additional savings can be obtained by considering approximations in the coupling blocks, e.g., using the interactions' locality.

#### 4. Exemplary working equations

The expressions above are valid for a general truncation level, and it is instructive to consider a few cases for which the actual expressions for the densities $\boldsymbol{\Gamma}^{\mathrm{CC}}$ are rather simple. For CCS, a fair amount of simplifications apply to Eqs. (114)–(118). The ground-state density that is to be used to evaluate the kernel integrals becomes then exactly that of the Hartree-Fock. The auxiliaries are, however, different,

$${}^{t}\Gamma^{\mathrm{CCS}}_{x_1y_1;i_1a_1} = \left(\frac{\partial D^{\Lambda\,(1)}_{x_1y_1}}{\partial t^{a_1\,(1)}_{i_1}}\right) = \langle\Lambda|[E_{x_1y_1}, E_{a_1i_1}]|\mathrm{CC}\rangle = \delta_{i_1x_1}\delta_{a_1y_1}, \tag{123}$$

$${}^{\bar{t}}\Gamma^{\mathrm{CCS}}_{x_1y_1;j_1b_1} = \left(\frac{\partial D^{\Lambda\,(1)}_{x_1y_1}}{\partial \bar{t}^{b_1\,(1)}_{j_1}}\right) = \langle^{\bar{b}_1}_{j_1}| \exp(-T) E_{x_1y_1}|\mathrm{CC}\rangle = \delta_{j_1x_1}\delta_{b_1y_1}, \tag{124}$$

$${}^{t\bar{t}}\Gamma^{\mathrm{CCS}}_{x_1y_1;i_1a_1,j_1b_1} = \left(\frac{\partial^2 D^{\Lambda\,(2)}_{x_1y_1}}{\partial t^{a_1\,(1)}_{i_1}\partial \bar{t}^{b_1\,(1)}_{j_1}}\right) = \langle^{\bar{b}_1}_{j_1}| \exp(-T)[E_{x_1y_1}, E_{a_1i_1}]|\mathrm{CC}\rangle = \delta_{i_1j_1}\delta_{x_1a_1}\delta_{y_1b_1} - \delta_{a_1b_1}\delta_{x_1i_1}\delta_{y_1j_1}. \tag{125}$$

This leads to the potential and kernel contribution to the coupled-cluster Jacobian,

$$\begin{aligned} {}^{t\bar{t}}A^{\mathrm{CCS,\,int}}_{i_1a_1,j_1b_1} &= \sum_{x_1y_1,s_1u_1} {}^{t}\Gamma^{\mathrm{CCS}}_{x_1y_1;j_1b_1}\,{}^{\bar{t}}\Gamma^{\mathrm{CCS}}_{s_1u_1;i_1a_1}(x_1y_1|w^{\mathrm{I,I}}_{\mathrm{emb}}|s_1u_1) \\ &\quad + \sum_{x_1y_1} {}^{t\bar{t}}\Gamma^{\mathrm{CCS}}_{x_1y_1;i_1a_1,j_1b_1}\langle x_1|v^{\mathrm{I}}_{\mathrm{emb}}|y_1\rangle \\ &= \delta_{i_1j_1}\langle a_1|v^{\mathrm{I}}_{\mathrm{emb}}|b_1\rangle - \delta_{ab}\langle i_1|v^{\mathrm{I}}_{\mathrm{emb}}|j_1\rangle + (i_1a_1|w^{\mathrm{I}}_{\mathrm{xck}}|b_1j_1) \end{aligned} \tag{126}$$

that are equivalent to the elements of the Hessian in the case of CIS-in-DFT, compare Eq. (82) and, if neglecting the ${}^{\mathrm{I,I}}\mathbf{B}^{\mathrm{HF}}$ blocks, Eq. (84). Coupling elements are obtained similarly,

$${}^{\mathrm{I,II}}A_{i_1a_1,j_{\mathrm{II}}b_{\mathrm{II}}} = \sum_{x_1y_1,t_{\mathrm{II}}u_{\mathrm{II}}} {}^{\bar{t}}\Gamma^{\mathrm{CC}}_{x_1y_1;i_1a_1}\,\Gamma^{\mathrm{KS}}_{t_{\mathrm{II}}u_{\mathrm{II}};j_{\mathrm{II}}b_{\mathrm{II}}}(x_1y_1|w^{\mathrm{I,II}}_{\mathrm{emb}}|t_{\mathrm{II}}u_{\mathrm{II}}) \tag{127}$$

$$= 2(i_1a_1|b_{\mathrm{II}}j_{\mathrm{II}})_{\mathrm{int}} + (i_1a_1|w_{\mathrm{xck}}|b_{\mathrm{II}}j_{\mathrm{II}}). \tag{128}$$

This example is of course only presented to illustrate the general theory, typically WFT descriptions will be aimed at improving upon a DFT description by using a method of at least CC2 quality.

## VI. RELATION TO QM/MM METHODS

The discussion above has mostly been concerned with the formalism and the connection to prior work within the context of DFT-in-DFT or WFT-in-DFT frozen-density embedding. Nevertheless, for the "embedding" mode of the formalism it is illustrative also to make a connection to other related approaches such as the SD/molecular mechanics (SD/MM) (Ref. 77) and coupled-cluster/molecular mechanics (CC/MM) (Refs. 10–13) methods, and the more recent polarizable embedding (PE) approaches, PE-SD (Ref. 78) and PE-CC,[14] proposed by Christiansen, Mikkelsen, Kongsted, and co-workers.

Conceptually similar is the treatment of Coulombic interactions. The major difference lies in the continuous electron density in case of FDE, whereas MM and PE use a discrete multipole expansion. In both cases, the quadratic density dependence in the energy transfers to a linear dependence in the embedding potential. The differences come from the exchange-correlation and kinetic energy contributions

that provide the Pauli repulsion that is lacking in the classical approaches. These also lead to a nonlinear dependency on changes in the (partitioned) density due to the perturbing field.

Despite these differences in the physical content in the FDE and CC/MM approaches, the working equations exhibit a number of similarities. An important example is, for instance, the matrix **J** which is absent for the vacuum case. For both FDE and CC/MM or PE-CC, **J** is responsible for coupling the response equations determining the perturbed coupled-cluster amplitudes and Lagrangian multipliers.[11,14] In our formalism, **J** describes changes in the response of the coupled-cluster system due to the environment that are caused by changes in the (intra-subsystem) non-additive exchange-correlation and kinetic energy contributions, while explicit "inter-subsystem" coupling effects are accounted for by the off-diagonal blocks of the matrices **A** and **F**. In the CC/MM or PE-CC, the "inter-subsystem" contributions are also present in **J**, since the interactions with the environment (including its response) are expressed as "effective" contributions to the QM part.

There is also a relation with respect to pole and residue analysis of the response function that is discussed in the context of classical embedding.[10–12] The specific coupling of the amplitudes and multipliers leads to poles of higher order compared to exact theory in vacuum. For CC/MM, it has been proposed to ignore the coupling of the $\mathbf{t}$ and $\bar{\mathbf{t}}$ responses so that the same formal expressions for transition properties are obtained compared to the vacuum case, with the difference that they include the embedding contributions. A similar approach could be followed here.

## VII. NUMERICAL EXAMPLE

Although the present article is mainly concerned with the theoretical aspects of FDE response formalism, a pilot application is presented to assess the importance of the various environment contributions. The model system is a solvated water molecule as investigated by Jacob *et al.* in the context of evaluating the performance of FDE for ground-state and response properties vs. the discrete reaction field method.[34]

The formalism presented is implemented in a library currently interfaced to a development version of the DIRAC program package,[79] following up on previous work,[50] and restricting the discussion to HF-in-DFT response. The details of the implementation and its use in connection to electron correlation methodologies will be addressed in subsequent publications. In our calculations we employ the PyADF scripting framework[80] in order to perform FDE calculations with the ADF code[33,81] (using the PBE functional for $E_{\mathrm{xc}}$ and $E_{\mathrm{xc}}^{\mathrm{nadd}}$, and PW91k for $T_{\mathrm{s}}^{\mathrm{nadd}}$, and a TZ2P basis set augmented with diffuse functions). Following one of the strategies discussed in Ref. 34, the frozen density is constructed as a superposition of fragment densities obtained for an isolated molecule. The DIRAC calculations are performed employing the aug-cc-pVTZ basis in combination with the Levy-Leblond (non-relativistic) Hamiltonian.

Our results for the three lowest singlet excitation energies are shown in Table III. We observe that our calculations and those reported in Ref. 34 yield similar trends, although quantitative differences occur. For instance, in both cases the shifts in the energies due to solvation are positive and show little variation, but the HF-in-DFT values are roughly two-thirds of those obtained by Jacob *et al.* As for the response (kernel) contributions, we can see that for the lowest singlet state both calculations yield similar results, namely, 0.06 eV for HF-in-DFT and 0.07 eV for DFT-in-DFT.

TABLE III. TDHF excitation energies (in eV) for the first three singlet states of a water molecule, isolated ($E_{\mathrm{iso}}$) and solvated ($E_{\mathrm{sol}}$) by 127 water molecules employing FDE. The FDE corrections are further subdivided into "diagonal"' and "response," i.e., arising from the potential and the kernel contributions, see Eqs. (80) and (83), denoted by $\Delta E_{\mathrm{env}}^{\mathrm{diag}}$ and $\Delta E_{\mathrm{env}}^{\mathrm{resp}}$, respectively.

| State | $E_{\mathrm{iso}}$ (eV) | $E_{\mathrm{sol}}$ (eV) | $\Delta E_{\mathrm{env}}^{\mathrm{diag}}$ (eV) | $\Delta E_{\mathrm{env}}^{\mathrm{resp}}$ (eV) |
|---|---|---|---|---|
| 1 | 8.65 | 9.24 | 0.53 | 0.06 |
| 2 | 10.33 | 10.85 | 0.47 | 0.05 |
| 3 | 10.94 | 11.47 | 0.47 | 0.06 |

Similar behavior is seen for the dipole moments. Using HF-in-DFT, we obtain a shift of +0.57 D from the value for the isolated molecule (1.98 D), whereas the DFT calculations in Ref. 34 show a shift of +0.65 D from the value of the isolated molecule (1.80 D).

## VIII. CONCLUSIONS

We have presented a formalism suitable for calculating the time-dependent molecular properties within a subsystem embedding framework, the key aspects of which are: First, the definition of a time-dependent Lagrangian expression that connects the energies of the isolated subsystems and their interaction energy. Second, the use of the time-averaged quasi-energy formalism in order to identify the molecular properties with the (time-averaged) derivatives of the Lagrangian with respect to the perturbing fields' strengths.[52] As usual, the time-independent properties are also accessible, as a special (zero-frequency) case.

The crucial ansatz in our formalism is the expression of the interaction contribution to the Lagrangian in a purely DFT fashion, that is, as a functional of the (time-dependent) electron density for the total system. In addition, we consider the number of particles in each subsystem as fixed, although for subsystems treated by DFT it may be possible to relax this constraint, see, e.g., Refs. 47, 55, and 56. However, these features provide several advantages: there is no double counting of electron correlation; the total density can be expressed as the sum of overlapping subsystem densities, and variational and non-variational WFT methods can be treated on the same footing. Furthermore, it offers the pathway to an efficient description with a large number of subsystems.[82] The calculation of the interaction contributions is straightforward, being limited primarily by the accuracy of the approximate exchange-correlation and kinetic energy functionals used to calculate the non-additive kinetic and exchange-correlation contributions.

While we have restricted the discussion to coupled-cluster as an example of non-variational wave functions and to Hartree-Fock as a simple example of variational methods, we note that the methodology presented can be applied to other non-variational methods such as MP2 as well as to variational methods such as MCSCF. Since the interaction contribution to the quasi-energy Lagrangian is a functional of the (total) electron density, we only require a formulation of the time-dependent electron density using the method of choice. We believe that for time-independent properties, the simplest WFT-in-DFT model in practice should be MP2-in-DFT as simpler models will not improve upon the DFT description. For time-dependent properties, we propose the CC2-in-DFT—although the pole structure is more complicated to the vacuum case, approximations offer the possibility to correct this deficiency, whereas MP2 itself exhibits inherently a wrong pole structure.

In the preceding discussion, we have hinted at some strategies to take advantage of the subsystem formulation in the calculation of the response parameters and (coupled) excitation energies. In the future work, we plan to investigate this further and implement efficient approximate embedding treatments.

## ACKNOWLEDGMENTS

S.H. has been supported by the European Commission under a Marie-Curie Intra-European Fellowship (Contract No. PIEF-GA-2010-274224). L.V. has been supported by a VICI grant by the Netherlands Organisation for Scientific Research (NWO).

[1]C. Hättig and A. Köhn, J. Chem. Phys. **117**, 6939 (2002).
[2]D. Kats, T. Korona, and M. Schütz, J. Chem. Phys. **125**, 104106 (2006).
[3]D. Kats, T. Korona, and M. Schütz, J. Chem. Phys. **127**, 064107 (2007).
[4]A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[5]P. Sherwood, in *Modern Methods and Algorithms of Quantum Chemistry*, NIC Series Vol. 1, edited by J. Grotendorst (John von Neumann Institute for Computing, Jülich, 2000), pp. 257–277.
[6]J. Gao, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd, (VCH, New York, 1995), Vol. 7, pp. 119–185.
[7]H. M. Senn and W. Thiel, Angew. Chem. **48**, 1198 (2009).
[8]S. Humbel, S. Sieber, and K. Morokuma, J. Chem. Phys. **105**, 1959 (1996).
[9]M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, J. Phys. Chem. **100**, 19357 (1996).
[10]O. Christiansen and K. V. Mikkelsen, J. Chem. Phys. **110**, 8348 (1999).
[11]J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, Mol. Phys. **100**, 1813 (2002).
[12]J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, J. Phys. Chem. A **107**, 2578 (2003).
[13]J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, J. Chem. Phys. **118**, 1620 (2003).
[14]K. Sneskov, T. Schwabe, J. Kongsted, and O. Christiansen, J. Chem. Phys. **134**, 104108 (2011).
[15]T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).
[16]G. Senatore and K. Subbaswamy, Phys. Rev. B **34**, 5754 (1986).
[17]P. Cortona, Phys. Rev. B **46**, 2008 (1992).
[18]T. A. Wesolowski, J. Phys. A **36**, 10607 (2003).
[19]O. V. Gritsenko and L. Visscher, Phys. Rev. A **82**, 032519 (2010).
[20]C. R. Jacob, S. M. Beyhan, and L. Visscher, J. Chem. Phys. **126**, 234116 (2007).
[21]A. Savin and T. A. Wesolowski, in *Progress in Theoretical Chemistry and Physics*, Advances in the Theory of Atomic and Molecular Systems Vol. 19, edited by P. Piecuch, J. Maruani, G. Delgado-Barrio, and S. Wilson (Springer, New York, 2010), pp. 311-326.
[22]J. D. Goodpaster, N. Ananth, F. R. Manby, and I. T. F. Miller, J. Chem. Phys. **133**, 084103 (2010).
[23]T. A. Wesolowski, H. Chermette, and J. Weber, J. Chem. Phys. **105**, 9182 (1996).
[24]T. A. Wesolowski and J. Weber, Int. J. Quantum Chem. **61**, 303 (1997).
[25]A. W. Götz, S. M. Beyhan, and L. Visscher, J. Chem. Theory Comput. **5**, 3161 (2009).
[26]S. Fux, K. Kiewisch, C. R. Jacob, J. Neugebauer, and M. Reiher, Chem. Phys. Lett. **461**, 353 (2008).
[27]S. M. Beyhan, A. W. Gotz, C. R. Jacob, and L. Visscher, J. Chem. Phys. **132**, 044114 (2010).
[28]M. Iannuzzi, B. Kirchner, and J. Hutter, Chem. Phys. Lett. **421**, 16 (2006).
[29]C. R. Jacob, J. Neugebauer, and L. Visscher, J. Comput. Chem. **29**, 1011 (2008).
[30]M. E. Casida and T. A. Wesolowski, Int. J. Quantum Chem. **96**, 577 (2004).
[31]J. Neugebauer, J. Chem. Phys. **126**, 134116 (2007).
[32]J. Neugebauer, J. Chem. Phys. **131**, 084104 (2009).
[33]G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).
[34]C. R. Jacob, J. Neugebauer, L. Jensen, and L. Visscher, Phys. Chem. Chem. Phys. **8**, 2349 (2006).
[35]J. Neugebauer, Phys. Rep. **489**, 1 (2010).
[36]J. Neugebauer, M. J. Louwerse, E. J. Baerends, and T. A. Wesolowski, J. Chem. Phys. **122**, 094115 (2005).
[37]J. Neugebauer, Ch. R. Jacob, T. A. Wesolowski, and E. J. Baerends, J. Phys. Chem. A **109**, 7805 (2005).
[38]J. M. García-Lastra, T. Wesolowski, M. T. Barriuso, J. A. Aramburu, and M. Moreno, J. Phys.: Condens. Matter **18**, 1519 (2006).
[39]M. Zbiri, C. A. Daul, and T. A. Wesolowski, J. Chem. Theory Comput. **2**, 1106 (2006).
[40]R. E. Bulo, C. R. Jacob, and L. Visscher, J. Phys. Chem. A **112**, 2640 (2008).
[41]J. Neugebauer and E. J. Baerends, J. Phys. Chem. A **110**, 8786 (2006).
[42]O. Christiansen, H. Koch, and P. Jørgensen, Chem. Phys. Lett. **243**, 409 (1995).
[43]N. Govind, Y. A. Wang, A. J.R. da Silva, and E. A. Carter, Chem. Phys. Lett. **295**, 129 (1998).
[44]N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **110**, 7677 (1999).
[45]P. Huang and E. A. Carter, J. Chem. Phys. **125**, 084102 (2006).
[46]S. Sharifzadeh, P. Huang, and E. A. Carter, Chem. Phys. Lett. **470**, 347 (2009).
[47]C. Huang, M. Pavone, and E. A. Carter, J. Chem. Phys. **134**, 154110 (2011).
[48]T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, Phys. Rev. Lett. **86**, 5954 (2001).
[49]T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **116**, 42 (2002).
[50]A. S. P. Gomes, C. R. Jacob, and L. Visscher, Phys. Chem. Chem. Phys. **10**, 5353 (2008).
[51]Y. G. Khait and M. R. Hoffmann, J. Chem. Phys. **133**, 044107 (2010).
[52]O. Christiansen, P. Jørgensen, and C. Hättig, Int. J. Quantum Chem. **68**, 1 (1998).
[53]P. Sałek, T. Helgaker, and T. Saue, Chem. Phys. **311**, 187 (2005).
[54]N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **110**, 7677 (1999).
[55]P. Elliott, M. H. Cohen, A. Wasserman, and K. Burke, J. Chem. Theory Comput. **5**, 827 (2009).
[56]P. Elliott, K. Burke, M. H. Cohen, and A. Wasserman, Phys. Rev. A **82**, 024501 (2010).
[57]C. Huang and E. A. Carter, J. Chem. Phys. **135**, 194104 (2011).
[58]T. A. Wesolowski, J. Chem. Phys. **106**, 8516 (1997).
[59]U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen, and K. Ruud, J. Chem. Theory Comput. **6**, 1971 (2010).
[60]R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
[61]S. Sharifzadeh, P. Huang, and E. A. Carter, Chem. Phys. Lett. **470**, 347 (2009).
[62]T. A. Wesolowski, Phys. Rev. A **77**, 012504 (2008).
[63]F. Aquilante and T. A. Wesolowski, J. Chem. Phys. **135**, 084120 (2011).
[64]O. Christiansen, Theor. Chem. Acc. **116**, 106 (2006).

[65]T. Saue and H. J.A. Jensen, J. Chem. Phys. **118**, 522 (2003).
[66]S. Coriani, S. Høst, B. Jansík, L. Thøgersen, J. Olsen, P. Jørgensen, S. Reine, F. Pawłowski, T. Helgaker, and P. Sałek, J. Chem. Phys. **126**, 154108 (2007).
[67]T. Saue and T. Helgaker, J. Comput. Chem. **23**, 814 (2002).
[68]P. Sałek, O. Vahtras, T. Helgaker, and H. Ågren, J. Chem. Phys. **117**, 9630 (2002).
[69]A. Dreuw and M. Head-Gordon, Chem. Rev. **105**, 4009 (2005).
[70]C. Hättig and P. Jørgensen, J. Chem. Phys. **109**, 2762 (1998).
[71]O. Christiansen, C. Hättig, and J. Gauss, J. Chem. Phys. **109**, 4745 (1998).
[72]O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **103**, 7429 (1995).
[73]T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley VCH, Berlin, 2000).
[74]H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).
[75]H. Koch, H. J. A. Jensen, P. Jørgensen, T. Helgaker, G. E. Scuseria, and H. F. Schaefer, J. Chem. Phys. **92**, 4924 (1990).
[76]O. Christiansen, A. Halkier, H. Koch, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **108**, 2801 (1998).
[77]C. B. Nielsen, O. Christiansen, K. V. Mikkelsen, and J. Kongsted, J. Chem. Phys. **126**, 154112 (2007).
[78]J. M. Olsen, K. Aidas, and J. Kongsted, J. Chem. Theory Comput. **6**, 3721 (2010).
[79]R. Bast, H. J. Aa. Jensen, T. Saue, and L. Visscher, with contributions from V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, T. Fleig, O. Fossgaard, A. S. P. Gomes, T. Helgaker, J. K. Lærdahl, J. Henriksson, M. Iliaš, Ch. R. Jacob, S. Knecht, C. V. Larsen, H. S. Nataraj, P. Norman, G. Olejniczak, J. Olsen, J. K. Pedersen, M. Pernpointner, K. Ruud, P. Sałek, B. Schimmelpfennig, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, S. Villaume, O. Visser, T. Winther, and S. Yamamoto, DIRAC, a relativistic *ab initio* electronic structure program, Release DIRAC11 (2011), see http://dirac.chem.vu.nl.
[80]Ch. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S.P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, J. Comput. Chem. **32**, 2328 (2011).
[81]ADF, Amsterdam density functional program, Theoretical Chemistry, Vrije Universiteit Amsterdam, see http://www.scm.com, 2011.
[82]J. Neugebauer, J. Phys. Chem. B **112**, 2207 (2008).

## D.4 Paper XI

THE JOURNAL OF CHEMICAL PHYSICS **137**, 084308 (2012)

# The electronic spectrum of $CUONg_4$ (Ng = Ne, Ar, Kr, Xe): New insights in the interaction of the CUO molecule with noble gas matrices

Paweł Tecmer,[1] Henk van Lingen,[1] André Severo Pereira Gomes,[2] and Lucas Visscher[1]
[1]*Amsterdam Center for Multiscale Modeling (ACMM), VU University Amsterdam, Theoretical Chemistry, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*
[2]*Laboratoire PhLAM, CNRS, UMR 8523, Université de Lille 1, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France*

(Received 26 June 2012; accepted 23 July 2012; published online 28 August 2012)

The electronic spectrum of the CUO molecule was investigated with the IHFSCC-SD (intermediate Hamiltonian Fock-space coupled cluster with singles and doubles) method and with TD-DFT (time-dependent density functional theory) employing the PBE and PBE0 exchange–correlation functionals. The importance of both spin–orbit coupling and correlation effects on the low-lying excited-states of this molecule are analyzed and discussed. Noble gas matrix effects on the energy ordering and vibrational frequencies of the lowest electronic states of the CUO molecule were investigated with density functional theory (DFT) and TD-DFT in a supermolecular as well as a frozen density embedding (FDE) subsystem approach. This data is used to test the suitability of the FDE approach to model the influence of different matrices on the vertical electronic transitions of this molecule. The most suitable potential was chosen to perform relativistic wave function theory in density functional theory calculations to study the vertical electronic spectra of the CUO and $CUONg_4$ with the IHFSCC-SD method. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4742765]

## I. INTRODUCTION

The chemistry of uranium remains fascinating and challenging, with implications that go beyond the use of fissionable $^{235}$U in nuclear energy, with e.g., a growing interest in possible applications in catalysis.[1,2] This is due to uranium's wide range of oxidation states (+6, +5, +4 and +3) and coordination geometries.[3–6] These features arise from the availability of the spatially rather different 5*f*-, 6*d*-, 7*s*-, and 7*p*-orbitals that are energetically close and do participate in the chemical bonding.

Many uranium complexes contain the uranyl ($UO_2^{2+}$) species, perhaps the best studied of its molecular oxides due to its predominance in the aqueous chemistry of uranium. For this ion (but also for non-oxide isoelectronic species such as NUN) the uranium metal center forms strong triple bonds with the axial ligands.[3,7–14] The high stability and linear structure of these U(VI) species are well explained by the "pushing from below" model proposed by Tatsumi and Hoffmann.[15] In the linear structure the strong interaction between the semi-core uranium 6*p*-orbitals and the oxygen 2*p*-orbitals brings the energy of the $\sigma_u$ HOMO (highest occupied molecular orbital) close to that of the 5*f*-orbitals of the uranium leading to a significant 5*f* participation in this orbital.[16] The accompanying lowering of the energy leaves a large HOMO-LUMO gap for the cation, making uranyl an archetypical closed-shell species. For other isoelectronic species such as $NUO^+$ or CUO the picture is slightly different since the bonds are not equal, even though both can still be regarded as triple bonds. To the diminishing electronegativity of N and C with respect to O there corresponds a decreasing HOMO-LUMO gap, as seen in various theoretical studies[13,17] as a marked decrease of ionization potentials and the lowest electronic excitation energies on going from $UO_2^{2+}$ to NUN, $NUO^+$, and up to CUO.[18,19]

For CUO in particular, the HOMO has a predominantly U–C bonding character with the U–O bond lying at significantly lower energy.[17] Moreover, a point is reached where low-lying triplet excited states are so close to the ground-state that some theoretical approaches predict them to be the electronic ground-state.[12,17] This suggests that even weakly bound equatorial ligands can play an important role in establishing the precise ordering of states in the lower part of the spectrum. This agrees with the findings in the extensive experimental studies of CUO trapped in noble gas matrices that spurred the interest of theoreticians in this molecule. Tague *et al.*[20] excited uranium atoms in excess of CO by laser ablation and found that uranium can insert in the triple CO bond and form the CUO molecule. The experimentally measured U–C and U–O stretching vibrations are 853 $cm^{-1}$ and 804 $cm^{-1}$ in an argon matrix, but when the same experiment was carried out by Zhou *et al.*[21] in a solid neon matrix, they found the U–C and U–O frequencies at 1047 $cm^{-1}$ and 872 $cm^{-1}$. Later on, a similar experiment was carried out for krypton and xenon matrices as well. Again, a large red shift ($\approx$70 and $\approx$200 $cm^{-1}$ for the U–O and U–C stretching modes, respectively) was observed, relative to the data obtained in the neon matrix.[22–24] This large red shift in the vibrational spectra suggest that the ground-state of CUO depends on the noble gas environment, with the weakly interacting neon atoms the system favors a singlet ground-state while the stronger interaction in the heavier noble gas matrices, i.e., argon, krypton, or xenon,[23,25,26] produces a triplet ground-state in which also a nonbonding uranium $5f_\phi$ or $5f_\delta$ is occupied (or a mixture thereof if spin–orbit coupling is accounted for[17]). Direct

validation for this hypothesis could come from electronic spectroscopy on CUO in these matrices. Such data are available for other uranium compounds[27,28] in noble gas matrices, but to the best of our knowledge CUO has not yet been studied in this manner. One thus has to rely on theoretical predictions for the electronic spectra. In the theoretical treatment one needs to accurately describe not only electron correlation but also, because of the heavy atoms involved (U, Xe), both scalar and spin–orbit relativistic effects.

These requirements pose significant challenges[29] to calculations, especially because the electronic states may have a marked multi-reference character. Suitable methodologies are, for instance, the relativistic formulations[30–32] of the Fock-space coupled cluster singles and doubles (FSCC-SD (Refs. 33–35)) method, or the spin–orbit complete active space second-order perturbation theory (SO-CASPT2 (Refs. 36 and 37)). The recent work of Infante *et al.*,[38] who employed the spin–orbit coupling restricted active space second order perturbation (SOC-RASPT2 (Ref. 39)) method to investigate the $UO_2Ar_4$ species, can be considered to be at the limit of what is currently feasible employing wave function based approaches. Since the CUO species has less symmetry, an all-atom treatment is still out of reach. The Fock-space based approaches are less affected by active space limitations, but are computationally more expensive than the PT2 ones and at present are also not able to tackle a full seven atom model (both experimental and theoretical studies suggest coordination of four noble gas atoms in the equatorial plane perpendicular to the CUO unit[23,25,26]).

This leaves as alternative the use of so-called embedding approaches in which interactions from the environment are modeled in a simplified manner.[40] Since both the molecule of interest and the environment consist of neutral molecules, and because the dipole moments of the two states of interest of CUO do not differ much (3.5 D in the singlet state and 2.4 D in the excited triplet state[24]), the most important factor is probably a combination of Pauli repulsion and weak coordination from the equatorial ligands. Most embedding methods are not able to capture such subtle effects as they are usually designed to describe primarily electrostatic interactions. A method that should in principle be able to capture both Pauli repulsion and (to lesser extent) weak coordination bonds is the so-called WFT-in-DFT (wave function theory in density functional theory) embedding, first proposed by Carter and co-workers,[41–43] for studying the electronic spectra of impurities on solids and surfaces.[44,45] In this method, one divides the system of interest in an active part (for which the spectra are to be calculated with a given wave function method) and an environment that is described by DFT. This type of approach has been applied successfully in describing the effect of chloride-actinyl interactions on the *f-f* spectra of the $NpO_2^{2+}$ cation.[46] The WFT-in-DFT frozen density embedding (FDE) scheme is theoretically well-defined and suitable for extension to coupled subsystems[47] and can provide an enormous reduction of both computational cost and the complexity of the data that is to be analyzed. We therefore think it is of interest to use the CUO noble gas interaction as another test case for the feasibility of the approach in describing uranium coordination chemistry. As we will use the method in its uncoupled formulation, we need to ascertain that there is no coupling between the electronic excitations on CUO with those of the environment,[48] something that can be explored at DFT level. In addition, we will use a monomer expansion of the wave function which will also limit the possibilities to describe donation from the noble gas orbitals into the CUO orbitals. This system therefore constitutes a stringent test for the applicability of this simplest form of WFT-in-DFT embedding, but one in which there is a fair chance of success given the fact that DFT-in-DFT typically works well for such weak interactions.[49–54]

In this work our initial goal was to investigate the use of WFT-in-DFT embedding for determining the electronic structure of the model $CUONg_4$ (Ng=Ne, Ar, Kr, and Xe) systems, with particular emphasis on the still debated issue of whether or not there is a change in the nature of the ground-state (between singlet and triplet) for the different noble gas ligands. While carrying out the TD-DFT calculations that were used to validate the embedding procedure, we obtained new insights in the effect of including exact exchange in the TD-DFT description of this molecule that are worthwhile to report as well. These will be discussed in Sec. III A.

## II. COMPUTATIONAL DETAILS

Our investigations required three types of calculations: geometry optimization and calculation of vibrational frequencies for the two possible ground-states of the CUO molecule, TD-DFT calculations of the isolated, embedded, and supermolecular species, and IHFSCC-SD calculation of the isolated and embedded species. For the latter two we also need to discuss the generation of the embedding potential that was used. We will partition the section on computational details accordingly.

### A. Geometry optimization and harmonic frequencies

We performed spin–free DFT geometry optimization and analytical frequency calculations with the TURBOMOLE 5.10 package[55–57] employing the PBE0 exchange–correlation (xc) functional[58,59] in conjunction with the def-TZVP (triple-$\zeta$ valence polarization) basis sets.[60] For the heavy elements (uranium and xenon) small-core effective potentials were utilized.[61,62]

Geometries and structures were obtained for the $^1\Sigma^+$,$^3\Phi$ states of CUO and the $^1A_1$ ($^1\Sigma^+$) and $^3E$ ($^3\Phi$) states for the $CUONg_4$ (Ng=Ne, Ar, Kr, Xe) models. All the DFT calculations were performed in $C_{2v}$ point group symmetry with the multiple grid option m5, that is, a coarser grid during SCF iterations and a more precise grid at the final SCF iteration and the gradient evaluation, as implemented in the TURBOMOLE 5.10 program package.

### B. Time-dependent density functional theory and time-dependent Hartree–Fock

All–electron spin–free TD-DFT (Refs. 63 and 64) calculations were carried out with the ADF2010 package[65–67] in $C_{2v}$ point group symmetry with the PBE (Ref. 68) and PBE0

(Refs. 58 and 59) xc functional and the TZ2P (triple–$\zeta$ double polarization) basis sets from the ADF repository.[69] Scalar relativistic effects were incorporated through the ZORA (zeroth-order regular approximation) Hamiltonian.[70] The spin–free TD-DFT vertical excitation energies were obtained at the optimized geometries within the adiabatic local density approximation (ALDA), in which the local (in time and space) functional derivative of the LDA functional is used in the TD-DFT kernel. The exact (not fitted) density was used to evaluate all xc contributions, with the numerical integration accuracy option set to a value of 10.

For every molecule we determined the 30 lowest-lying (singlet and triplet) states, which was sufficient to characterize the 22 transitions corresponding to excitations from the $25a_1$ ($f_\sigma\{U\} + p_\sigma\{C\}$) and $12b_1/b_2(f_\pi\{U\})$ orbitals to the $26a_1(f_\sigma\{U\} + s_\sigma\{C\} + p_\sigma\{C\})$, $13b_1/b_2(f_\phi\{U\})$, $27a_1/5a_2(f_\delta\{U\})$ and $28a_1/6a_2(f_\delta\{U\})$ orbitals of the CUO molecule and the same kind of orbitals for the $CUONg_4$ model systems. In addition we carried out time-dependent Hartree–Fock (TD-HF) calculations with the same basis set and setup as for the TD-DFT runs.

### C. Frozen density embedding

DFT-in-DFT (Refs. 71–75) calculations were performed with the ADF2010 (Refs. 65–67) package, using a development version of the PYADF (Ref. 76) scripting framework. The same settings as specified above for the supermolecular Kohn–Sham (KS) runs were employed for the DFT part.

In the FDE setup the total electronic density of the CUO complexes ($CUONg_4$, Ng=Ne, Ar, Kr, Xe) was partitioned into the density of CUO and a $Ng_4$ environment and subsequently updated in nine freeze-and-thaw cycles[75] to allow for polarization of both the environment and the active system. We utilized the monomolecular basis set expansion, in which only basis functions belonging to the currently active system are used. This expansion introduces an additional approximation with respect to the supermolecular calculations which is typically small for the cases with limited overlap between the embedded subsystem and the environment.[77] Since the PBE0 xc functional is orbital-dependent it cannot be used to provide an xc contribution to the embedding potential. We chose to use the PBE (Ref. 68) xc for this contribution, following earlier work on orbital–dependent functionals.[46,78–80] For the kinetic energy component of the embedding potential we tested out a number of different functionals: the Thomas–Fermi[81–83] functional, the NDSD functional of Wesołowski and co-workers[84] (which contains a TF component but was developed especially for FDE), the PW91K (Ref. 85) functional, and PW91K with the long-distance correction proposed by Jacob and Visscher[78] (PW91K-CJCORR). This procedure yielded converged embedding potentials that were subsequently used in WFT-in-DFT calculations as well as DFT-in-DFT calculations to allow for comparison with supermolecular results.

These FDE TD-DFT calculations were carried out in the uncoupled approach[86] in which the response of the environment to the electronic excitation is neglected and in which charge-transfer excitations are not taken into account. The validity of the first approximation was investigated by performing exploratory calculations with the coupled excitation framework of Neugebauer[87] from which we found that these couplings are indeed negligibly small.

### D. Wave function theory

The WFT-in-DFT calculations were done using the FDE–implementation[46] in the DIRAC10 (Ref. 88) relativistic quantum chemical package.

The IHFSCC-SD (intermediate Hamiltonian Fock-space coupled cluster with singles and doubles[30–32]) method with a Dirac–Coulomb (DC) Hamiltonian, where $(SS|SS)$ integrals have been approximated by a point charge model,[89] was used. The valence double–$\zeta$ basis set of Dyall[90] (dyall.v2z) for the uranium atom and the augmented correlation consistent polarized valence double–$\zeta$ basis sets of Dunning[91] (aug-cc-pVDZ) for oxygen and carbon atoms were employed.

In the Fock-space coupled cluster[34,40,92,93] method a diagonalization of an effective Hamiltonian yields amplitudes and eigenvalues of the set of states that is related to a closed-shell reference system (sector $(0h, 0p)$) by electron annihilation ($h$), creation ($p$), or both.[94–98]

For CUO, the conventional approach would be to obtain the excitation energies that we are interested in by selecting the appropriate active space for a sector $(1h, 1p)$ ("one hole, one particle") calculation of the Fock-space. This procedure does, however, lead to convergence problems while determining the "one particle" (electron affinity or sector $(0h, 1p)$ of Fock-space) amplitudes that are required as input for the sector $(1h, 1p)$ calculation. These problems could be traced back to the presence of a number of low-lying Rydberg orbitals, that appear at lower energies than the local orbitals of interest (i.e., those that are involved in the lowest electronic transitions) in the CUO molecule. These problems with Rydberg orbitals were not encountered in the earlier work of Infante and Visscher[17] due to lack of very diffuse functions in their basis. Since we are now interested in the interaction with the environment, augmented basis sets are clearly preferable, so we need to find a way around this problem. This can be done by calculating the electronic spectrum via sector $(0h, 2p)$ of Fock-space, i.e., starting from the $CUO^{2+}$ as a reference system and calculating excitation energies as differences between the second electron affinities that are obtained in this sector. With this choice of reference system, the local orbitals are all below the diffuse orbitals and fully converged results can be obtained.

For the systems we investigated, all spinors with energies $\epsilon \in [-3.00; 80.00]$ a.u. were correlated, which correspond to 11 occupied (22 electrons) and about 350 virtual spinors. This corresponds to slightly larger active spaces with respect to those employed in our earlier work on the $UO_2^{2+}$, NUN, and $NUO^+$ molecules,[18,99] and significantly enlarged the valence active space in comparison with the previous work of Infante and Visscher.[17] A detailed description of the active spaces for each particular molecule is given in the supplementary material.[110]

## III. RESULTS AND DISCUSSION

We start with a brief discussion of the electronic structure of the CUO molecule, for which we obtained TD-DFT results with functionals that were not yet applied to this molecule. In order to focus on the essential physical aspects at play, in the discussion that follows we will employ a spin–free model in which the question of ordering of the lowest electronic states reduces to the discussion of a gap between a closed-shell singlet and an open-shell triplet state. We may thereby take over some of the knowledge gained in non relativistic calculations on light molecules to the current species.

### A. Electronic structure of the CUO molecule

The relative energies of the singlet $\sigma^2$ and triplet $\sigma^1\phi^1$ states of the CUO compound have been intensively debated for a number of years as the different electronic structure methods that have been used did not give a consistent picture. While disagreeing on the precise energy difference, all methods give qualitatively the same picture with respect to the bond length and vibrational frequency difference between these two states. The $\sigma^1\phi^1$ ($^3\Phi$) state has longer U–O and in particular U–C bond lengths than the $\sigma^2$ ($^1\Sigma^+$) state. This is easily explained by the transfer of an electron from a bonding U–C orbital to the non-bonding $5f_\phi$-orbital of the uranium atom. Inclusion of spin–orbit coupling (SOC) gives the same qualitative picture with some admixture of the $^3\Delta_2$ state into the lower $^3\Phi_2$ state that lowers this state relative to the $^1\Sigma_0^+$ state which is in first order not affected by SOC.

Roos *et al.*[12] performed CASPT2 calculations with *a posteriori* included SOC and predicted the $^3\Phi_2$ to be the ground-state of CUO. Later on, Infante and Visscher[17] applied a relativistic coupled cluster with singles doubles and perturbative triples—CCSD(T) method—as well as the Fock-space coupled cluster with singles and doubles—FSCC-SD—method. They found a strong stabilization of the $^1\Sigma_0^+$ state by the dynamic correlation effects and concluded that this state should be the ground-state for an isolated molecule. This supports the explanation that was put forward to explain the experimental findings. Recently Yang *et al.*[13] published SOC configuration interaction with singles and doubles (CISD (Ref. 100)) calculations that are in qualitative agreement with these coupled cluster results. They determined the $^1\Sigma_0^+$ state to be the lowest state in both $^1\Sigma_0^+$ and $\Phi_2$ optimized geometries. Noticeable is the large discrepancy (amounting to about 1 eV) between the different electronic structure methods employed so far. From the results of Infante and Visscher it appears that adding dynamical correlation, by correlating more electrons or improving the basis set, favors the more compact $^1\Sigma_0^+$ state, thus yielding a larger gap. It is therefore of interest to see what a DFT treatment of electron correlation gives.

Taking the $^1\Sigma_0^+$ as the reference state we chose the DFT-optimized structure of this state, with U–O and U–C bonds of 1.779 and 1.733 Å, respectively, for spin–free vertical excitation calculations. In the following discussion we will focus on excitations from the $\sigma$ (HOMO) orbital to the virtual $\sigma$, $\delta$, and $\phi$ molecular orbitals. These transitions give rise to the $^{1,3}\Sigma$, $^{1,3}\Delta$, and $^{1,3}\Phi$ excited-states. While the $\delta$- and $\phi$-orbitals also play an important role in some molecules isoelectronic to the $UO_2^{2+}$ such as NUN and $NUO^+$, CUO is special in having also a relatively low–lying $\sigma$ virtual orbital. In most studies this orbital has been ignored, but we will see that it can play a role in some of the calculations.

In agreement with ΔSCF calculations with the same functionals that place the $^3\Phi$ below the reference $^1\Sigma^+$ state, TD-DFT calculations with the PBE (Ref. 68) and PBE0 (Refs. 58 and 59) xc functionals indicate a negative excitation energy for this triplet state. As states with negative excitation energies are not calculated by the standard TD-DFT algorithm,[64] we added a very small admixture of Hartree–Fock (HF = 0.1%) exchange to the PBE xc functional to force the program to use the algorithm for hybrid functionals. As shown in Table I this admixture does not affect the energies of remaining states, but indicates a singlet-triplet gap of −0.13 eV. As found elsewhere (see e.g.,[101] for a recent discussion), such artificially low triplet states are more likely to occur when the fraction of exact exchange is larger. With 25% of exact exchange (the PBE0 xc functional) the $^3\Phi$ state lies 0.20 eV below the $^1\Sigma^+$ state and the separation increases to 0.55 eV for the 100% of exact exchange in the TD-HF method. At the same time the exchange splitting between the $^3\Phi$ and $^1\Phi$ increases dramatically from 0.23 eV (PBE) to 1.96 eV (TD-HF). Concerning the trend noted previously in the WFT calculations, in which an improvement of the description of dynamical correlation lowered the $^1\Sigma$ relative to the $^3\Phi$ state, we find qualitative agreement by defining TD-HF as the most uncorrelated method and TD-DFT with PBE as the most correlated method.

TABLE I. Spin–free vertical excited-states of the CUO molecule with respect to the $^1\Sigma^+$ ground-state (in eV): a comparison of different methods.

| | TD-DFT(PBE) | | | | | | CCSD(T) | | |
|---|---|---|---|---|---|---|---|---|---|
| Symmetry | HF=0% | HF=0.1% | HF=25% | TD-HF | CASPT2[a] | CCSD[b] | Ref. 17 | Ref. 108 | IHFSCC-SD |
| $^3\Sigma^+$ | 0.80 | 0.80 | 1.03 | 1.68 | … | … | … | … | 0.88 |
| $^3\Phi$ | … | −0.13 | −0.20 | −0.55 | 0.09 | 0.77 | 0.83 | 0.68 | 1.34 |
| $^3\Delta$ | 0.41 | 0.41 | 0.46 | 2.35 | 0.54 | … | … | … | 1.36 |
| $^1\Sigma^+$ | 1.47 | 1.48 | 1.77 | 2.85 | … | … | … | … | 1.50 |
| $^1\Delta$ | 0.71 | 0.71 | 0.98 | 2.83 | 0.72 | … | … | … | 1.53 |
| $^1\Phi$ | 0.10 | 0.10 | 0.50 | 1.41 | 0.59 | … | … | … | 1.67 |

[a]Ref. 12.
[b]Ref. 17.

TABLE II. Spin–orbit vertical excited-states of the CUO molecule with respect to the $^1\Sigma_0^+$ ground-state (in eV): a comparison of different methods.

| State[a] | IHFSCC-SD(*22e*) | CCSD(T)(*34e*)[b] | CASPT2(*12e*)[c] | SO-CISD(*24e*)[d] |
|---|---|---|---|---|
| $^3\Sigma_1^+$ | 0.81 | ... | ... | ... |
| $^3\Phi_2$ | 0.94 | 0.60 | –0.36 | 0.30 |
| $^3\Phi_3$ | 1.01 | ... | –0.15 | 0.57 |
| $^3\Delta_1$ | 1.22 | ... | 0.19 | 0.58 |
| $^3\Delta_2$ | 1.28 | ... | 0.31 | ... |

[a]Subscripts denote Ω-values and upperscripts spin multiplicity.
[b]Ref. 17.
[c]Ref. 12.
[d]Ref. 13.

For the WFT-based methods we note that the spin–free CASPT2 calculations of Roos *et al.*[12] predict the $^3\Phi$ state slightly above the singlet state by 0.09 eV, but this is not enough for this state to remain the lowest when SOC is accounted for. Since dynamical correlation is only included up to the second-order in perturbation theory, one can again fit these results in a trend leading from HF (predicting the $^3\Phi$ state to be the ground-state by 0.5 eV) to the coupled cluster methods that predict a $^1\Sigma_0^+$ ground-state. This trend is even stronger when including our new IHFSCC-SD (Refs. 30–32) data that were obtained using a larger basis set and active space than was feasible earlier.[17] From these calculations we obtain the $^3\Phi$ state at even more than 1.3 eV above the closed-shell $^1\Sigma^+$ state, with an open-shell $^3\Sigma^+$ state being the first excited-state at 0.88 eV.

We believe that the new DFT and CC results are interesting because they clearly illustrate the difficulties in finding a proper description of the balance between exchange and correlation (in a DFT picture) and static and dynamic correlation (in a WFT picture) for actinides. The model xc hole used in the DFT approach will have to provide an equally good description of the strong angular correlation found in the radially localized $\phi$-orbitals as well as for the qualitatively different correlation in the $\sigma$ bonding orbital. This is difficult to achieve on basis of information from only the electron density and its gradient, while adding exchange in a hybrid approach has a limited value as this brings along the large error in the HF description. For the WFT methods the challenge is equally large as an electron in the $\phi$-orbital will have a stronger interaction with the semicore electrons than an electron in a more extended $\sigma$-orbital, requiring a well-balanced basis set and a substantial number of electrons to be correlated. We plan to study these aspects in more detail in another publication.

Coming back to the main topic of this paper, we also analyzed the influence of the spin–orbit coupling on the transitions. The results are listed in Table II in which we kept the major LS-designation but added the Ω values, which is the only proper quantum number after inclusion of SOC. Besides the relativistic splitting of all the triplet states, there is no qualitative difference with the spin–free transition energies.

## B. The $CUONg_4$ models

In this section we investigate how the interaction with the environment affects the lowest electronic states. We focus on the $^1\Sigma^+$ and $^3\Phi$ states, but also consider the fate of the $^3\Sigma^+$ state when the system is confined in a matrix.

### 1. Geometries and vibrational frequencies

Table III lists all the geometries and vibrational frequencies obtained in this work. We observe a slight increase of the U–C and U–O bond distances due to the environment that becomes stronger when we attach heavier, more polarizable, noble gas (Ng) atoms. The four attached noble gas atoms move slightly out of the equatorial plane with a C–U–Ng angle slightly larger than 90° in the singlet states and slightly smaller than 90° in the triplet states of the $CUOXe_4$ complex.

TABLE III. Optimized structures and vibrational spectra calculated using DFT with PBE0 xc functional in the $C_{2v}$ point group symmetry. Bond distances are given in Å, angles in degrees, and frequencies in $cm^{-1}$. Reference theoretical studies are presented in round brackets and experimental data in square brackets.

| System | d(U–C) | d(U–O) | d(U–Ng) | ∠ CUNg | $\nu_{UC}$ | $\nu_{UO}$ |
|---|---|---|---|---|---|---|
| $CUO(^1\Sigma^+)$ | 1.733 (1.714[a], 1.770[b], 1.757[c], 1.738[d]) | 1.779 (1.759[a], 1.795[b], 1.798[c], 1.785[d]) | | | 1175 (1269[a], 1089[c], 1183[f], 1182[d]) | 927 (960[a], 870[d], 917[e], 922[f]) |
| $CUONe_4(^1\Sigma^+)$ | 1.734 (1.769[c]) | 1.782 (1.805[c]) | 3.315 | 99.7 | 1170 (983[c]) [1047[c]] | 923 (847[c]) [872[c]] |
| $CUOAr_4(^1\Sigma^+)$ | 1.738 (1.774[c]) | 1.788 (1.811[c]) | 3.416 | 95.8 | 1156 (963[c]) | 913 (840[c]) |
| $CUOKr_4(^1\Sigma^+)$ | 1.741 (1.775[c]) | 1.790 (1.812[c]) | 3.455 | 94.5 | 1145 (947[c]) | 908 (839[c]) |
| $CUOXe_4(^1\Sigma^+)$ | 1.747 (1.782[c]) | 1.793 (1.813[c]) | 3.554 | 92.4 | 1127 (942[c]) | 901 (839[c]) |
| $CUO(^3\Phi)$ | 1.836 (1.814[a], 1.871[b], 1.857[c]) | 1.808 (1.789[a], 1.818[b], 1.825[c]) | | | 948 (893[c]) | 879 (828[c]) |
| $CUONe_4(^3\Phi)$ | 1.840 (1.862[c]) | 1.811 (1.828[c]) | 3.147 | 91.7 | 942 (884[c]) | 872 (819[c]) |
| $CUOAr_4(^3\Phi)$ | 1.845 (1.868[f], 1.871[c]) | 1.815 (1.843[f], 1.833[c]) | 3.309 | 90.4 | 932 (881[f], 869[c]) [835[f], 853[c]] | 868 (824[f], 814[c]) [793[f], 804[c]] |
| $CUOKr_4(^3\Phi)$ | 1.848 (1.873[f], 1.874[c]) | 1.817 (1.839[f], 1.835[c]) | 3.374 | 89.9 | 926 (876[f], 864[c]) [832[f], 842[c]] | 864 (822[f], 810[c]) [790[f], 797[c]] |
| $CUOXe_4(^3\Phi)$ | 1.853 (1.875[f], 1.879[c]) | 1.818 (1.833[f], 1.836[c]) | 3.514 | 88.6 | 917 (870[f], 860[c]) [830[f]] | 860 (819[f], 808[c]) [789[f]] |

[a]MRSOCISD (Ref. 13).
[b]SO-CASPT2 (Ref. 12).
[c]DFT/PW91 (Ref. 23).
[d]DFT/PW91 (Ref. 26).
[e]DFT/PBE0 (Ref. 109).
[f]DFT/PBE0 (Ref. 102).

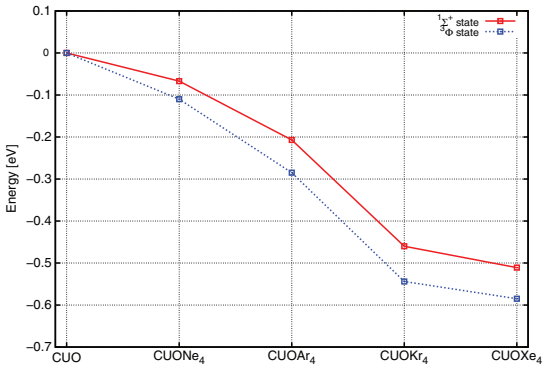



FIG. 1. Total bonding energies of the $CUONg_4$ complexes with respect to to the $^1\Sigma^+$ (solid line) and $^3\Phi$ (dashed line) states of the CUO system (DFT calculations in ADF).

The weakening of the U–C and (to less extent) U–O bond in the triplet state is most clearly visible in the U–C and U–O stretch vibrations. For the bare molecule the U–C stretch is lower by 227 $cm^{-1}$ in the triplet state than in the singlet state. Compared to this difference the matrix effects are much weaker with a shift of 31 (48) $cm^{-1}$ of the U–C stretch for the singlet (triplet) state in the model for the xenon matrix. The experimentally observed shift of 238 $cm^{-1}$ of the U–C stretch upon going from neon to argon does therefore best match with the 195 $cm^{-1}$ difference between the vibration of the singlet state in the neon model and the triplet state in the argon model. This is already discussed in detail by Andrews *et al.*[23] who obtained slightly different values but a qualitatively similar trend with the PW91K functional. The new data with hybrid functionals provides further support for their interpretation. Note that only trends can be reliably compared, for a more complete comparison with the experimental values it might be necessary to include more than just four noble gases around the CUO to model the matrix. Moreover, it might be necessary to consider also anharmonic effects.[102]

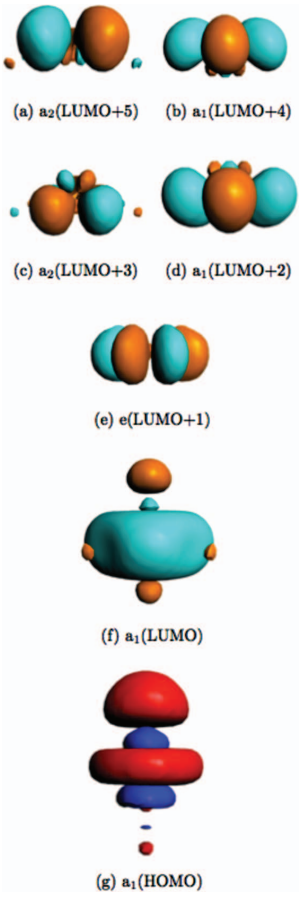



FIG. 2. $CUONe_4$ orbitals involved in the lowest-lying electronic transitions visualized with an isosurface value of 0.03 in the ADFGUI.[107]

### 2. Ground-state DFT study

To simplify the comparison between the CUO and $CUONg_4$ complexes, we will use (idealized) $C_{4v}$ point group symbols, with in parenthesis the $C_{\infty v}$ designations to indicate the parentage of the states or orbitals. We define the zero of our energy scale as the $^1A_1(^1\Sigma^+)$ state of CUO with a cage of four Ng atoms at infinite distance and plot in Figure 1 the variation of the binding energy of the $^1A_1$ ($^1\Sigma^+$) and the $^3E(^3\Phi)$ states as a function of Ng type. This interaction with the cage is indeed slightly more pronounced in the triplet state with the largest difference occurring when moving from neon to argon, but the difference is only 0.04 eV. This can only change the order of two states if they are already very nearly degenerate. As discussed already in the section of TD-DFT, the PBE0 functional places the triplet slightly below the singlet already for the bare CUO molecule and considering optimized structures for the triplet state reinforces that conclusion.

### 3. Excitation energies from TD-DFT

To understand the trends in excitation energies, it is instructive to look at the changes in the valence orbital energies of CUO induced by the attachment of the Ng atoms. We consider the closed-shell singlet calculation in which we have as HOMO a bonding U–C $\sigma$-orbital, as LUMO a nonbonding $\sigma$-orbital, and at higher energy a $\phi$- and two $\delta$-orbitals. These orbitals are depicted in Figure 2. When adding the Ng atoms the two $\delta$-orbitals are split by the ligand field into $27a_1$ and $5a_2$, and the $28a_1$- and $6a_2$-orbitals, respectively. In Figure 3(a) we plot the energy difference of all relevant virtual orbitals relative to the HOMO to get a first indication of the effect of the environment on excitation energies. The diffuse non-bonding $\sigma$-orbital, which is the LUMO in bare CUO and $CUONe_4$, is pushed up by the repulsive interaction with the closed-shell ligands and rises in the argon system above the $\phi$-orbitals, and in krypton and xenon also above the lower $\delta$-orbital $27a_1$. This indicates that this orbital is indeed of less interest in explaining the electronic structure of the molecule in the matrix.

Spin–free excitation energies are listed in Table IV, with the negative transition energies indicating a triplet ground-state. In agreement with the ΔSCF calculations, we find small

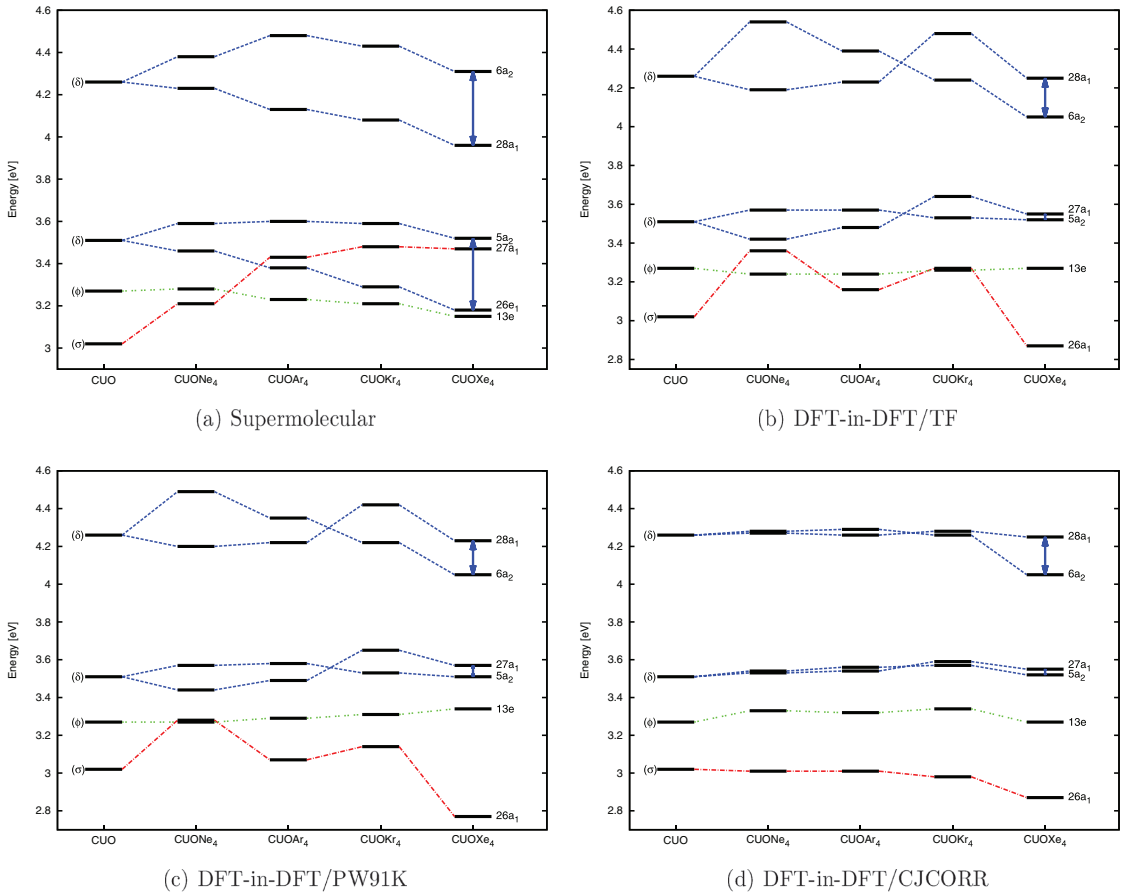



FIG. 3. Lowest-lying valence orbital energies of the CUO ($^1\Sigma^+$) and the CUONg$_4$($^1$A$_1$) complexes from the supermolecular and DFT-in-DFT/PBE0 calculations using different kinetic energy functionals in ADF. Orbital energies are given relative to the energy of HOMO for all the compounds.

effects on the transition energies, except for the large effect on the b $^3$A$_1$ ($^3\Sigma^+$) and d $^1$A$_1$ ($^1\Sigma^+$) states that correspond to excitation to the diffuse $\sigma$-orbital. These transitions rapidly shift to higher energies when adding the noble gas cage. Other changes in the electronic spectrum introduced by the noble gas environment are relatively small for the lowest-lying excitations (difference less than 0.1 eV). Those results are shown in supermolecular case, where the differences in the electronic transitions between the bare CUO and the CUONg$_4$ complexes are plotted. We note a relatively large difference between the Ne and Ar cage, but as already seen in the orbital energy differences, an overall surprisingly small effect of the matrix given the experimental findings.

### C. DFT-in-DFT electronic structure of the CUONg$_4$ compounds

#### 1. The quality of embedding potential

We now investigate whether the subtle effect on the environment can be represented by an FDE embedding potential. Two criteria may be used to quantify the quality of the FDE-approach: (1) the reproduction of the ground-state density of the complex and (2) the error in the calculated transition energies. These errors can be evaluated exactly in DFT and are supposed to carry over to the WFT description.

To assess the first error we compare the electron density calculated within the DFT-in-DFT framework to the supermolecular density. This criterium tests the accuracy of the approximation used for the non-additive parts of the kinetic energy and xc functionals. For the latter we restrict ourselves to the PBE functional to remain as close as possible to the PBE0 hybrid functional used in the supermolecular calculation. For the repulsive kinetic energy contribution to the embedding potential that models the Pauli repulsion of the Ng cage, we used the simplest local Thomas–Fermi (TF) (Refs. 81–83) and the often more robust[77] gradient-corrected functional PW91K.[85] Since underestimation of the Pauli repulsion may give rise to unphysical transfer of electron density from the active center to the environment subsystem—the so called "electron leak" problem,[52,78,103] we also considered the NDSD

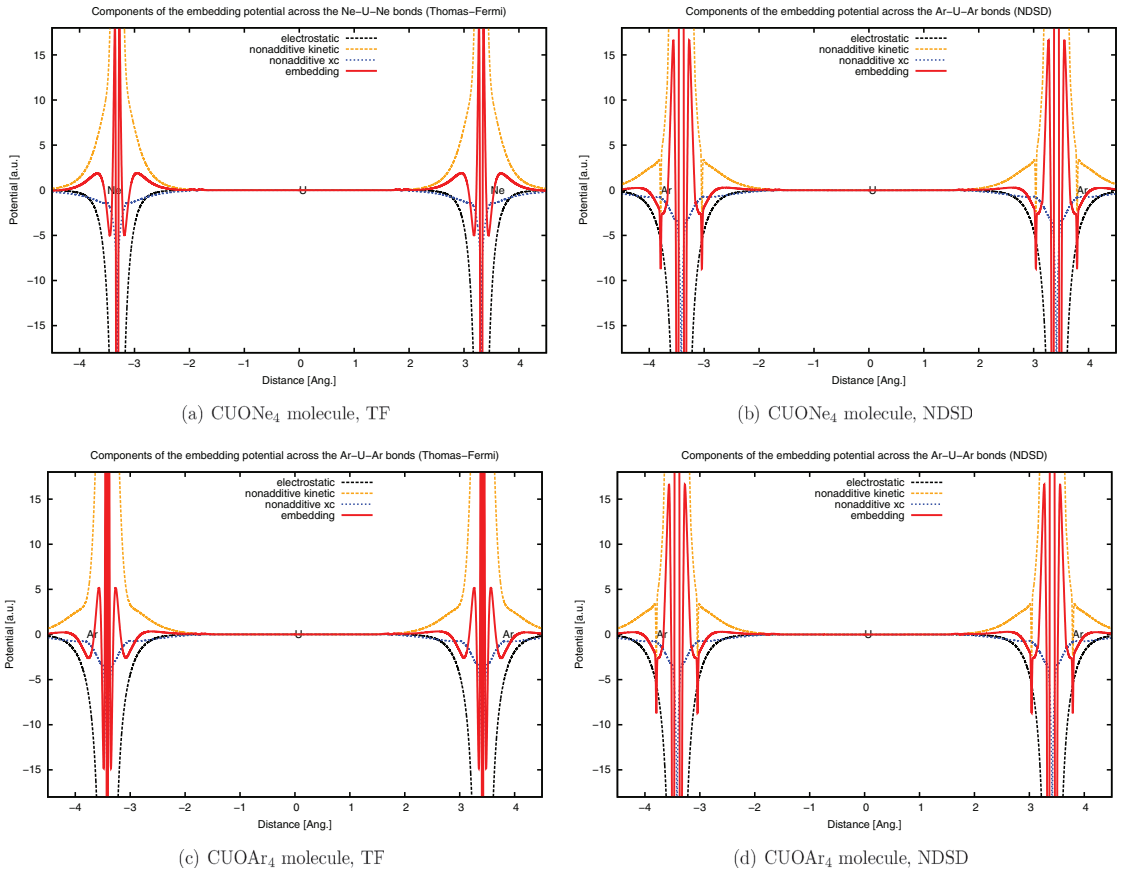
TABLE IV. Spin-free vertical excitation energies of the CUO($^1\Sigma^+$) and CUONg$_4$($^1A_1$) from the KS/TD-DFT/PBE0 approach (in eV).

| Symmetry[a] | CUO | CUONe$_4$ | CUOAr$_4$ | CUOKr$_4$ | CUOXe$_4$ |
|---|---|---|---|---|---|
| a$^3$E ($^3\Phi$) | −0.204 | −0.222 | −0.233 | −0.253 | −0.231 |
| a$^3$A$_1$ ($^3\Delta$) | 0.464 | 0.431 | 0.384 | 0.211 | 0.368 |
| a$^3$A$_2$ ($^3\Delta$) | 0.464 | 0.440 | 0.399 | 0.259 | 0.385 |
| a$^1$E ($^1\Phi$) | 0.504 | 0.472 | 0.434 | 0.333 | 0.411 |
| a$^1$A$_1$ ($^1\Delta$) | 0.983 | 0.951 | 0.915 | 0.839 | 0.891 |
| a$^1$A$_2$ ($^1\Delta$) | 0.983 | 0.978 | 0.960 | 0.892 | 0.949 |
| b$^3$A$_1$ ($^3\Sigma^+$) | 1.029 | 1.245 | 1.598 | 1.930 | 1.826 |
| c$^3$A$_1$ ($^3\Gamma$) | 1.317 | 1.289 | 1.274 | 1.234 | 1.281 |
| b$^3$A$_2$ ($^3\Gamma$) | 1.317 | 1.289 | 1.275 | 1.236 | 1.284 |
| d$^3$A$_1$ ($^3\Delta$) | 1.324 | 1.296 | 1.281 | 1.245 | 1.289 |
| c$^3$A$_2$ ($^3\Delta$) | 1.324 | 1.296 | 1.281 | 1.243 | 1.289 |
| b$^1$A$_1$ ($^1\Gamma$) | 1.484 | 1.455 | 1.438 | 1.405 | 1.441 |
| b$^1$A$_2$ ($^1\Gamma$) | 1.484 | 1.455 | 1.438 | 1.404 | 1.440 |
| c$^1$A$_1$ ($^1\Delta$) | 1.726 | 1.693 | 1.666 | 1.653 | 1.620 |
| c$^1$A$_2$ ($^1\Delta$) | 1.726 | 1.710 | 1.698 | 1.671 | 1.692 |
| b$^3$E ($^3\Phi$) | 1.773 | 1.748 | 1.736 | 1.698 | 1.745 |
| c$^3$E ($^3\Phi$) | 1.775 | 1.765 | 1.761 | 1.737 | 1.773 |
| d$^1$A$_1$ ($^1\Sigma^+$) | 1.771 | 1.938 | 2.156 | 2.323 | 2.248 |

[a]We use the $C_{4v}$ point group symmetry notation with $C_{\infty v}$ irreps indicated in parenthesis.

(Ref. 84) functional and the zero-overlap correction of Jacob and Visscher[78] that eliminate such problems. We used the error measures defined by Bernard *et al.*:[104] integrated absolute errors in the electron density, the integrated root mean square errors in the electron density, and the magnitude of the errors in the dipole moment. While the former two strictly depend on the absolute size of the error in the electron density, the latter also provides information on its spatial redistribution. The size of the errors is taken relative to the sum of fragment density obtained by simply superimposing the density of the cage and the density of the bare CUO molecule.

The simple TF potential provides a reasonable description and has even the smallest errors measures for the complex with xenon (see Table V). For the neon cage PW91K provides the best description, while for the other two cases, argon and krypton, no unambiguously best functional can be selected. In correcting for charge-leak artifacts, the Jacob and Visscher[78] correction reduces the embedding potential too much, underestimating the interaction between CUO and the cage. This can have an adverse effect on the dipole moment, when compared to its parent functional, PW91K. For the NDSD (Ref. 84) kinetic energy functional that is based on the TF



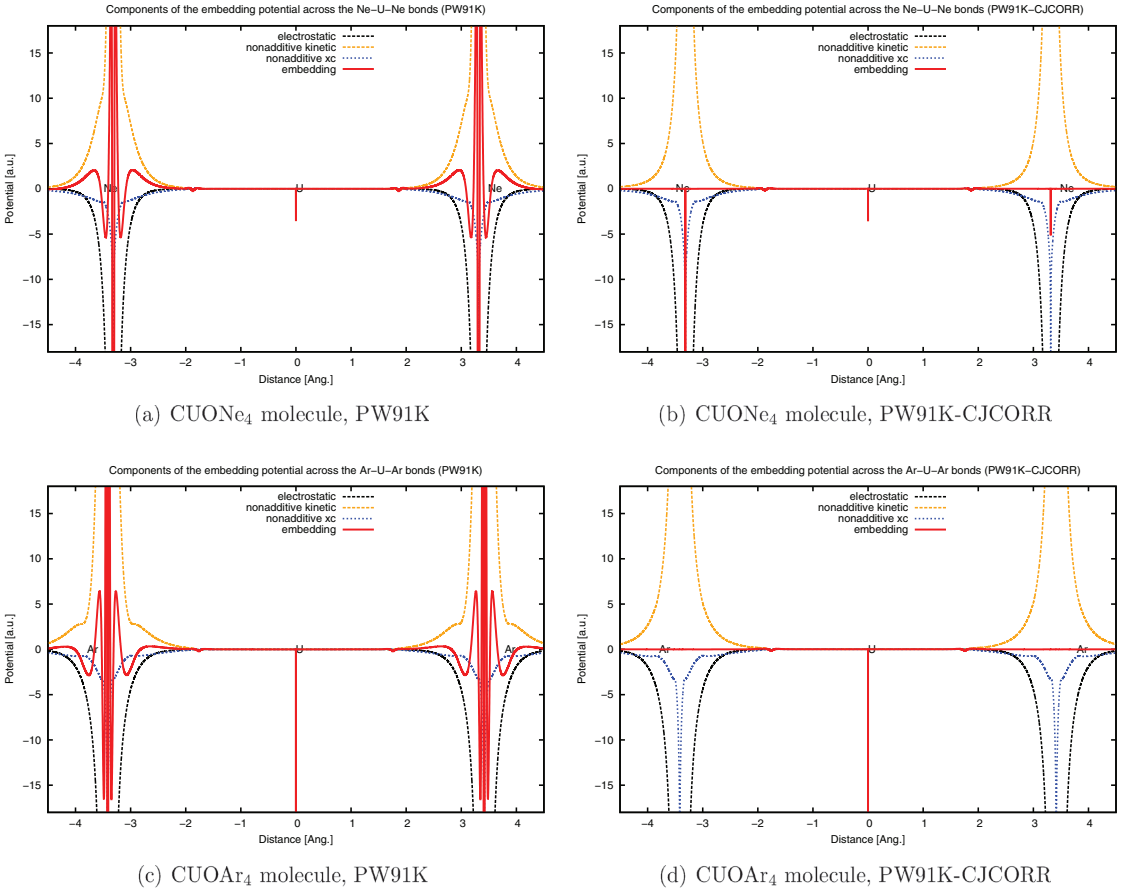

(a) CUONe$_4$ molecule, TF

(b) CUONe$_4$ molecule, NDSD

(c) CUOAr$_4$ molecule, TF

(d) CUOAr$_4$ molecule, NDSD

FIG. 4. Contributions to the total embedding potential from the electrostatic, non-additive xc and non-additive kinetic energy—TF to the left and NDSD to the right—components along the Ng–U–Ng axis in the CUONg$_4$($^1A_1$) complexes (in hartree).




(a) $CUONe_4$ molecule, PW91K (b) $CUONe_4$ molecule, PW91K-CJCORR

(c) $CUOAr_4$ molecule, PW91K (d) $CUOAr_4$ molecule, PW91K-CJCORR

FIG. 5. Contributions to the total embedding potential from the electrostatic, non-additive xc and non-additive kinetic energy—PW91K to the left and PW91K-CJCORR to the right—components along the Ng–U–Ng axis in the $CUONg_4(^1A_1)$ complexes (in hartree).

functional we see an improvement relative to TF for the Ar and Kr complexes but a slightly worse performance for the Ne and Xe complexes.

The best overall performance for the two cases that are of most interest, neon and argon, is therefore obtained with the PW91K kinetic energy functional. More interesting than these errors in the density is, however, the effect on the valence orbital energies of the DFT-in-DFT $CUONg_4$ complexes. These results are indicative of the errors that can be expected in the supermolecular spectra. This data are given in Figures 3(b)–3(d). From the figures we observe that the effect of the cage is generally underestimated by the FDE approach: while in KS/DFT the $26a_1$-orbital is the LUMO only for the CUO and $CUONe_4$ molecules, with FDE it is, e.g., always found as LUMO for $CUOXe_4$. Even more troublesome is the fact that the order of the $26a_1$- and $13b_1/b_2$-orbitals is reversed (crossing of the blue dashed lines between Ar and Kr) for the heavier noble gas cages. To check whether these artifacts are due to the freeze-and-thaw procedure we also did some test calculations in which the density of unperturbed Ng atoms was used to represent the density of the cage, but this gave the same picture.

More insight can be obtained by analyzing the three separate components of the potential: the electrostatic (Coulomb and nuclear), and the non-additive xc and kinetic energy components. In Figures 4 and 5 these are plotted along the Ne–U–Ne and Ar–U–Ar lines in an idealized geometry (with the Ng all exactly in the equatorial plane). In all pictures we see that it is the balance between the large attractive electrostatic and the repulsive kinetic energy components that determines the potential. This balance is more difficult to obtain when the electrostatic terms become larger, as is the case for the heavier Ng atoms. The exact potential should furthermore reflect the shell structure of the atoms that also becomes more pronounced for the heavier atoms. For the uncorrected TF and PW91K potentials the shell structure mainly arises from the compensation of the monotonously negative Coulomb potential by the monotonously positive kinetic energy component, while for the NDSD approach the switching function that is used to interpolate between two functional forms also introduces oscillations in the kinetic energy component for the heavier Ng atoms. The Jacob and Visscher correction[78] operates on the full potential and reduces this in the vicinity of the nuclei to obey an exact limit for non-overlapping systems.

TABLE V. Integrated errors in the electron density: $\Delta^{abs}$ (absolute) and $\Delta^{rms}$ (root mean square), magnitude of the error in the dipole moment $|\Delta_\mu|$ for the sum of fragments, and the DFT-in-DFT/PBE0 calculations with a different approximate non-additive kinetic energy functionals. The most accurate values are marked in boldface.

| System | | $\Delta^{abs} \times 10^{-3}$ | $\Delta^{rms} \times 10^{-3}$ | $\lvert\Delta_\mu\rvert$(D) |
|---|---|---|---|---|
| $^1$CUONe$_4$ | Sum of fragments | 1.35 | 0.99 | 0.183 |
| | FDE (Thomas–Fermi) | 4.57 | 0.03 | 0.029 |
| | FDE (NDSD) | 4.84 | 0.03 | 0.033 |
| | FDE (PW91K) | **2.32** | **0.02** | **0.010** |
| | FDE (PW91K-CJCORR) | 6.23 | 0.04 | 0.171 |
| $^1$CUOAr$_4$ | Sum of fragments | 2.76 | 1.39 | 0.354 |
| | FDE (Thomas–Fermi) | 1.42 | 1.39 | 0.146 |
| | FDE (NDSD) | 1.40 | **0.06** | **0.139** |
| | FDE (PW91K) | **1.39** | 0.07 | 0.261 |
| | FDE (PW91K-CJCORR) | 1.66 | 0.09 | 0.370 |
| $^1$CUOKr$_4$ | Sum of fragments | 2.71 | 1.23 | 0.225 |
| | FDE (Thomas–Fermi) | 1.08 | **0.06** | **0.089** |
| | FDE (NDSD) | 1.10 | 0.06 | 0.115 |
| | FDE (PW91K) | **1.03** | 0.06 | 0.188 |
| | FDE (PW91K-CJCORR) | 1.33 | 0.85 | 0.348 |
| $^1$CUOXe$_4$ | Sum of fragments | 2.27 | 0.11 | 0.205 |
| | FDE (Thomas–Fermi) | **1.36** | **0.07** | **0.178** |
| | FDE (NDSD) | 1.41 | 0.07 | 0.222 |
| | FDE (PW91K) | 1.40 | 0.08 | 0.302 |
| | FDE (PW91K-CJCORR) | 1.40 | 0.08 | 0.307 |

For the current system this leads to an almost complete reduction of the potential which explains the too small orbital energy shifts for this approach.

#### 2. Excitation energies from embedded TD-DFT

As the goal of the embedding approach is to reproduce the supermolecular approach, that is the data in Table IV, we directly compare the DFT-in-DFT to the reference supermolecular excitation energies in Figure 6. As was expected from the errors seen in the orbital energies, none of the kinetic energy approximations is able to reproduce the KS/TD-DFT/PBE0 results for the heavier Ng matrices. Errors are largest in absolute magnitude for the $^{1,3}A_1$ and $^{1,3}A_2$ excited-states in heavier noble gases. All embedding potentials do shift the $^3E$ state above the $^1A_1$ ground-state. For all excitations the effect of the environment is small and the PW91K functional provides the relatively best performance, in agreement with the density error analysis discussed above.

### D. WFT-in-DFT electronic structure of the CUONg$_4$ compounds

Considering all the information discussed above we find that the PW91K kinetic energy functional can give a qualitative description of the neon and to a less extent argon environments, although quantitative agreement and correct trends are not reached. We decided to select this embedding potential to perform WFT-in-DFT calculations and check whether the same environment effects are found.

The SOC WFT-in-DFT excitation energies are listed in Table VI. The effect of noble gas environment on the transitions is clearly much too small to cause a ground-state change in the WFT calculations, but just sufficient, in the case of neon, to place the $\Phi_2$ excited-state below the two $^3\Sigma^+$ states. For argon this trend is reversed, in contrast to the supermolecular DFT calculations, which is due to the larger error in the embedding potential for this system. This error is larger

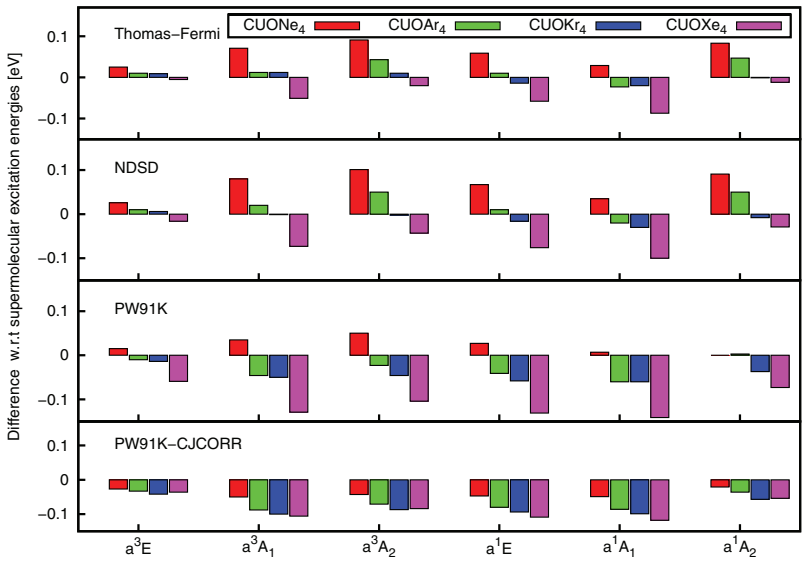



FIG. 6. Accuracy of different approximate non-additive kinetic energy functionals on the spin–free vertical electronic spectra of the CUONg$_4$($^1$A$_1$) complexes with respect to the supermolecular (KS/TD-DFT/PBE0) reference data (in eV). Red bars correspond to CUONe$_4$, green to CUOAr$_4$, blue to CUOKr$_4$ and violet to CUOXe$_4$ complexes.

TABLE VI. SOC vertical excitation energies of the CUO and $CUONg_4$ complexes from the WFT-in-DFT: IHFSCC-SD via sector (0*h*, 2*p*) for the bare CUO subsystem and embedded in the $Ng_4$ DFT-in-DFT/PBE0/PW91K potential.

| State[a] | CUO | $CUONe_4$ | $CUOAr_4$ | $CUOKr_4$ | $CUOXe_4$ |
|---|---|---|---|---|---|
| $^3\Sigma_{0,1}$[b] | 0.81 | 1.02 | 0.85 | 0.76 | 0.65 |
| $^3\Phi_2$ | 0.94 | 0.84 | 0.89 | 0.93 | 0.97 |
| $^3\Phi_3$ | 1.01 | 0.94 | 0.98 | 1.02 | 1.06 |
| $^3\Delta_1$ | 1.23 | 1.07 | 1.27 | 1.30 | 1.33 |
| $^3\Delta_2$ | 1.28 | 1.17 | 1.31 | 1.34 | 1.37 |
| $^1\Sigma_0$ | 1.37 | 1.51 | 1.36 | 1.31 | 1.21 |
| $^3\Phi_4$ | 1.60 | 1.53 | 1.57 | 1.61 | 1.64 |
| $^3\Delta_3$ | 1.69 | 1.58 | 1.58 | 1.67 | 1.64 |
| $^1\Delta_2$ | 1.72 | 1.68 | 1.75 | 1.78 | 1.76 |
| $^1\Phi_3$ | 1.79 | 1.73 | 1.76 | 1.79 | 1.80 |

[a]Subscripts denote Ω-values and upperscripts spin multiplicity.
[b]The differences between $^3\Sigma_0^+$ and $^3\Sigma_1^+$ electronic states are less than 0.01 eV.

still when utilizing an embedding with heavier Ng atoms, and leads to meaningless results as expected on basis of the DFT-in-DFT results.

## IV. CONCLUSIONS AND OUTLOOK

We have optimized geometries and calculated vibrational frequencies of the CUO and $CUONg_4$ complexes in the $^1\Sigma^+$ and $^3\Phi$ states using the PBE0 hybrid xc functional. The significant difference in characteristic U–O and U–C vibrational frequencies that was observed in experiments, is in agreement with previous work, explained by a difference of electronic ground-state ($^1\Sigma^+$ to $^3\Phi$) in the neon and argon matrix, respectively. The direct effect of environment is smaller than 50 $cm^{-1}$ for the U–O and U–C stretching frequencies.

Since analysis of the vibrational spectra provides only an indirect measure for the energetic ordering of the electronic states, we also investigated this ordering directly for CUO and the $CUONg_4$ model compounds. The spin–free and spin–orbit IHFSCC-SD excitation energies of the CUO molecule, indicate that this molecule has a singlet ground-state, which is in line with earlier Fock-space coupled cluster studies by one of us[17] and the CISD calculations of Yang *et al.*[13] These coupled cluster energies were used to assess the accuracy of hybrid TD-DFT/PBE0 excitation energies. It turns out that both components of PBE0 xc functional—DFT and HF part—are insufficiently accurate. This issue can be related to the well–known "triplet instability" problem,[101,105,106] caused by the large exchange splitting in the 5*f*-shell, that is strongly overestimated in Hartree–Fock theory.

Compared to the large energy splittings predicted by the WFT approaches, the low-lying excited-states show only minor perturbations due to the noble gas environment. An exception are the $^{3,1}\Sigma$ excited-states that originate from electron transfer to the diffuse virtual $f_\sigma$-orbital. This diffuse orbital, and the corresponding transitions, are shifted to a much higher energy by the equatorial ligands.

Since the perturbations are small, we expected that these could be captured by the FDE approach, but it turns out that none of the currently available kinetic energy functionals is able to yield results with the desired accuracy of less than 0.1 eV. For the $CUONe_4$ model the PW91K kinetic energy functional provides a reasonable agreement with a supermolecular approach, but for the complexes with heavier noble gases all kinetic energy functional produce too small ligand field splittings and a qualitatively incorrect result. WFT-in-DFT calculations on the $CUONg_4$ systems should, however, be feasible as soon as embedding potentials improve because the analysis of the densities shows that a density partitioning into a CUO and an environment density is indeed well possible.

As a side result of our work we note the importance of further development of xc functionals that can provide qualitatively correct xc splittings for actinides. While for the $UO_2^{2+}$, NUN, and $NUO^+$ molecules good agreement between the TD-DFT and IHFSCC-SD electronic excitations can be reached, this is currently not the case for the CUO molecule.

## ACKNOWLEDGMENTS

The authors thank Dr. Erik van Lenthe and Mirco Franchini for helpful discussions on the practical aspects of carrying out the ADF calculations and Dr. Christoph R. Jacob for assistance with plotting different components of embedding potential within a development version of PYADF. P.T. and L.V. acknowledge the financial support from The Netherlands Organization for Scientific Research (NWO) via the Vici program and from the Dutch National Computing Facilities (NCF) for computational resources at the LISA and HUYGENS facilities at SARA. A.S.P.G. acknowledges support from PhLAM (Laboratoire de Physique des Lasers, Atomes et Molécules, Unité Mixte de Recherche de l'Université de Lille 1 et du CNRS).

[1]G. J. Hutchings, C. S. Heneghan, I. D. Hudson, and S. H. Taylor, Nature (London) **384**, 341 (1996).
[2]Z. T. Zhang, M. Konduru, S. Dai, and S. H. Overbury, Chem. Commun. **20**, 2406 (2002).
[3]R. G. Denning, J. Phys. Chem. A **111**, 4125 (2007).
[4]M. Pepper and B. E. Bursten, Chem. Rev. **91**, 719 (1991).
[5]S. Fortier, G. Wu, and T. W. Hayton, J. Am. Chem. Soc. **132**, 6888 (2010).
[6]P. C. Burns, Y. Ikeda, and K. Czerwinski, MRS Bull. **35**, 868 (2010).
[7]S. Matsika, Z. Zhang, S. R. Brozell, J.-P. Blaudeau, Q. Wang, and R. M. Pitzer, J. Phys. Chem. A **105**, 3825 (2001).
[8]M. Matsika and M. Pitzer, J. Phys. Chem. A **105**, 637 (2001).
[9]P. Pyykkö, J. Li, and N. Runeberg, J. Phys. Chem. **98**, 4809 (1994).
[10]N. Kaltsoyannis, Inorg. Chem. **39**, 6009 (2000).
[11]L. Gagliardi and B. O. Roos, Chem. Soc. Rev. **36**, 893 (2007).
[12]B. O. Roos, P.-O. Widmark, and L. Gagliardi, Faraday Discuss. **124**, 57 (2003).
[13]T. Yang, R. Tyagi, Z. Zhang, and R. M. Pitzer, Mol. Phys. **107**, 1193 (2009).
[14]X. Wang, L. Andrews, P.-A. Malmqvist, B. O. Roos, A. P. Gonçalves, C. C. L. Pereira, J. Marçalo, C. Godart, and B. Villeroy, J. Am. Chem. Soc. **132**, 8484 (2010).
[15]K. Tatsumi and R. Hoffmann, Inorg. Chem. **19**, 2656 (1980).
[16]K. Dyall, Mol. Phys. **96**, 511 (1999).
[17]I. Infante and L. Visscher, J. Chem. Phys. **121**, 5783 (2004).
[18]P. Tecmer, A. S. P. Gomes, U. Ekström, and L. Visscher, Phys. Chem. Chem. Phys. **13**, 6249 (2011).
[19]P. Tecmer, R. Bast, K. Ruud, and L. Visscher, J. Phys. Chem. A **116**, 7397 (2012).
[20]T. J. Tague, L. Andrews, and R. D. Hunt, J. Phys. Chem. **97**, 10920 (1993).
[21]M. Zhou, L. Andrews, J. Li, and B. E. Bursten, J. Am. Chem. Soc. **121**, 9712 (1999).

[22]B. Liang, L. Andrews, J. Li, and B. E. Bursten, J. Am. Chem. Soc. **124**, 9016 (2002).
[23]L. Andrews, B. Liang, J. Li, and B. E. Bursten, J. Am. Chem. Soc. **125**, 3126 (2003).
[24]J. Li, B. E. Bursten, B. Liang, and L. Andrews, Science **295**, 2242 (2002).
[25]B. Liang, L. Andrews, J. Li, and B. E. Bursten, Chem.-Eur. J. **9**, 4781 (2003).
[26]B. Liang, L. Andrews, J. Li, and B. E. Bursten, Inorg. Chem. **43**, 882 (2004).
[27]V. Goncharov, L. A. Kaledin, and M. C. Heaven, J. Chem. Phys. **125**, 133202 (2006).
[28]J. Jin, R. Gondalia, and M. C. Heaven, J. Phys. Chem. A **113**, 12724 (2009).
[29]T. Fleig, Chem. Phys. **395**, 2 (2012).
[30]A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **113**, 9905 (2000).
[31]A. Landau, E. Eliav, Y. Ishikawa, and U. Kaldor, J. Chem. Phys. **115**, 6862 (2001).
[32]L. Visscher, E. Eliav, and U. Kaldor, J. Chem. Phys. **115**, 9720 (2001).
[33]R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[34]V. V. Ivanov, D. I. Lyakh, and L. Adamowicz, Phys. Chem. Chem. Phys. **11**, 2355 (2009).
[35]L. Meissner and M. Musiał, *Recent Progress in Coupled Cluster Methods* (Springer, 2010), p. 395.
[36]K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Woliński, J. Phys. Chem. **94**, 5483 (1990).
[37]K. Andersson, P.-A. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).
[38]I. Infante, L. Andrews, X. Wang, and L. Gagliardi, Chem. Eur. J. **16**, 12804 (2010).
[39]P.-A. Malmqvist, K. Pierloot, A. R. M. Shahi, C. J. Cramer, and L. Gagliardi, J. Chem. Phys. **128**, 204109 (2008).
[40]A. S. P. Gomes and C. R. Jacob, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem. **108**, 222 (2012).
[41]N. Govind, Y. A. Wang, A. J. R. da Silva, and E. Carter, Chem. Phys. Lett. **295**, 129 (1998).
[42]N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **110**, 7677 (1999).
[43]P. Huang and E. A. Carter, J. Chem. Phys. **125**, 084102 (2006).
[44]T. Klüner, Y. A. Wang, N. Govind, and E. A. Carter, Phys. Rev. Lett. **86**, 5954 (2001).
[45]T. Klüner, N. Govind, Y. A. Wang, and E. A. Carter, J. Chem. Phys. **116**, 42 (2002).
[46]A. S. P. Gomes, C. R. Jacob, and L. Visscher, Phys. Chem. Chem. Phys. **10**, 5353 (2008).
[47]S. Höfener, A. S. P. Gomes, and L. Visscher, J. Chem. Phys. **136**, 044104 (2012).
[48]J. Neugebauer, ChemPhysChem **10**, 3148 (2009).
[49]T. A. Wesołowski, Chem. Phys. Lett. **311**, 87 (1999).
[50]T. A. Wesołowski and F. Tran, J. Chem. Phys. **118**, 2072 (2003).
[51]C. R. Jacob, T. A. Wesolowski, and L. Visscher, J. Chem. Phys. **123**, 174104 (2005).
[52]S. Fux, K. Kiewisch, C. R. Jacob, J. Neugebauer, and M. Reiher, Chem. Phys. Lett. **461**, 353 (2008).
[53]K. Kiewisch, G. Eickerling, M. Reiher, and J. Neugebauer, J. Chem. Phys. **128**, 044114 (2008).
[54]S. M. Beyhan, A. W. Götz, C. R. Jacob, and L. Visscher, J. Chem. Phys. **132**, 044114 (2010).
[55]TURBOMOLE v5.10 2008, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available on http://www.turbomole.com.
[56]O. Treutler and R. Ahlrichs, J. Chem. Phys. **102**, 346 (1995).
[57]M. von Arnim and R. Ahlrichs, J. Comput. Chem. **19**, 1746 (1998).
[58]M. Ernzerhof and G. Scuseria, J. Chem. Phys. **110**, 5029 (1999).
[59]S. Grimme, J. Comput. Chem. **25**, 1463 (2004).
[60]K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, Theor. Chem. Acc. **97**, 119 (1997).
[61]X. Cao and M. Dolg, J. Mol. Struct.: THEOCHEM **673**, 203 (2004).
[62]A. Nicklass, M. Dolg, H. Stoll, and H. Preuss, J. Chem. Phys. **102**, 8942 (1995).
[63]E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
[64]M. E. Casida, *Recent Advances in Density Functional Methods, Part I* (Word Scientific, 1995).
[65]ADF2010.01, SCM, Theoretical chemistry, Vrije Universiteit, Amsterdam, The Netherlands, see http://www.scm.com.
[66]G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. F. Guerra, E. J. Baerends, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).
[67]C. F. Guerra, J. G. Snijders, G. te Velde, and E. J. Baerends, Theor. Chem. Acc. **99**, 391 (1998).
[68]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[69]E. van Lenthe and E. J. Baerends, J. Comput. Chem. **24**, 1142 (2003).
[70]E. van Lenthe, E. J. Baerends, and J. G. Snijders, J. Chem. Phys. **99**, 4597 (1993).
[71]P. Cortona, Phys. Rev. B **44**, 8454 (1991).
[72]P. Cortona, Phys. Rev. B **46**, 2008 (1992).
[73]G. Senatore and K. R. Subbaswamy, Phys. Rev. B **34**, 5754 (1986).
[74]T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).
[75]T. A. Wesolowski and J. Weber, Chem. Phys. Lett. **248**, 71 (1996).
[76]C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, J. Comput. Chem. **32**, 2328 (2011).
[77]A. W. Götz, S. M. Beyhan, and L. Visscher, J. Chem. Theory Comput. **5**, 3161 (2009).
[78]C. R. Jacob, S. M. Beyhan, and L. Visscher, J. Chem. Phys. **126**, 234116 (2007).
[79]S. Laricchia, E. Fabiano, and F. D. Sala, J. Chem. Phys. **133**, 164111 (2010).
[80]M. Pavanello and J. Neugebauer, J. Chem. Phys. **135**, 134113 (2011).
[81]L. H. Thomas, Math. Proc. Cambridge Philos. Soc **23**, 542 (1927).
[82]I. von Fermi, Z. Phys. **36**, 73 (1928).
[83]E. H. Lieb, Rev. Mod. Phys. **53**, 603 (1981).
[84]J. M. G. Lastra, J. W. Kamiński, and T. A. Wesolowski, J. Chem. Phys. **129**, 074107 (2008).
[85]A. Lembarki and H. Chermette, Phys. Rev. A **50**, 5328 (1994).
[86]M. E. Casida and T. A. Wesolowski, Int. J. Quantum Chem. **96**, 577 (2004).
[87]J. Neugebauer, J. Phys. Chem. B **112**, 2207 (2008).
[88]T. Saue, L. Visscher, and H. J. Aa. Jensen *et al.*, DIRAC, a relativistic *ab initio* electronic structure program, release DIRAC10 (2010), see http://dirac.chem.vu.nl.
[89]L. Visscher, Theor. Chem. Acc. **98**, 68 (1997).
[90]K. G. Dyall, Theor. Chem. Acc. **491**, 483 (2007).
[91]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[92]I. Lindgren and D. Mukherjee, Phys. Rep. **151**, 93 (1987).
[93]D. I. Lyakh, M. Musiał, V. F. Lotrich, and J. Bartlett, Chem. Rev. **112**, 182 (2012).
[94]L. Meissner, J. Chem. Phys. **103**, 8014 (1995).
[95]L. Meissner, J. Chem. Phys. **108**, 9227 (1998).
[96]K. Jankowski, L. Meissner, and K. Rubiniec, Int. J. Quantum Chem. **67**, 239 (1998).
[97]L. Meissner and P. Malinowski, Phys. Rev. A **61**, 062510 (2000).
[98]L. Meissner, P. Malinowski, and A. Nowaczyk, J. Chem. Phys. **116**, 7362 (2002).
[99]F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet, and E. Eliav, J. Phys. Chem. A **113**, 12504 (2009).
[100]S. Yabushita, Z. Zhang, and R. M. Pitzer, J. Phys. Chem. A **103**, 5791 (1999).
[101]M. J. G. Peach, M. J. Williamson, and D. J. Tozer, J. Chem. Theory Comput. **7**, 3578 (2011).
[102]C. Clavaguéra-Sarrio, N. Ismail, C. J. Marsden, D. Bégue, and C. Pouchan, Chem. Phys. **302**, 1 (2004).
[103]M. Dułak and T. A. Wesołowski, J. Chem. Phys. **124**, 164101 (2006).
[104]Y. A. Bernard, M. Dułak, J. W. Kamiński, and T. A. Wesołowski, J. Phys A: Math. Theor. **41**, 055302 (2008).
[105]J. Čížek and J. Paldus, J. Chem. Phys. **47**, 3976 (1967).
[106]O. B. Lutnæs, T. Helgaker, and M. Jaszuński, Mol. Phys. **108**, 2579 (2010).
[107]ADFGUI 2011, SCM, Amsterdam, The Netherlands, see http://www.scm.com.
[108]B. E. Bursten, M. L. Drummond, and J. Li, Faraday Discuss. **124**, 1 (2003).
[109]S. O. Odoh and G. Schreckenbach, J. Phys. Chem. A **114**, 1957 (2010).
[110]See supplementary material at http://dx.doi.org/10.1063/1.4742765 for the precise definition of the active space used in the IHFSCC-SD calculations.

## D.5 Paper XII

THE JOURNAL OF CHEMICAL PHYSICS **139**, 104106 (2013)

# Solvatochromic shifts from coupled-cluster theory embedded in density functional theory

Sebastian Höfener,[1] André Severo Pereira Gomes,[2] and Lucas Visscher[1]
[1]*Amsterdam Center for Multiscale Modelling (ACMM), VU University Amsterdam, Theoretical Chemistry Section, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands*
[2]*Laboratoire PhLAM, CNRS UMR 8523, Université de Lille 1, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France*

(Received 6 June 2013; accepted 23 August 2013; published online 11 September 2013)

Building on the framework recently reported for determining general response properties for frozen-density embedding [S. Höfener, A. S. P. Gomes, and L. Visscher, J. Chem. Phys. **136**, 044104 (2012)], in this work we report a first implementation of an embedded coupled-cluster in density-functional theory (CC-in-DFT) scheme for electronic excitations, where only the response of the active subsystem is taken into account. The formalism is applied to the calculation of coupled-cluster excitation energies of water and uracil in aqueous solution. We find that the CC-in-DFT results are in good agreement with reference calculations and experimental results. The accuracy of calculations is mainly sensitive to factors influencing the correlation treatment (basis set quality, truncation of the cluster operator) and to the embedding treatment of the ground-state (choice of density functionals). This allows for efficient approximations at the excited state calculation step without compromising the accuracy. This approximate scheme makes it possible to use a first principles approach to investigate environment effects with specific interactions at coupled-cluster level of theory at a cost comparable to that of calculations of the individual subsystems in vacuum. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4820488]

## I. INTRODUCTION

Electronic excitations of valence electrons play a central role in photochemical and photophysical processes. To understand the properties of species in a condensed phase, one may consider two extremes: one in which there is little interaction between the constituents and the spectra are close to those in vacuum; and the other where interactions between the species and its environment are strong enough to produce qualitatively different spectra.

The most direct way to model spectra in the latter case is to employ a supermolecular approach and create a model including the species of interest as well as the closest neighbor molecules in the environment. This model is then treated with one of the standard electronic structure methods, such as the generalizations to the time-dependent case of density-functional theory (TD-DFT), see, e.g., Ref. 1 and references therein, or wave function theory (WFT)-based methods as in the linear response (LR) coupled-cluster (CC) approach.[2]

While coupled-cluster methods are often preferable to TD-DFT due to better accuracy and reliability, their rather high computational costs limit their applicability. For instance, TD-DFT exhibits a formal computational scaling of up to $N^4$, where $N$ is a measure of the system size, but in practice it is as low as $N^2$,[3] so that single-point calculations on systems of up to 400 atoms[4] can be carried out routinely. By contrast, coupled-cluster with singles and doubles excitations (CCSD) exhibits a formal computational scaling of $N^6$, so that molecules up to some tens of atoms are feasible.[5] Certain approximate methods can show more favorable scalings, e.g., $N^4$ in case of a scaled opposite spin (SOS)-Laplace CC2 of Winter and Hättig,[6] but even in these "best-case" scenarios systems only up to about 150 atoms can be routinely treated, in particular if excited-state properties are sought. Effectively, this prevents their use in models in which a large number of solvent molecules (e.g., a few hundred or more) must be taken into account in order to describe long-range effects.

An alternative to the supermolecule approach is found in embedding approaches, where instead of treating the whole system as a single entity, it is partitioned into a number of interacting subsystems. While from a formal perspective one can formulate embedding approaches equivalent to the supermolecular case, for instance by considering the case of (TD-) DFT with exact functionals,[7] their real strength lies in the possibility of introducing systematic approximations allowing for reducing the overall computational cost while maintaining a desired degree of accuracy.

Among the embedding approaches for calculating electronic spectra, those in the QM/MM (quantum mechanics/molecular mechanics) family are perhaps the most widely used in connection to model electronic spectra in solution so far. In QM/MM, the system is partitioned into an active center to be calculated with QM methods, and its environment represented by a classical force field, so that the interaction between the two is obtained as an interaction between a QM density and a representation in terms of multipole expansions. Combined with electronic structure methods based on response theory,[8–11] polarizable force fields have been shown to incorporate effects such as the polarization of the environment as well as its back-polarization effect on the ground and excited states of the solute rather accurately,[12,13] provided

that reliable parameters can be generated for the system to be studied.

A dependence on force fields can be circumvented by employing purely quantum-mechanical embedding approaches[7] (denoted QM/QM) such as frozen-density embedding (FDE). Originally formulated within a ground-state DFT framework[14–16] and later extended to the time-dependent domain,[17–21,28] FDE is based on the partitioning of the total electron density into that of subsystems, and allows to describe the electronic structure of one subsystem with the others represented by a so-called embedding potential.

In the FDE framework it is also possible to combine WFT and DFT, obtaining a methodology that can be accurate and relatively inexpensive as the WFT-based method can be used to describe only a subsystem of interest, while the others and the interaction between subsystems are treated at DFT level. This approach was followed in the WFT-in-DFT method pioneered by Carter and co-workers,[22] who employed variational approaches such as multireference (MR) configuration interaction (CI), obtaining the embedding potential in an iterative manner. Recently, a less costly computational approach was presented, in which an embedding potential obtained from purely DFT calculations is used as a fixed one-electron potential in the WFT calculation, allowing for the use of electronic structure methods for which electron densities are not easily available such as non-variational methods.[23] These WFT-in-DFT approaches typically account for environment effects, including the polarization of the environment, for the ground-state but do not include polarization effects due to electronic excitations that can become important, depending on the property of interest and on the nature of the system.[18,20,24,25]

In this article, we therefore aim to provide a first step towards the incorporation of excited-state polarizations in WFT-in-DFT calculations, akin to what is currently possible with purely DFT approaches, by providing a first implementation and applications of a CC-in-DFT approach based on the generalization[26] of the FDE subsystem response formalism. In this paper, we restrict the discussion to the response of the active subsystem treated at the CC level, whereas the response of the environment treated with DFT is disregarded. This means that excited-state polarization within the active subsystem will be taken into account for the first time in a CC-in-DFT calculation. In future work we shall address the coupling between subsystems in the excited states.

This article is organized as follows: it starts by summarizing the essential aspects of FDE-DFT and CC-in-DFT, followed by a discussion of the calculation of the individual contributions to, e.g., the coupled-cluster Jacobian. After that numerical results for the influence of different levels of approximations on excitation energies are discussed, followed by conclusions and outlook.

## II. THEORY

In this section the essentials of DFT-in-DFT and CC-in-DFT embedding are outlined, and the reader is referred to recent publications[7,27,28] for further details.

### A. FDE and subsystem DFT

The starting point of DFT FDE is the representation of the ground-state electron density for the total system, $\rho_{\text{tot}}$, as a sum of the densities of a number of subsystems:

$$\rho_{\text{tot}} = \rho_{\text{I}} + \sum_{i \neq \text{I}} \rho_i = \rho_{\text{I}} + \rho_{\text{II}}, \tag{1}$$

where $\rho_{\text{I}}$ denotes the density for the so-called "active" subsystem of interest, and $\rho_{\text{II}}$ the density of its surrounding, denoted the "frozen" subsystem. With such a partitioning, the total energy of the system, $E_{\text{tot}}[\rho_{\text{tot}}]$, can then be written as a sum of subsystem energies and one interaction energy,

$$E_{\text{tot}}[\rho_{\text{I}} + \rho_{\text{II}}] = E_{\text{I}}[\rho_{\text{I}}] + E_{\text{II}}[\rho_{\text{II}}] + E_{\text{int}}[\rho_{\text{I}}, \rho_{\text{II}}]. \tag{2}$$

The subsystem energy expression for each of the subsystems is given as

$$E_i[\rho_i] = \int \rho_i(\mathbf{r}) v^i_{\text{nuc}}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho_i(\mathbf{r})\rho_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}\, d\mathbf{r}' + E_{\text{xc}}[\rho_i] + T_s[\rho_i] + E^i_{\text{nuc}}, \tag{3}$$

where $v^i_{\text{nuc}}$ is the nuclear potential and $E^i_{\text{nuc}}$ the nuclear repulsion energy. The interaction energy is given by the expression

$$E_{\text{int}}[\rho_{\text{I}}, \rho_{\text{II}}] = \int \rho_{\text{I}}(\mathbf{r}) v^{\text{II}}_{\text{nuc}}(\mathbf{r}) d\mathbf{r} + \int \rho_{\text{II}}(\mathbf{r}) v^{\text{I}}_{\text{nuc}}(\mathbf{r}) d\mathbf{r} + E^{\text{I,II}}_{\text{nuc}} + \iint \frac{\rho_{\text{I}}(\mathbf{r})\rho_{\text{II}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E^{\text{nadd}}_{\text{xck}}[\rho_{\text{I}}, \rho_{\text{II}}], \tag{4}$$

where the non-additive contributions are defined as

$$E^{\text{nadd}}_{\text{xck}}[\rho_{\text{I}}, \rho_{\text{II}}] = E_{\text{xck}}[\rho_{\text{I}} + \rho_{\text{II}}] - E_{\text{xck}}[\rho_{\text{II}}] - E_{\text{xck}}[\rho_{\text{II}}]. \tag{5}$$

Here, $E_{\text{xck}}[\rho]$ denotes the sum of exchange-correlation and kinetic-energy contributions, which are calculated using density functionals. The explicit dependence on the density gradients of the different densities in Eq. (5) is omitted for clarity and is taken into account for, e.g., generalized gradient approximation (GGA) functionals.

Minimizing the total energy of the system with respect to $\rho_{\text{I}}$ yields the Euler-Lagrange equation,

$$\frac{\delta E_{\text{I}}[\rho_{\text{I}}]}{\delta \rho_{\text{I}}} + \frac{\delta E_{\text{int}}[\rho_{\text{I}}, \rho_{\text{II}}]}{\delta \rho_{\text{I}}} = \mu_{\text{I}}, \tag{6}$$

with the Lagrange multiplier $\mu_{\text{I}}$ introduced to keep the number of electrons in system I constant. The derivative of the interaction energy functional that carries the intersystem dependence is the embedding potential

$$v^{\text{I}}_{\text{int}}(\mathbf{r}) = \frac{\delta E_{\text{int}}[\rho]}{\delta \rho_{\text{I}}(\mathbf{r})} = v^{\text{II}}_{\text{nuc}}(\mathbf{r}) + \int \frac{\rho_{\text{II}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \left[ \frac{\delta E_{\text{xck}}}{\delta \rho} \Big|_{\rho_{\text{tot}}} - \frac{\delta E_{\text{xck}}}{\delta \rho} \Big|_{\rho_{\text{I}}} \right]. \tag{7}$$

It can be decomposed into the electrostatic interactions with the environment due to the nuclei and electron density of the frozen subsystem, as well as derivatives of the non-additive exchange-correlation and kinetic energy contributions. The density of subsystem I is obtained by constructing

a non-interacting reference system and employing the Kohn-Sham equation for a constrained electron density (KSCED),[16] yielding

$$F^{\mathrm{KS;I}}\phi_p^I(\mathbf{r}) = \Bigg[ -\frac{1}{2}\Delta + v_{\mathrm{nuc}}^I(\mathbf{r}) + \int \frac{\rho_{\mathrm{I}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{\mathrm{xc}}[\rho_I](\mathbf{r}) + v_{\mathrm{int}}^I(\mathbf{r}) \Bigg] \phi_p^I(\mathbf{r}) = \varepsilon_p^{\mathrm{KS;I}} \phi_p^I(\mathbf{r}). \quad (8)$$

In order for the subsystem densities to satisfy Eq. (1), the frozen density $\rho_{\mathrm{II}}$ must be smaller than $\rho_{\mathrm{tot}}$ everywhere and $v$-representable.[29] As these conditions often cannot be fulfilled in practice if $\rho_{\mathrm{II}}$ remains fixed, a solution is to minimize the total energy with respect to $\rho_{\mathrm{I}}$ as well as $\rho_{\mathrm{II}}$.

This implies that instead of a single Euler-Lagrange equation, one is dealing with a system of equations coupled through the interaction term. The most widely used procedure to solve such a system is the so-called freeze-thaw scheme, an iterative procedure where one solves for one of the subsystem densities at a time, considering the others frozen and subsequently exchanging roles until self-consistency. The freeze-thaw procedure can be therefore be thought of as a way to polarize the environment at the ground-state.

FDE and DFT-in-DFT have been shown to be very efficient in the calculation of excitation energies in the presence of rather large environments, and FDE in particular can be easily deployed in combination with molecular dynamics (MD).[30–32] Central to this efficiency is the use of disjoint sets of molecular orbitals (MOs) for the different subsystems, which greatly reduces the computational effort in calculating two-electron integrals, and the use of pure density functional approximations to the exact kinetic energy functional contributing to the interaction energy and its derivatives.

These two key aspects are at the same time major sources of numerical differences to the supermolecule treatment. The use of disjoint sets of orbitals, and the usual restriction that the subsystems do not exchange particles, may prevent a proper description of electronic excitations for which there are significant contributions form the environment, unless one also couples the subsystems in the excited states.[18–20,28] The currently available kinetic density functionals, in their turn, are only sufficiently accurate to describe relatively weak interactions[29,33,34] and impose limitations on the possible partitioning into subsystems, e.g., preventing partitions that separate atoms making covalent bonds, unless special care is taken.[35] Such shortcomings have motivated the development of approaches circumventing the use of kinetic energy functionals in the construction of the embedding potential,[29,36–39] but these schemes are not yet mature enough to be used in the calculation of excited states.

### B. The coupled-cluster ansatz

The discussion of the theoretical aspects in the present work is restricted to the orbital-unrelaxed singles and doubles ansatz (CCSD), where the cluster operator $T$ is truncated after two-particle excitations, and approximations to it.[40] The starting point is the definition of the CCSD Lagrangian $\mathcal{L}$ as the energy contribution obtained from the projection of the Schrödinger equation to different excited determinants:

$$\mathcal{L} = \langle \mathrm{HF}| \exp(-T_2)\hat{H}\exp(T_2)|\mathrm{HF}\rangle + \sum_{\mu_i} \bar{t}_{\mu_i}\Omega_{\mu_i}, \quad (9)$$

$$\Omega_{\mu_i} = \langle \mu_i| \exp(-T_2)\hat{H}\exp(T_2)|\mathrm{HF}\rangle. \quad (10)$$

Here, $T_2$ denotes the double-excitation cluster operator and $\mu_i$ the excited-determinant excitation parameterization. In this article, hats label $T_1$-similarity transformed quantities, e.g., in case of the Hamilton operator:

$$\hat{H} = \exp(-T_1)H\exp(T_1). \quad (11)$$

As an approximate method, the CC2 model is considered in the present work, for which the amplitude equations for the single excitations ($\Omega_{\mu_1}$) remain as in CCSD, but those defining the doubles excitations ($\Omega_{\mu_2}$) are simplified in such a way that the fluctuation potential is only transformed with the $T_1$ cluster operator:

$$\Omega_{\mu_2}^{\mathrm{CC2}} = \langle \mu_2|[F, T_2] + \hat{\Phi}|\mathrm{HF}\rangle, \quad (12)$$

where $H$ has been decomposed into the Fock operator $F$ and the fluctuation potential $\Phi$:

$$H = F + \Phi. \quad (13)$$

### C. The CC-in-DFT response ansatz

In the CC-in-DFT response ansatz,[26] instead of the expression for the total energy in Eq. (2), a Lagrangian $L$ for the total system is defined, constructed from the Lagrangian of subsystem I ($\mathcal{L}_{\mathrm{I}}$), as well as that for subsystem II ($\mathcal{L}_{\mathrm{II}}$) and the interaction contribution ($Q_{\mathrm{int}}$):

$$L = \mathcal{L}_{\mathrm{I}} + Q_{\mathrm{int}} + \mathcal{L}_{\mathrm{II}}. \quad (14)$$

$Q_{\mathrm{int}}$ has the same form of as $E_{\mathrm{int}}$ in Eq. (4), but now both $\rho_{\mathrm{I}}$ and $\rho_{\mathrm{II}}$ can be time-dependent. The electron density for subsystem I is given as

$$\rho_{\mathrm{I}} = \rho_{\mathrm{CC}} = \langle \Lambda|\rho^{\mathrm{I}}|\mathrm{CC}\rangle, \quad (15)$$

with the bra state defined as

$$\langle \Lambda| = \langle \mathrm{HF}| + \sum_{\nu} \bar{t}_{\nu}\langle \nu|\exp(-T). \quad (16)$$

If the frozen-core approximation is applied in this context, the density of the corresponding core orbitals is taken from the Hartree-Fock level of theory for the construction of the embedding potential.

### D. Ground-state amplitudes and multipliers in CC-in-DFT

The coupled-cluster ground-state amplitudes and Lagrange multipliers are obtained from the condition that the zeroth-order quasienergy Lagrangian is stationary with respect to them.[40] However, embedding coupled-cluster approaches differ from a conventional treatment in the fact that these parameters cannot be determined independently.[12,26,41]

TABLE I. Ground-state quantities for CC-in-DFT. Superscripts in parenthesis denote the order of perturbation, $\{\mu, \nu\}$ refer to singly or doubly excited determinants, $\hat{T}^{(0)}$ cluster operator built with ground-state amplitudes, $E_{pq}$ is a spin-free excitation operator, and $\{p, q\}$ run over orbital indices. Integrals over $v_{\text{int}}$ and $w_{\text{int}}$ are defined in Eqs. (18) and (27), respectively. See Ref. 26 and references therein for further details.

| | | | Contributions | |
|---|---|---|---|---|
| Quantity | determined from | | Vacuum | Embedding |
| $\mathbf{t}^{(0)}$ | $0 \stackrel{!}{=} \frac{\partial\{L^{(0)}\}_T}{\partial \bar{\mathbf{t}}_\nu^{(0)}}$ | | $\Omega_\nu$ | $\sum_{pq}\langle p\|v_{\text{int}}\|q\rangle {}^{\bar{t}}\Gamma_{pq;\nu}$ |
| $\bar{\mathbf{t}}^{(0)}$ | $0 \stackrel{!}{=} \frac{\partial\{L^{(0)}\}_T}{\partial \mathbf{t}_\mu^{(0)}}$ | | $\sum_\nu \bar{t}_\nu^{(0)} \frac{\partial\Omega_\nu}{\partial t_\mu^{(0)}} + {}^t\eta_\mu$ | $\sum_{pq}\langle p\|v_{\text{int}}\|q\rangle {}^t\Gamma_{pq;\mu}$ |
| | | $\equiv$ | $\sum_\nu \bar{t}_\nu^{(0)} \mathbf{A}_{\mu\nu}^{\text{vac}} + \eta_\mu^{\text{vac}}$ | $\sum_\nu \bar{t}_\nu^{(0)} \mathbf{A}_{\mu\nu}^{\text{int}} + \eta_\mu^{\text{int}}$ |
| $\mathbf{A}_{\mu\nu}$ | | | $\langle\nu\|e^{-\hat{T}^{(0)}}[\hat{H}, \tau_\mu]\|\text{CC}\rangle$ | $\sum_{pq}\langle p\|v_{\text{int}}\|q\rangle\langle\nu\|e^{-\hat{T}^{(0)}}[E_{pq}, \tau_\mu]\|\text{CC}\rangle$ |
| $\boldsymbol{\eta}_\mu$ | | | $\langle\text{HF}\|e^{-\hat{T}^{(0)}}[\hat{H}, \tau_\mu]e^{\hat{T}^{(0)}}\|\text{HF}\rangle$ | $\sum_{pq}\langle p\|v_{\text{int}}\|q\rangle \times$ $\langle\text{HF}\|e^{-\hat{T}^{(0)}}[E_{pq}, \tau_\mu]e^{\hat{T}^{(0)}}\|\text{HF}\rangle$ |
| ${}^t\Gamma_{pq;\mu}$ | $\frac{\partial\rho_{\text{CC}}^{(0)}}{\partial \bar{t}_\mu^{(0)}}$ | | | $\langle\Lambda\|[E_{pq}, \tau_\mu]\|\text{CC}\rangle$ |
| ${}^{\bar{t}}\Gamma_{pq;\mu}$ | $\frac{\partial\rho_{\text{CC}}^{(0)}}{\partial t_\mu^{(0)}}$ | | | $\langle\mu\|e^{-\hat{T}^{(0)}}E_{pq}\|\text{CC}\rangle$ |

#### 1. Amplitude equations

The coupled-cluster amplitudes are obtained from the derivative of the Lagrangian with respect to the Lagrange multipliers. In the following, the short-hand notation of Ref. 26 will be used, yielding

$$0 = \frac{\partial L}{\partial \bar{t}_{\mu_i}} = \Omega_{\mu_i} + \sum_{pq}\langle p|v_{\text{int}}|q\rangle {}^{\bar{t}}\Gamma_{pq;\mu_i}, \tag{17}$$

where

$$\langle p|v_{\text{int}}|q\rangle = \int \phi_p^\dagger(\mathbf{r}) v_{\text{int}}^{\text{I}}(\mathbf{r})\phi_q(\mathbf{r})d\mathbf{r}, \tag{18}$$

and ${}^{\bar{t}}\Gamma_{pq;\mu_i}$ is a density matrix, obtained by expanding the CC density to different orders and collecting the contributions of same order. Hence, in this case one obtains

$${}^{\bar{t}}\Gamma_{pq;\mu_i} = \langle\mu_i|\exp(-T)E_{pq}|\text{CC}\rangle. \tag{19}$$

The elements of ${}^{\bar{t}}\Gamma_{pq;\mu_i}$ can be easily obtained, since they exhibit the same structure as the conventional amplitude equations. At the level of working equations, similar contributions are obtained.

#### 2. Lagrangian multipliers

The Lagrangian multipliers are obtained from the derivative of the Lagrangian with respect to the amplitudes:

$$0 = \frac{\partial L}{\partial t_{\mu_i}} = \langle\Lambda|[H, \tau_{\mu_i}]|\text{CC}\rangle + \sum_{pq}\langle p|v_{\text{int}}|q\rangle {}^t\Gamma_{pq;\mu_i}, \tag{20}$$

where the density matrix elements ${}^t\Gamma_{pq;\mu_i}$ are given as

$${}^t\Gamma_{pq;\mu_i} = \langle\Lambda|[E_{pq}, \tau_{\mu_i}]|\text{CC}\rangle. \tag{21}$$

It is possible to rearrange Eq. (20) so that it can be expressed in matrix form as

$$-(\boldsymbol{\eta}_{\text{vac}} + \boldsymbol{\eta}_{\text{int}}) = \bar{\mathbf{t}}\left(\mathbf{A}^{\text{vac}} + {}^t\mathbf{A}\right). \tag{22}$$

Expressions for $\mathbf{A}$ and $\boldsymbol{\eta}$ can be found in Table I.

#### 3. Computational protocol

In the FDE approach, the electron densities for the subsystems are the central quantities, as they fully determine the embedding potential $v_{\text{int}}$. Thus, for CC-in-DFT it is the ground-state coupled-cluster density $\rho_{\text{CC}}$ that takes this central role, and, since $\rho_{\text{CC}}$ has a non-linear dependence on the set of ground-state CC amplitudes and multipliers, one has to employ an iterative procedure akin to the self-consistent field procedure for its determination.

The focus in the present work is on the polarization of the coupled-cluster density due to a frozen environment, and the options available for this step are illustrated in the computational protocol in Fig. 1. Thus, one starts by obtaining (step (a) in Fig. 1) the electrostatic potential and electron density $\rho_{\text{II}}$ for the environment, where the latter can be determined either from a calculation in the original DFT FDE formulation,[16] or from calculations in which it is allowed to be polarized via a freeze-thaw procedure.[22,23] Following that, one can follow two approaches. The first one consists of performing a calculation of subsystem I without any environment (in vacuum, step (b)), in order to obtain a trial $\rho_{\text{CC}}$ (step (c)) used to construct a trial embedding potential (step (d)). The $\rho_{\text{CC}}$ and $v_{\text{int}}$ are subsequently updated in a self-consistent manner (steps (c)–(e)) but only at the coupled cluster level, without modifying the (vacuum) MOs. This approach emphasizes the fact that in WFT-in-DFT, it is the correlated electron density that interacts with the DFT density and regards the Hartree-Fock orbitals only as auxiliary quantities.

In the second approach, the embedding potential is included in the Hartree-Fock step (b) at every iteration of the

(a) Obtain a reasonable environment density $\rho_{\mathrm{II}}$, store environment density and its electrostatic (nuclear + Coulomb) potential

(b) Calculate orbitals of subsystem I

(c) Calculate coupled-cluster density of subsystem I

(d) Update embedding potential

(e) Goto (b) or (c) until converged

(f) Calculate desired properties

FIG. 1. Computational protocol for ground-state CC-in-DFT with an active system in the presence of a frozen environment. In the first iteration, calculations in vacuum are carried out, whereas in subsequent calculations the embedding potential is included for subsystem I.

self-consistent procedure to determine $v_{\mathrm{int}}$ (steps (b)–(e)), using, for instance, the $v_{\mathrm{int}}$ from a prior embedding calculation in step (a) as an initial guess. In this way, the MOs are modified at each iteration by the interaction with the environment, and therefore can be though of as being relaxed with respect to (or perturbed by) the environment. The motivation for this approach comes from the observation that in CC/MM[12] calculations introducing the interaction potential at the self-consistent field (SCF) stage can sensibly improve the calculated excitation energies by providing some additional orbital relaxation, thus avoiding the need to introduce it at the linear response level. This enables the method to retain a pole structure close to exact theory. In addition to a possibly better description of orbital relaxation, the use of the same embedding potential in the SCF and coupled-cluster treatment for the second procedure allows for further simplifications in the working equations. Important for the computational efficiency is that this scheme makes it possible to employ a T2-direct algorithm without further approximations for CC2 – something that would not be justified if the embedding contributions were split into SCF and CC contributions.

With respect to their computational costs, the two approaches have similar formal computational scalings. Let $N_{\mathrm{I}}$ and $N_{\mathrm{II}}$ be a measure for the size of the active and environment subsystem, respectively. For step (a), the cost of DFT FDE or subsystem DFT calculations will grow at most as $O(n_{\mathrm{I}}[N_{\mathrm{I}}^4 + N_{\mathrm{II}}^4])$, where $n_i$ denotes the number of freeze-thaw iterations (and are equal to one for DFT FDE). In turn, for steps (b)–(e), the CC-in-DFT calculations will scale as $O(m_{\mathrm{I}}N_{\mathrm{I}}^6 + \delta)$ for CCSD or $O(m_{\mathrm{I}}N_{\mathrm{I}}^5 + \delta)$ for CC2, where $m_{\mathrm{I}}$ is the number of steps required to achieve self-consistency for $\rho_{\mathrm{CC}}$ and $v_{\mathrm{int}}$ and $\delta$ indicates the overhead, e.g., due to the calculation of the embedding potentials. It is nevertheless difficult to estimate which of the two approaches is faster in practice, as the convergence behavior of the self-consistent procedure is not necessarily the same for the two and therefore the number of iterations $m_{\mathrm{I}}$ might differ.

It is clear, however, that CC-in-DFT scales much more favorably than coupled cluster supermolecular calculations, whose cost grows as $O([N_{\mathrm{I}} + N_{\mathrm{II}}]^6)$ for CCSD or $O([N_{\mathrm{I}} + N_{\mathrm{II}}]^5)$ for CC2, and will be computationally advantageous whenever the environment consists of more than a few molecules. For relatively small systems, on the other hand, it may be that the overhead associated with the embedding calculation offsets the computational advantage of treating smaller subsystems, and supermolecular calculations remain competitive.

### E. CC-in-DFT excitation energies

As discussed in Ref. 26, excitation energies in the subsystem formulation correspond to the poles of the linear response function,

$$\langle\langle X;Y\rangle\rangle_{\omega} = \frac{d^2\{^{2n+1}\mathcal{L}_{\mathrm{I}}^{(2)}\}_T}{d\epsilon_X(\omega)d\epsilon_Y(-\omega)} + \frac{d^2\{^{2n+1}Q_{\mathrm{int}}^{(2)}\}_T}{d\epsilon_X(\omega)d\epsilon_Y(-\omega)} + \frac{d^2\{^{2n+1}\mathcal{L}_{\mathrm{II}}^{(2)}\}_T}{d\epsilon_X(\omega)d\epsilon_Y(-\omega)}. \tag{23}$$

The contributions from $\{^{2n+1}\mathcal{L}_{\mathrm{II}}^{(2)}\}_T$, representing the response of the environment, will not be taken into account here. In this case, the (orbital-unrelaxed) CC-in-DFT linear response function becomes (compare Table II)

$$\langle\langle X;Y\rangle\rangle_{\omega} = \frac{1}{2}C^{\pm\omega}P\left(X(\omega),Y(-\omega)\right)\left\{\frac{1}{2}\mathbf{J}\,\bar{\mathbf{t}}^X(\omega)\bar{\mathbf{t}}^Y(-\omega) + \left[\boldsymbol{\eta}^X + \frac{1}{2}\mathbf{F}\,\mathbf{t}^X(\omega)\right]\mathbf{t}^Y(-\omega)\right\}. \tag{24}$$

with $C^{\pm\omega}$ a symmetrization operator[40] and $\bar{\mathbf{t}}^{X,Y}(\pm\omega)$, $\mathbf{t}^{X,Y}(\pm\omega)$ the first-order Lagrange multipliers and CC amplitudes, respectively, obtained by solving the response equations,[26,40] and $\mathbf{F}$ and $\mathbf{J}$ are matrices defined on the basis of subsystem I (see Table II).

For the vacuum case,[40] the matrix ${}^{\bar{t}t}\mathbf{J}$ is identically zero and it is straightforward to show that the poles of the CC response function are found to occur at the same frequencies as the eigenvalues of the CC Jacobian matrix ${}^{\bar{t}t}\mathbf{A}$. For CC-in-DFT as well as CC/MM, on the other hand, $\mathbf{J}$ may be

TABLE II. Linear response quantities for CC-in-DFT. Superscripts in parenthesis denote the order of perturbation, $\{\mu, \nu\}$ refer to singly or doubly excited determinants, $\hat{T}^{(0)}$ cluster operator built with ground-state amplitudes, $E_{pq}$ is a spin-free excitation operator, and $\{p, q\}$ run over orbital indices. Integrals over $v_{\text{int}}$ and $w_{\text{int}}$ are defined in Eqs. (18) and (27), respectively. See Ref. 26 and references therein for further details.

| | | Contributions | |
|---|---|---|---|
| Quantity | determined from | Vacuum | Embedding |
| $\eta^X_\mu$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\epsilon_X(\omega)\partial t^{(1)}_\mu(-\omega)}$ | $\langle\Lambda\vert[\hat{X},\tau_\mu]\vert\text{CC}\rangle$ | |
| ${}^{tt}F_{\mu\nu}$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial t^{(1)}_\mu(\omega)\partial t^{(1)}_\nu(-\omega)}$ | $\langle\Lambda\vert[[\hat{H},\tau_\mu],\tau_\nu]\vert\text{CC}\rangle$ | $\sum_{pq}\langle p\vert v_{\text{int}}\vert q\rangle^{tt}\Gamma_{pq;\mu\nu}$ |
| | | | $+\frac{1}{2}\sum_{pqrs}(pq\vert w_{\text{int}}\vert rs)^{t}\Gamma_{pq;\mu}{}^{t}\Gamma_{rs;\nu}$ |
| ${}^{t\bar{t}}A_{\mu\nu}(\omega)$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial t^{(1)}_\mu(\omega)\partial\bar{t}^{(1)}_\nu(-\omega)}$ | $\langle\nu\vert e^{-\hat{T}^{(0)}}[\hat{H},\tau_\mu]\vert\text{CC}\rangle$ | $\sum_{pq}\langle p\vert v_{\text{int}}\vert q\rangle^{t\bar{t}}\Gamma_{pq;\mu\nu}$ |
| | | | $+\frac{1}{2}\sum_{pqrs}(pq\vert w_{\text{int}}\vert rs)^{t}\Gamma_{pq;\mu}{}^{\bar{t}}\Gamma_{rs;\nu}$ |
| ${}^{\bar{t}\bar{t}}J_{\mu\nu}$ | $\frac{\partial^2\{L^{(2)}\}_T}{\partial\bar{t}^{(1)}_\mu(\omega)\partial\bar{t}^{(1)}_\nu(-\omega)}$ | $\mathbf{0}$ | $\frac{1}{2}\sum_{pqrs}(pq\vert w_{\text{int}}\vert rs)^{\bar{t}}\Gamma_{pq;\mu}{}^{\bar{t}}\Gamma_{rs;\nu}$ |
| ${}^{tt}\Gamma_{pq;\mu\nu}$ | $\frac{\partial^2\rho^{(2)}_{\text{CC}}}{\partial t^{(1)}_\mu(\omega)\partial t^{(1)}_\nu(-\omega)}$ | | $\langle\Lambda\vert[[E_{pq},\tau_\mu],\tau_\nu]\vert\text{CC}\rangle$ |
| ${}^{t\bar{t}}\Gamma_{pq;\mu\nu}$ | $\frac{\partial^2\rho^{(2)}_{\text{CC}}}{\partial t^{(1)}_\mu(\omega)\partial\bar{t}^{(1)}_\nu(-\omega)}$ | | $\langle\nu\vert e^{-\hat{T}^{(0)}}[E_{pq},\tau_\mu]\vert\text{CC}\rangle$ |
| ${}^{t}\Gamma_{pq;\mu}$ | $\frac{\partial\rho^{(1)}_{\text{CC}}}{\partial t^{(1)}_\mu(\omega)}$ | | $\langle\Lambda\vert[E_{pq},\tau_\mu]\vert\text{CC}\rangle$ |
| ${}^{\bar{t}}\Gamma_{pq;\mu}$ | $\frac{\partial\rho^{(1)}_{\text{CC}}}{\partial\bar{t}^{(1)}_\mu(\omega)}$ | | $\langle\mu\vert e^{-\hat{T}^{(0)}}E_{pq}\vert\text{CC}\rangle$ |

non-zero due to contributions arising from $\{Q^{(2)}_{\text{int}}\}_T$. This makes the expressions for the poles of the response function somewhat more complicated than the simple relation between excitation energies and the eigenvalues of **A** found for the vacuum case. However, one can recover such a relation by following the computational strategy employed in the CC/MM schemes[42] and neglect the contributions from **F** and **J** when calculating excitations.

Thus, from the definition of the CC-in-DFT Jacobian,

$$ {}^{\bar{t}t}A_{\mu_i,\nu_j}(\omega_Y) = \frac{\partial^2\{^{2n+1}L^{(2)}\}_T}{\partial\bar{t}^{(1)}_{\mu_i}(\omega)\partial t^{(1)}_{\nu_j}(-\omega)}, \tag{25}$$

one identifies the FDE contributions as

$$\begin{aligned}
{}^{\bar{t}t}A_{\mu_i,\nu_j} \leftarrow & \sum_{pq}\langle p\vert v_{\text{int}}\vert q\rangle^{t\bar{t}}\Gamma_{pq;\mu_i\nu_j} \\
& +\frac{1}{2}\sum_{pqrs}(pq\vert w_{\text{int}}\vert rs)^{t}\Gamma_{pq;\mu_i}{}^{\bar{t}}\Gamma_{rs;\nu_j} \\
= & \sum_{pq}\langle p\vert v_{\text{int}}\vert q\rangle\langle\mu_i\vert\exp(-T)[E_{pq},\tau_{\nu_j}]\vert\text{CC}\rangle \\
& +\frac{1}{2}\sum_{pq,rs}(pq\vert w_{\text{int}}\vert rs)\quad\langle\mu_i\vert\exp(-T)E_{pq}\vert\text{CC}\rangle \\
& \langle\Lambda\vert[E_{rs},\tau_{\nu_j}]\vert\text{CC}\rangle,
\end{aligned} \tag{26}$$

where $T$ here denotes the ground-state CC amplitudes and the superscripts on the density matrices defined in Table II are omitted for brevity. In Eq. (26) apart from a term including the potential seen also in Eq. (20), a new contribution arises which is denoted kernel contribution, because it is the analogue of the exchange-correlation (XC) kernel in DFT-based response theory:

$$\begin{aligned}
& (pq\vert w_{\text{int}}\vert rs) \\
& = \iint\phi^\dagger_p(\mathbf{r}')\phi_q(\mathbf{r}')\left[\left.\frac{\delta^2E_{\text{xck}}}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})}\right|_{\rho_{\text{tot}}} - \left.\frac{\delta^2E_{\text{xck}}}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})}\right|_{\rho_{\text{t}}}\right] \\
& \quad\times\phi^\dagger_r(\mathbf{r})\phi_s(\mathbf{r})d\mathbf{r}d\mathbf{r}'.
\end{aligned} \tag{27}$$

Furthermore, given that, in general, the coupled-cluster eigenvalue problem is not solved exactly but with approximate schemes for the lowest-lying solutions, in which a matrix **A** is contracted with a reduced-space trial vector $\vec{R}$:[43]

$$\mathbf{A}\vec{R} = \vec{\varrho}, \tag{28}$$

this leads to the following contributions to the singles part:

$$\begin{aligned}
{}^1\varrho^{\text{emb}}_{ia} \leftarrow & \sum_b v_{ab}R^b_i - \sum_j v_{ji}R^a_i \\
& +\frac{1}{2}\sum_{jb}\sum_{pq,rs}(pq\vert w_{\text{int}}\vert rs)^{\bar{t}}\Gamma_{pq;ia}{}^{t}\Gamma_{rs;jb}R^b_j,
\end{aligned} \tag{29}$$

$$\begin{aligned}
{}^2\varrho^{\text{emb}}_{ia} \leftarrow & \sum_{ck}\left(2R^{ac}_{ik} - R^{ac}_{ki}\right)v_{kc} \\
& +\frac{1}{2}\sum_{jkbc}\sum_{pq,rs}(pq\vert w_{\text{int}}\vert rs)^{\bar{t}}\Gamma_{pq;ia}{}^{t}\Gamma_{rs;jbkc}R^{bc}_{jk},
\end{aligned} \tag{30}$$

while doubles part is augmented with the following contributions:

$$ {}^1\varrho^{\text{emb}}_{ia,jb} \leftarrow \frac{1}{2}\sum_{kc}\sum_{pq,rs}(pq\vert w_{\text{int}}\vert rs)^{\bar{t}}\Gamma_{pq;iajb}{}^{t}\Gamma_{rs;kc}R^c_k, \tag{31}$$

$$
{}^{2}\varrho^{\mathrm{emb}}_{ia,jb} \leftarrow \sum_{c} \left( R^{ac}_{ij} \hat{h}_{bc} + R^{bc}_{ij} \hat{h}_{ac} \right) - \sum_{k} \left( R^{ab}_{ik} \hat{h}_{kj} + R^{ab}_{jk} \hat{h}_{ki} \right) + \frac{1}{2} \sum_{jkbc} \sum_{pq,rs} (pq|w_{\mathrm{int}}|rs)^{\bar{t}} \Gamma_{pq;iajb}{}^{t} \Gamma_{rs;jbkc} R^{bc}_{jk}. \tag{32}
$$

Apart from the choice of density functional (local density approximation (LDA), GGA) for the calculation of the kernel contributions, different additional approximations are possible. The first and most severe is to drop the contributions completely,[23] whereas a second is the use of approximate densities for evaluating the functional derivatives. Coupled cluster singles (CCS) densities are then the simplest choice, since the ${}^{2}\varrho^{\mathrm{emb}}_{ia}$, ${}^{1}\varrho^{\mathrm{emb}}_{ia,jb}$, and ${}^{2}\varrho^{\mathrm{emb}}_{ia,jb}$ contributions are zero and only one term remains:

$$
{}^{1}\varrho^{\mathrm{emb}}_{ia} \leftarrow \sum_{jb} (ia|w_{\mathrm{int}}|jb) R^{b}_{j}. \tag{33}
$$

One could also consider the use of other approximations, such as using CCS-like densities as recently proposed by Schwabe *et al.*, where the T1 amplitudes and multipliers from CC2 are used.[44]

## III. COMPUTATIONAL DETAILS

### A. FDE calculations

This work is concerned with the treatment of an active system whose density is allowed to be polarized in the presence of an environment whose density is kept frozen. Thus, all preparatory FDE DFT or DFT-in-DFT calculations, necessary to obtain the environment density and "fixed" embedding potentials, were performed using the ADF program package[45,46] via the PyADF scripting framework.[47]

All coupled-cluster calculations were performed with a locally modified version of the Dalton 2011 code.[48] Furthermore, all time-dependent Hartree-Fock calculations have been performed with the Dirac program,[49] using the Levy-Leblond (non-relativistic) Hamiltonian.[50]

In discussing the FDE calculations below, the abbreviation "LDA" refers to an embedding potential created with Slater exchange and Thomas-Fermi kinetic energy, whereas "GGA" refers to an embedding potential created with the PBE[51] exchange-correlation functional in combination with the PW91[52] kinetic energy functional (PW91k).

Given the different possibilities to determine the ground-state parameters as well as other computational approximations, one may employ combinations of the following:

- Perturbed/Unperturbed HF orbitals:
  Orbitals obtained via pathways 2 and 1, respectively, see Sec. II D 3.
- Recalculated/Fixed embedding potential:
  Embedding potential is/is not updated using the coupled-cluster density for subsystem I.
- No/ALDA/GGA kernel contributions in the CC Jacobian:
  If a kernel is used, the approximate CCS densities are employed in the present work.

For instance, the abbreviation CCSD-`emb{p,r,GGA}` denotes CCSD level of theory with perturbed HF orbitals and a recalculated (using the corresponding CC density) embedding potential, where the excitation energies were calculated with kernel contributions with a GGA kernel (evaluated with a CCS density). If no kernel contributions are present, a dash is used, e.g., `emb{p,r,-}`.

### B. Water in water

The geometry for the solvated water system was taken from Ref. 24 which corresponds to one snapshot from a MD simulation, where one water is surrounded by an environment consisting of 127 other water molecules. For this test case, the aug-cc-pCVTZ basis set[53] was employed, and no orbitals were frozen in the coupled-cluster calculations. The density for the environment is constructed from the superposition of densities for the individual molecules, employing the AUG/ATZ2P basis.

### C. Uracil in water

For the uracil-water system, two different structural models were explored: a "static" one, along the lines of the microsolvated uracil described in Ref. 54, where the geometries used for uracil and six surrounding water molecules correspond to those optimized for the cluster; and a "dynamic" one, where 120 snapshots from a MD simulation[55,56] are used. In this simulation, the uracil geometry was kept fixed and the number of water molecules taken into account for each snapshot (about 240) corresponds to retaining all solvent molecules within a 12 Å radius from the uracil, and has been verified to yield excitation energies converged with respect to the number of waters.[56] The environment densities were obtained by ADF FDE calculations employing a TZP basis.

The basis sets used in the coupled-cluster calculations for the "static" and "dynamic" models are, respectively, cc-pVDZ[57] and aug-cc-pVDZ basis.[58] The smaller basis used for the former is due to the fact that it turned out that calculations on the uracil-water cluster were not feasible in a large basis.

## IV. RESULTS

### A. Water in water

The results for the lowest singlet excited state for pure water case can be found in Table III, along with the results of Ref. 24 for the model potential for the statistical average of orbital potentials (SAOP),[59–61] specifically designed to get the correct asymptotic behavior of the exchange-correlation potential, used in combination with FDE and the QM/MM discrete reaction field (DRF)[62] method, where the classical solvent molecules are represented using distributed atomic charges and polarizabilities, and the experimental values.[63] Before analyzing the effects of the parameters in the FDE calculations (such as orbitals, embedding potential, and kernel contributions) in more detail, first some general trends are discussed.

The different electronic structure methods show a typical behavior, with the uncorrelated methods (CCS,

TABLE III. Water in water: Influence of different FDE parameters on the calculated solvated excitation energies (Solv.) and shifts (in eV). Here, `emb{a,b,c}` denotes embedding contributions using `a`: orbitals determined (`p`) in the presence of $v_{\text{int}}(\mathbf{r})$ or (`u`) in vacuum; `b`: embedding potential (`s`) determined in DFT FDE calculations or (`r`) recalculated from $\rho_{\text{CC}}$ and $\rho_{\text{II}}$; and `c`: with or without (`-`) kernel contributions to the excitations.

| Method | Potential | | Vacuum | Solv. | Shift | Δ |
|---|---|---|---|---|---|---|
| HF | GGA | emb{p,f,-} | 8.66 | 9.31 | +0.65 | … |
| | LDA | emb{p,r,-} | | 9.21 | +0.55 | … |
| | GGA | emb{p,r,-} | | 9.29 | +0.63 | … |
| | GGA | emb{p,r,ALDA} | | 9.34 | +0.68 | 0.04[a] |
| | GGA | emb{p,r,GGA} | | 9.35 | +0.69 | 0.02[b] |
| CCS | GGA | emb{p,f,-} | 8.70 | 9.34 | +0.64 | … |
| | LDA | emb{p,r,-} | | 9.58 | +0.88 | … |
| | GGA | emb{p,r,-} | | 9.27 | +0.57 | … |
| | GGA | emb{p,r,ALDA} | | 9.32 | +0.62 | 0.05[a] |
| CC2 | GGA | emb{p,f,-} | 7.28 | 7.79 | +0.51 | … |
| | LDA | emb{u,r,-} | | 8.17 | +0.89 | … |
| | LDA | emb{p,r,-} | | 8.14 | +0.86 | … |
| | GGA | emb{u,r,-} | | 7.80 | +0.52 | … |
| | GGA | emb{p,r,-} | | 7.78 | +0.50 | … |
| | GGA | emb{p,r,ALDA} | | 7.81 | +0.53 | 0.03[a] |
| CCSD | GGA | emb{p,f,-} | 7.64 | 8.23 | +0.59 | … |
| | LDA | emb{u,r,-} | | 8.55 | +0.91 | … |
| | LDA | emb{p,r,-} | | 8.52 | +0.88 | … |
| | GGA | emb{u,r,-} | | 8.23 | +0.59 | … |
| | GGA | emb{p,r,-} | | 8.20 | +0.56 | … |
| | GGA | emb{p,r,ALDA} | | 8.24 | +0.60 | 0.04[a] |
| SAOP | GGA(PBE) | emb{p,r,ALDA} | 7.75 | 8.48 | +0.73 | 0.06[a] |
| SAOP[24] | GGA(PW91) | emb{p,r,ALDA} | 7.76 | 8.71 | +0.95 | ×[c] |
| Expt.[63] | | | 7.4 | 8.2 | +0.8 | |

[a]Difference to no kernel (-).
[b]Difference to ALDA.
[c]Not given.

time-dependent Hartree-Fock (TDHF)) showing excitation energies that are significantly larger than for correlated methods (CC2 and CCSD) in vacuum as well as in solution. Between CC2 and CCSD, it can be seen that the former tends to be about 0.5 eV lower than the latter, in line with trends from other studies (see for instance Ref. 12).

Even though only one structure for the solvated water is considered and therefore one cannot draw definitive conclusions on the method's ability to reproduce experimental results, it can in any case be noted that the SAOP, CC2, and CCSD excitation energies and solvatochromic shifts are quite compatible to the experimental one, with the exception of the SAOP/FDE values of Ref. 24, for which the solvation effect is much larger than for the others. We believe that this is largely due to the description of the frozen waters, for which the density was obtained with the LDA functional and the DZP basis, as the SAOP calculation, employing the PBE functional to calculate the non-additive exchange-correlation contributions, and the larger basis set to represent the frozen waters leads to a shift of 0.73 eV, roughly 0.2 eV lower than in Ref. 24.

### 1. *Effect of the embedding potential and kernel*

Like in DFT, where the kernel contributes 0.06 eV to the shift, kernel contributions are generally small for this excitation irrespective of the electronic structure method used. A comparison between the GGA and adiabatic LDA (ALDA) kernels shows the latter already describes the bulk of the effect (in this case, about 75% of the final value) and should be sufficient for most purposes. The influence of the (perturbed) density used to evaluate the kernel contributions is also rather small, for the wave function-based methods a variation between 0.03 eV and 0.05 eV is observed. The SAOP value falls outside this range, but note that this value was calculated with ADF and a Slater type basis set which will also have an effect.

Relatively small values for the kernel contributions in cases such as this are expected from a physical standpoint, since the interactions are rather weak and the kernel contributions represent a second-order effect. The embedding potential appears in first order, and more important differences are observed when comparing potentials obtained with LDA and GGA functionals. For the coupled cluster methods GGA potentials give excitation energy shifts that are about 0.3 eV smaller than the corresponding LDA shifts. In the Hartree-Fock calculation the GGA embedding potential does, on the other hand, give a slightly larger shift than the LDA one. This demonstrates that solvation and correlation shifts are in general non-additive because the two types of wave functions probe the embedding potential in a different way.

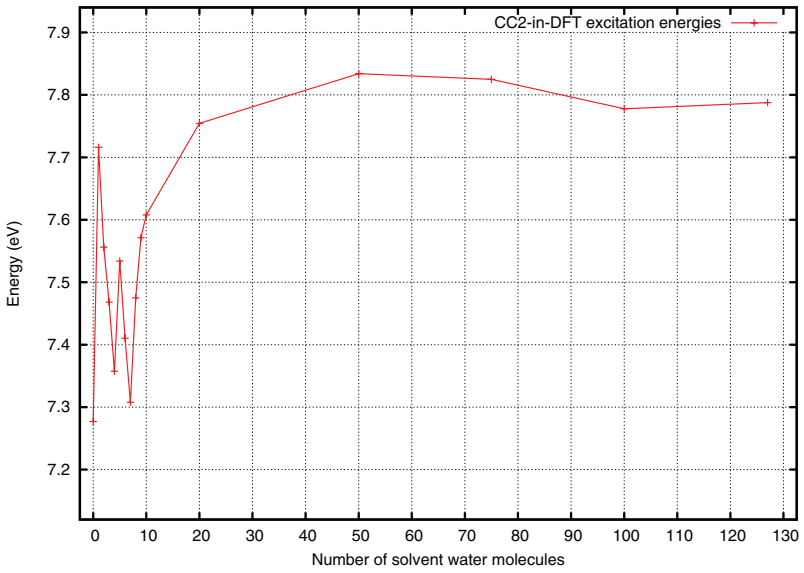



FIG. 2. Convergence of aug-cc-pCVTZ/CC2-in-DFT emb{p,f,-} excitation energies with respect to the number of solvent molecules. For up to about 20 water molecules, strong charge-transfer character is observed and the FDE approximation thus exhibits larger error bars with respect to supermolecular results.

For the embedded excitation energies, the results exhibit fortunate error compensation for the CC2 with LDA excitation energy. This value is close to those of the most sophisticated (CCSD with GGA) method as well as experimental results. By comparing the contributions one may easily see, however, that this is due to an overestimation of the shift by the LDA potential which is compensated by an underestimation of the excitation energy by the CC2 method.

### 2. Choice of orbitals

Considering the choice of orbitals in the computational scheme 1, one can observe that in this case the differences between emb{u,r,-} and emb{p,r,-} excitation energies are rather small and usually of the order of 0.03 eV, irrespective of whether a LDA or GGA embedding potential is used. This is in line with findings in the literature[56,65] for this system, and is an indication that the approximate orbital relaxation brought about by the T1 amplitudes in CC-in-DFT calculations was nearly sufficient, in the emb{u,r,-} case, to compensate for the use of its vacuum starting orbitals. As such small differences can hardly be used to favor one approach over the other, the discussion returns to this point in Sec. IV B, where these are more pronounced.

Apart from the two approaches discussed above, it remains to compare them to the "fixed potential" approach,[23] represented here as emb{p,f,-}. As expected, since the DFT-in-DFT potential is a GGA one, one sees that, the emb{p,f,-} results are close to the GGA emb{p,r,-} or emb{u,r,-}. Surprising is, however, that adding the kernel contributions in the emb{p,r,GGA} or emb{p,r,ALDA} approach gives an even closer agreement with roughly 0.02 eV discrepancies to emb{p,f,-} for the different electronic structure methods. While this must be due to error cancellation, the simple "fixed potential" approach should be a good choice for exploratory work, or for cases in which the iterated calculation of (orbitals and) coupled-cluster densities is not feasible.

### 3. Convergence of the excitation energies with number of solvent molecules

Apart from the effect of the different parameters discussed above, it is also interesting to see what are the effects on the excitation energies of the size the environment in the structural models. In order to illustrate this in the water case, in Fig. 2, the evolution of the CC2-in-DFT excitation energy is collected, with embedding contributions calculated at the GGA emb{p,f,-} level, as the number of frozen water molecules is increased from one (the nearest neighbor to the active water) to the full microsolvation model for the selected MD snapshot. From the figure it can be observed that the FDE excitation energies tend to increase as the size of the environment is increased, with relatively large variations for small models with up to eight solvent molecules. For larger microsolvation models the converge with system size is smoother and for at about twenty solvent molecules there are no significant variations in comparison to the result obtained for the complete set.

While at this point one could also think of comparing FDE and supermolecular results for smaller clusters in order to see how the differences between the two varies with environment size, we believe that this is difficult for this system. This is due to the delocalized nature of the lowest-lying excited states for this system found in supermolecular CCSD calculations.[64] Such a delocalisation can not be described with the uncoupled FDE approach. It is interesting to note, however, that these CCSD results suggest a lowering

TABLE IV. Microsolvated model for uracil in water, CC2 excitation energies ($E_{solv}$) and shifts (calculated using as vacuum structure for uracil: the supermolecular one, $\Delta E_S$; or optimized in vacuum, $\Delta E_O$) in eV for different methods and basis sets. Geometries are taken from Ref. 54. Here, `emb{a,b,c}` denotes embedding contributions using `a`: orbitals determined (`p`) in the presence of $v_{int}(\mathbf{r})$ or (`u`) in vacuum; `b`: embedding potential (`s`) determined in DFT FDE calculations or (`r`) recalculated from $\rho_{CC}$ and $\rho_{II}$; `c`: with or without (`-`) kernel contributions to the excitations.

| | | $n \rightarrow \pi^*$ | | | $\pi \rightarrow \pi^*$ | | |
|---|---|---|---|---|---|---|---|
| | | $E_{solv}$ | $\Delta E_S$ | $\Delta E_O$ | $E_{solv}$ | $\Delta E_S$ | $\Delta E_O$ |
| cc-pVDZ | emb{u,f,-} | 5.57 | +0.92 | +0.52 | 5.47 | −0.01 | −0.22 |
| aug-cc-pVDZ | emb{u,f,-} | 5.50 | +0.95 | +0.57 | 5.22 | +0.02 | −0.18 |
| cc-pVDZ | emb{p,f,-} | 5.42 | +0.77 | +0.37 | 5.46 | −0.02 | −0.23 |
| aug-cc-pVDZ | emb{p,f,-} | 5.37 | +0.82 | +0.44 | 5.19 | −0.01 | −0.21 |
| cc-pVDZ | emb{u,r,-} | 5.57 | +0.92 | +0.52 | 5.47 | −0.02 | −0.22 |
| aug-cc-pVDZ | emb{u,r,-} | 5.50 | +0.95 | +0.57 | 5.22 | +0.02 | −0.18 |
| cc-pVDZ | emb{p,r,-} | 5.42 | +0.77 | +0.37 | 5.46 | −0.02 | −0.23 |
| aug-cc-pVDZ | emb{p,r,-} | 5.37 | +0.82 | +0.44 | 5.19 | −0.01 | −0.21 |
| cc-pVDZ | emb{p,r,ALDA} | 5.42 | +0.77 | +0.37 | 5.46 | −0.02 | −0.23 |
| aug-cc-pVDZ | emb{p,r,ALDA} | 5.37 | +0.82 | +0.44 | 5.20 | −0.00 | −0.20 |
| cc-pVDZ | supermolecule | 5.31 | +0.66 | +0.26 | 5.36 | −0.12 | −0.33 |
| aug-cc-pVDZ | supermolecule | 5.19 | +0.64 | +0.26 | 5.12 | −0.08 | −0.28 |

$E_S^{n\rightarrow\pi^*}$: 4.65 eV (VDZ), 4.55 eV (aVDZ).
$E_O^{n\rightarrow\pi^*}$: 5.05 eV (VDZ), 4.93 eV (aVDZ).
$E_S^{\pi\rightarrow\pi^*}$: 5.48 eV (VDZ), 5.20 eV (aVDZ).
$E_O^{\pi\rightarrow\pi^*}$: 5.69 eV (VDZ), 5.40 eV (aVDZ).

of the excitation energies as the number of waters increases, which is at odds with what is measured experimentally and calculated with CC2-in-DFT. A more thorough investigation of these trends using ensemble-averaged coupled cluster calculations would be desirable but is beyond the scope of the present work.

## B. Uracil in water

The calculation of solvatochromic shifts for the lowest two excitations of uracil is a well-studied benchmark test case, to which various theoretical approaches have been applied, ranging from continuum to explicit solvation methods.[54,56,65,66] One important aspect is that for these two excitations two quite different shifts are obtained upon solvation: The $n \rightarrow \pi^*$ transition, the lowest excitation in vacuum, experiences a strong blue shift of about +0.5 eV while a clear red shift of about −0.2 eV is found for the $\pi \rightarrow \pi^*$ transition, the second-lowest excitation, eventually leading to an inverted order for the two connected excited states due to the fact that they are only separated by about 0.4 eV in vacuum. For uracil it is thus in particular necessary to obtain also the correct order of these two states, which can be used to assess the accuracy of a method.

### 1. Microsolvation model

In Table IV vertical CC2 excitation energies are shown for microsolvated uracil at its optimized geometry,[54] along with two solvatochromic shifts: the first, denoted by $\Delta E_S$, is the difference between excitations calculated with solvent ($E_{solve}$) and in vacuum (but employing the uracil structure in the supermolecular cluster, $E_S$), whereas in the second, denoted by $\Delta E_O$, the excitations of uracil in vacuum are obtained for a structure optimized without the solvent ($E_O$). While only $\Delta E_O$ can be compared to experimentally determined solvatochromic shifts, a separated analysis of the two serves to better understand the effect of structural relaxation of uracil upon solvation.

For this section, only the CC2 method was applied, as this level of theory is estimated to be sufficient to reveal and investigate the two main differences between FDE and supermolecule calculations. First, an intrinsic limitation originating from the fact that in FDE calculations excitations are strongly localized on a fragment, whereas in the supermolecule calculations the delocalization over the nearby solvent molecules includes more of environment response. Second, the accuracy is limited by the available kinetic energy functionals which can make the supermolecular and embedded ground state densities differ significantly.

The calculations reveal a slightly increased difference between unperturbed (`emb{u,r,-}`) and perturbed (`emb{p,r,-}`) orbitals compared to the water-in-water case. For the $n \rightarrow \pi^*$ this amounts to 0.06 eV, whereas the difference is 0.1 eV in the $\pi \rightarrow \pi^*$ excitation, which indicates that the approach including the embedding potential at the Hartree-Fock step is indeed advantageous. The table also shows that the difference of a fixed (`emb{p,f,-}`) and a recalculated (`emb{p,r,-}`) embedding potential does not lead to large differences, which confirms and strengthens the use of "fixed" embedding potentials in large-scale applications.[23]

From Table IV we can see that $\Delta E_S$ and $\Delta E_O$ differ by a factor of 2 for the two electronic states, with the latter being lower than the former by about 0.40 eV for $n \rightarrow \pi^*$ and 0.20 eV for $\pi \rightarrow \pi^*$ excitations, respectively. This structural effect is essentially the same for the different embedding approaches and the supermolecular case, for which we have values of $\Delta E_S$ of +0.66 eV and −0.12 eV for the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ excitations, respectively. Furthermore, for both $\Delta E_S$

and $\Delta E_O$ one has that the methodologically most sophisticated FDE calculation, `emb{p,r,ALDA}`, is blue-shifted by about 0.10 eV for both excitations relative to the supermolecular calculations. This remaining difference is mainly due to the missing delocalization of the excitation[67] in the FDE case.

One can include some of this delocalization effect by extending the basis set with functions on the frozen systems, but this easily leads to problems with the approximate kinetic energy function which tends to give too attractive embedding potentials.[68] A more robust, but also more expensive, approach is to calculate the lowest excited states of the surrounding waters explicitly and then couple the most important excitations explicitly.[69]

Recently, DeFusco *et al.*, employing the effective fragment potential (EFP) method,[66] obtained solvatochromic shifts relative to the relaxed monomer of about −0.26 eV for $\pi \rightarrow \pi^*$ and about +0.43 eV for $n \rightarrow \pi^*$, which are well in line with CCSD FDE results, but slightly higher than CC2 FDE values, and somewhat higher than the COSMO CC2 values of Ref. 54. However, in agreement with Ref. 54, in the present work using FDE, the main effects can be captured with a small number of explicit solvation molecules, independent of the treatment in a supermolecular calculation or represented on a grid in FDE. Putting errors of the basis set, approximate functionals and truncation of the wave function aside, the missing effects compared to experimental results should arise from outer solvation shells and statistical averaging, which will be discussed in Sec. IV B 2.

### 2. Uracil in solution

In Sec. IV B 1, it was outlined that the microsolvation model is able to obtain excitation energies in reasonably good agreement with other theoretical results[54,56,65,66] and the experimental results for the $\pi \rightarrow \pi^*$ excitation.[70,71] The $n \rightarrow \pi^*$ excitation is too weakly absorbing to be readily detected, so that the discussion of this excitation needs to be done by comparing the theoretical results.

It remains to be done to investigate the importance of the outer solvation shells as well as temperature effects for both excitations via the statistical averaging over the configuration space spanned by the MD snapshots. In order to investigate averaging and outer solvation shells separately, the 120 snapshots have been used to construct two models: (a) where only the 6 closest water molecules were included in the FDE calculations, and (b) where all 240 water molecules were included in the FDE calculation. Supermolecule calculations which explicitly include such a larger explicit solvation model were not feasible and no comparison for case (b) could be made.

The averaged results of both are collected in Table V, while a detailed plot for model (b) of both excitations for each snapshot can be found in Fig. 3. Because the solute-solvent conformations were obtained with force field methods, shifts are only compared to the $\Delta E_O$ values from Table IV. It can be seen that model (a) performs rather poorly, as the obtained shifts are much smaller than experimental or previously calculated ones. That appears to be linked to the fact that the uracil geometry used in the MD simulation is close to the vacuum and the waters are in general further away from the uracil compared to the optimized minimum geometry in Sec. IV B 1—a combination which obviously leads to a reduced embedding contribution.

TABLE V. CC2/aug-cc-pVDZ excitation energies ($E_{solv}$) and shifts (relative to the geometry optimized in vacuum, $\Delta E_O$) in eV, for two models based on the structures from the 120 MD snapshots by Kongsted and co-workers:[55] (a) only the six nearest waters to uracil are included in the FDE (GGA potential, using orbitals determined (`p`) in the presence of $v_{int}(\mathbf{r})$, embedding potential (`f`) determined in DFT FDE calculations and without kernel contributions to the excitations, `emb{p,f,-}`) treatment; (b) all water molecules are included in the FDE treatment. $E_{av}$ denotes the averaged values, while $E_{min}$ and $E_{max}$ denote the lowest and highest excitations from the set of snapshots, respectively. The experimental excitation energy (shift) for $\pi \rightarrow \pi^*$ is measured to be 4.78 ± 0.01 eV (−0.29 ± 0.01 eV).[70,71]

| | | Model (a) | | Model (b) | |
|---|---|---|---|---|---|
| Excitation | | $E_{solv}$ | $\Delta E_O$ | $E_{solv}$ | $\Delta E_O$ |
| $n \rightarrow \pi^*$ | $E_{av}$ | 4.93 | 0.00 | 5.25 | +0.33 |
| | $E_{max}$ | 5.40 | +0.47 | 5.54 | +0.61 |
| | $E_{min}$ | 4.60 | −0.33 | 4.87 | −0.06 |
| $\pi \rightarrow \pi^*$ | $E_{av}$ | 5.35 | −0.05 | 5.30 | −0.10 |
| | $E_{max}$ | 5.50 | +0.10 | 5.44 | +0.03 |
| | $E_{min}$ | 5.12 | −0.28 | 5.14 | −0.26 |

Comparing models (a) and (b), it can be seen that the inclusion of outer solvation shells is very important, in particular for the $n \rightarrow \pi^*$ state, as it induces a systematic increase on the shifts of about +0.3 eV for model (b). The $\pi \rightarrow \pi^*$ is also affected, with a shift of opposite sign and less pronounced of about −0.05 eV. These differences in the states' sensitivity to conformational changes upon solvation are in line with what is observed in other calculations.[54,56] The present results nevertheless show absolute deviations of about 0.1 eV with respect to the CC2/MM results of Olsen *et al.*[56] While these are in line with the current findings concerning the difference between FDE and supermolecular calculations, in Fig. 3 it becomes clear that even for model (b) it remains that most configurations still exhibit too small solvatochromic shifts, while

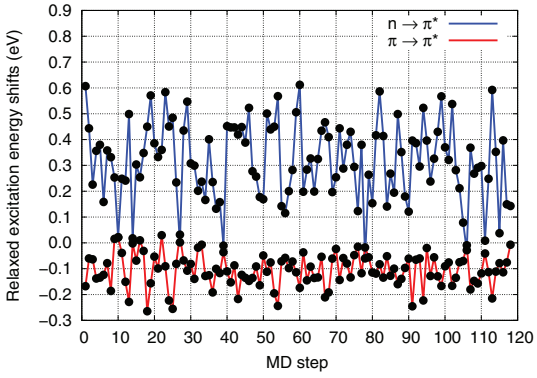



FIG. 3. Shifts of the lowest two excitations for 120 snapshots using aug-cc-pVDZ/CC2-in-DFT `emb{p,f,-}`. Taking into account the geometry relaxation of the monomer, the averaging yields shifts of +0.33 eV for $n \rightarrow \pi^*$, and −0.10 eV for $\pi \rightarrow \pi^*$.

TABLE VI. CC2/cc-pVDZ excitation energies in eV for the microsolvated uracil model with a varying number of molecules represented by FDE (GGA potential, using orbitals determined (p) in the presence of $v_{\text{int}}(\mathbf{r})$, embedding potential (f) determined in DFT FDE calculations and without kernel contributions to the excitations, emb{p,f,-}). The first and last lines correspond to the supermolecular and FDE results in Table IV, respectively.

| # $H_2O$ | | Excitation energies | |
|---|---|---|---|
| Explicit | FDE | $n \rightarrow \pi^*$ | $\pi \rightarrow \pi^*$ |
| 6 | 0 | 5.31 | 5.36 |
| 5 | 1 | 5.37 | 5.39 |
| 4 | 2 | 5.36 | 5.42 |
| 3 | 3 | 5.37 | 5.42 |
| 2 | 4 | 5.41 | 5.46 |
| 1 | 5 | 5.42 | 5.47 |
| 0 | 6 | 5.42 | 5.46 |

only some show results consistent with the results of the microsolvated cluster. It could be therefore interesting to investigate whether similar observations are made with structures originating from different MD approaches, such as *ab initio* MD (AIMD). With this, it would also be more straightforward to take into account the relaxation of the uracil geometry with respect to the surrounding water molecules, which has been kept frozen in the MD simulation by Olsen *et al.*[56]

### 3. Convergence of the excitation energies with number of solvent molecules

Unlike the solvated water case, for uracil the low-lying electronic excitations remain essentially localized on the uracil, and a comparison between supermolecule CC2 and CC2-in-DFT calculations is more instructive. Here we have opted to investigate, for the microsolvated model, the effect on the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ states of progressively shifting an outermost solvent molecule from the supermolecular description to the embedded one.

The results, found in Table VI, show first that supermolecule-FDE discrepancies are roughly equivalent for both excitations within each of the different models. Furthermore, one can observe three distinct groups, with respect to the magnitude of the supermolecule-FDE discrepancies: a first one for one FDE water, a second for two or three FDE waters for which FDE overestimates the supermolecule excitations by about 0.05 eV, and a third with four to six waters, where such overestimation reaches roughly 0.09 eV. These results are in line with a prior discussion of the relatively important contributions from the nearest solvent molecules to the excitations, absent by construction in the CC-in-DFT approach. They also indicate that one can obtain a non-negligible decrease (50% in this case) in the deviation from the reference by a relatively modest enlargement of the active subsystem.

## V. CONCLUSION

In this work, we have presented the first implementation of WFT-in-DFT embedding within the response theory framework recently introduced,[26] which allows for the calculation of general ground- and excited-state properties employing CC wave functions to describe the "wave-function" subsystem.

In contrast to Ref. 23, with the current implementation it is possible to calculate the embedding potential from coupled-cluster densities for the ground state, as well as to take into account the time-dependence of the density in the excited states via the calculation of the kernel contributions to the CC Jacobian matrix. This can be done either rigorously by calculating the electron density with the same level of theory as other parts of the calculation, or in a more approximate fashion using a computationally simple model such as CCS to obtain the densities.

The presented results for the solvatochromic shifts of two rather different model systems in water indicate, first, that WFT-in-DFT can capture quantitatively the changes in electronic spectra induced by the solvent molecules. The discrepancies with respect to supermolecular calculations of about 0.1 eV are smaller than errors due to basis set incompleteness or to the truncation of the excitation operator in the coupled-cluster treatment. Furthermore, one observes that the importance of embedding kernel contributions is for the systems considered relatively small (about 10% of the total shift value) compared to that of the embedding potential, and a major contribution of the effect can be captured via the adiabatic LDA approximation, while the embedding potential is preferably calculated with GGA functionals.

These findings not only validate the approximate scheme used in Ref. 23 but show that in cases where certain approximations are not valid, the methodological hierarchy allows for systematic steps towards the full model. An exception remains the limited accuracy of currently available orbital-free kinetic energy functionals common to all embedding approaches relying upon them.

In the present work, the polarization of the (active) coupled-cluster density with respect to a frozen environment is explicitly taken into account. The next extension consists of allowing the mutual polarization of the subsystems in a CC-in-DFT freeze-thaw procedure.

## ACKNOWLEDGMENTS

The authors are grateful to Erik D. Hedegård and Jacob Kongsted for kindly providing the snapshots of the uracil molecular dynamics simulation.

This study was supported by the EC-supported **ACTINET-i3** Integrated Infrastructure Initiative (JRP-C3-10, JRP-C5-08). A.S.P.G. acknowledges support from PhLAM (Laboratoire de Physique des Lasers, Atomes et Molécules, Unité Mixte de Recherche de l'Université de Lille 1 et du CNRS). S.H. has been supported by the European Commission under a Marie-Curie Intra-European Fellowship (Contract No. PIEF-GA-2010-274224). L.V. has been supported by a VICI grant by the Netherlands Organisation for Scientific Research (NWO).

[1]C. A. Ullrich, *Time-Dependent Density-Functional Theory: Concepts and Applications* (Oxford University Press, 2012).
[2]R. J. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).

[3]D. Foester, Phys. Rev. B **72**, 073106 (2005).
[4]C. Adamo and D. Jacquemin, Chem. Soc. Rev. **42**, 845 (2013).
[5]P. Tecmer, K. Kowalski, W. A. de Jong, and L. Visscher, J. Chem. Phys. **139**, 034301 (2013).
[6]N. O. Winter and C. Hättig, J. Chem. Phys. **134**, 184101 (2011).
[7]A. S. P. Gomes and C. R. Jacob, Annu. Rep. Prog. Chem., Sect. C: Phys. Chem. **108**, 222 (2012).
[8]J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, Mol. Phys. **100**, 1813 (2002).
[9]J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, J. Phys. Chem. A **107**, 2578 (2003).
[10]J. Kongsted, A. Osted, K. V. Mikkelsen, and O. Christiansen, J. Chem. Phys. **118**, 1620 (2003).
[11]C. B. Nielsen, O. Christiansen, K. V. Mikkelsen, and J. Kongsted, J. Chem. Phys. **126**, 154112 (2007).
[12]K. Sneskov, T. Schwabe, J. Kongsted, and O. Christiansen, J. Chem. Phys. **134**, 104108 (2011).
[13]K. Sneskov, T. Schwabe, O. Christiansen, and J. Kongsted, Phys. Chem. Chem. Phys. **13**, 18551 (2011).
[14]P. Cortona, Phys. Rev. B **44**, 8454 (1991).
[15]P. Cortona, Phys. Rev. B **46**, 2008 (1992).
[16]T. A. Wesolowski and A. Warshel, J. Phys. Chem. **97**, 8050 (1993).
[17]M. E. Casida and T. A. Wesolowski, Int. J. Quantum Chem. **96**, 577 (2004).
[18]J. Neugebauer, J. Chem. Phys. **126**, 134116 (2007).
[19]J. Neugebauer, J. Phys. Chem. B **112**, 2207 (2008).
[20]J. Neugebauer, J. Chem. Phys. **131**, 084104 (2009).
[21]C. König, N. Schlüter, and J. Neugebauer, J. Chem. Phys. **138**, 034104 (2013).
[22]P. Huang and E. A. Carter, J. Chem. Phys. **125**, 084102 (2006).
[23]A. S. P. Gomes, C. R. Jacob, and L. Visscher, Phys. Chem. Chem. Phys. **10**, 5353 (2008).
[24]C. R. Jacob, J. Neugebauer, L. Jensen, and L. Visscher, Phys. Chem. Chem. Phys. **8**, 2349 (2006).
[25]C. Daday, C. König, O. Valsson, J. Neugebauer, and C. Filippi, J. Chem. Theory Comput. **9**, 2355 (2013).
[26]S. Höfener, A. S. P. Gomes, and L. Visscher, J. Chem. Phys. **136**, 044104 (2012).
[27]J. Neugebauer, Phys. Rep. **489**, 1 (2010).
[28]C. König and J. Neugebauer, ChemPhysChem **13**, 386 (2012).
[29]S. Fux, K. Kiewisch, C. R. Jacob, J. Neugebauer, and M. Reiher, Chem. Phys. Lett. **461**, 353 (2008).
[30]J. Neugebauer, M. J. Louwerse, E. J. Baerends, and T. A. Wesolowski, J. Chem. Phys. **122**, 094115 (2005).
[31]J. Neugebauer, C. R. Jacob, T. A. Wesolowski, and E. J. Baerends, J. Phys. Chem. A **109**, 7805 (2005).
[32]J. Neugebauer and E. J. Baerends, J. Phys. Chem. A **110**, 8786 (2006).
[33]A. W. Götz, S. M. Beyhan, and L. Visscher, J. Chem. Theory Comput. **5**, 3161 (2009).
[34]S. Laricchia, E. Fabiano, L. A. Constantin, and F. Della Sala, J. Chem. Theory Comput. **7**, 2439 (2011).
[35]C. R. Jacob and L. Visscher, J. Chem. Phys. **128**, 155102 (2008).
[36]O. Roncero, A. Zanchet, P. Villarreal, and A. Aguado, J. Chem. Phys. **131**, 234110 (2009).
[37]P. de Silva and T. A. Wesolowski, J. Chem. Phys. **137**, 094110 (2012).
[38]F. R. Manby, M. Stella, J. D. Goodpaster, and T. F. Miller, J. Chem. Theory Comput. **8**, 2564 (2012).
[39]C. Huang, M. Pavone, and E. A. Carter, J. Chem. Phys. **134**, 154110 (2011).
[40]O. Christiansen, P. Jørgensen, and C. Hättig, Int. J. Quantum Chem. **68**, 1 (1998).
[41]O. Christiansen, Theor. Chem. Acc. **116**, 106 (2006).
[42]O. Christiansen and K. V. Mikkelsen, J. Chem. Phys. **110**, 8348 (1999).
[43]O. Christiansen, H. Koch, A. Halkier, P. Jørgensen, T. Helgaker, and A. S. de Merás, J. Chem. Phys. **105**, 6921 (1996).
[44]T. Schwabe, K. Sneskov, J. M. Haugaard Olsen, J. Kongsted, O. Christiansen, and C. Hättig, J. Chem. Theory Comput. **8**, 3274 (2012).
[45]ADF, Amsterdam density functional program, Theoretical Chemistry, Vrije Universiteit Amsterdam, see http://www.scm.com, 2007.
[46]G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).
[47]C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema, and L. Visscher, J. Comput. Chem. **32**, 2328 (2011).
[48]DALTON2011, a molecular electronic structure program (2011), see http://www.daltonprogram.org.
[49]DIRAC, a relativistic *ab initio* electronic structure program, Release DIRAC04.0 (2004), written by H. J. Aa. Jensen, T. Saue, and L. Visscher with contributions from V. Bakken, E. Eliav, T. Enevoldsen, T. Fleig, O. Fossgaard, T. Helgaker, J. Laerdahl, C. V. Larsen, P. Norman, J. Olsen, M. Pernpointner, J. K. Pedersen, K. Ruud, P. Salek, J. N. P. van Stralen, J. Thyssen, O. Visser, and T. Winther; see http://dirac.chem.sdu.dk), 2004.
[50]M. Levy, Phys. Rev. A **26**, 1200 (1982).
[51]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[52]A. Lembarki and H. Chermette, Phys. Rev. A **50**, 5328 (1994).
[53]D. E. Woon and T. H. Dunning, Jr., J. Phys. Chem. **103**, 4572 (1995).
[54]M. Etinski and C. M. Marian, Phys. Chem. Chem. Phys. **12**, 4915 (2010).
[55]E. D. Hedegård and J. Kongsted, private communication (2012).
[56]J. M. Olsen, K. Aidas, K. V. Mikkelsen, and J. Kongsted, J. Chem. Theory Comput. **6**, 249 (2010).
[57]T. H. Dunning, Jr., J. Phys. Chem. **90**, 1007 (1989).
[58]R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Phys. Chem. **96**, 6796 (1992).
[59]P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen, and E. J. Baerends, J. Chem. Phys. **112**, 1344 (2000).
[60]O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, Chem. Phys. Lett. **302**, 199 (1999).
[61]O. V. Gritsenko, P. R. T. Schipper, and E. J. Baerends, Int. J. Quantum Chem. **76**, 407 (2000).
[62]P. T. Van Duijnen, M. Swart, and L. Jensen, "The discrete reaction field approach for calculating solvent effects," in *Solvation Effects on Molecules and Biomolecules*, edited by S. Canuto (Springer, 2008).
[63]J. J. M. Heller, R. N. Hamm, R. D. Birkhoff, and L. R. Painter, J. Chem. Phys. **60**, 3483 (1974).
[64]O. Christiansen, T. M. Nymand, and K. V. Mikkelsen, J. Chem. Phys. **113**, 8101 (2000).
[65]C. Zazza, J. M. Olsen, and J. Kongsted, Computational and Theoretical Chemistry **974**, 109 (2011).
[66]A. DeFusco, J. Ivanic, M. W. Schmidt, and M. S. Gordon, J. Phys. Chem. A **115**, 4574 (2011).
[67]V. Ludwig, K. Coutinho, and S. Canuto, Phys. Chem. Chem. Phys. **9**, 4907 (2007).
[68]C. R. Jacob, S. M. Beyhan, and L. Visscher, J. Chem. Phys. **126**, 234116 (2007).
[69]J. Neugebauer, C. Curutchet, A. Muñoz-Losa, and B. Mennucci, J. Chem. Theory Comput. **6**, 1843 (2010).
[70]L. B. Clark, G. G. Peschel, and I. Tinoco, J. Phys. Chem. **69**, 3615 (1965).
[71]T. Gustavsson, A. Bányász, E. Lazzarotto, D. Markovitsi, G. Scalmani, M. J. Frisch, V. Barone, and R. Improta, J. Am. Chem. Soc. **128**, 607 (2006).

## D.6 Paper XIII

PCCP

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

# Towards systematically improvable models for actinides in condensed phase: the electronic spectrum of uranyl in $Cs_2UO_2Cl_4$ as a test case†

Cite this: *Phys. Chem. Chem. Phys.*, 2013, **15**, 15153

André Severo Pereira Gomes,*[a] Christoph R. Jacob,*[b] Florent Réal,[a] Lucas Visscher[c] and Valérie Vallet[a]

In this work we explore the use of frozen density embedding [Gomes *et al.*, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353] as a way to construct models of increasing sophistication for describing the low-lying electronic absorption spectra of $UO_2^{2+}$ in the $Cs_2UO_2Cl_4$ crystal. We find that a relatively simple embedding model, in which all but the $UO_2^{2+}$ unit are represented by an embedding potential, can already describe the main spectral features and the main environment effects can be attributed to the four chloride ions situated at the $UO_2^{2+}$ equatorial plane. Contributions from species further away, albeit small, are found to be important for reaching a close agreement with experimentally observed quantities such as the excited states' relative positions. These findings suggest that such an embedding approach is a viable alternative to supermolecular calculations employing larger models of actinyl species in condensed phase. Nevertheless, we observe a slight red shift of the excitation energies calculated with our models compared to experimental results, and attribute this discrepancy to inaccuracies in the underlying structural parameters.

Received 17th May 2013,
Accepted 11th July 2013

DOI: 10.1039/c3cp52090k

www.rsc.org/pccp



## 1 Introduction

Optical spectroscopy is a powerful probe of the interactions between the constituents of molecular complexes containing actinide species, as well as the interaction of such complexes and their surroundings. However, actinide species are often difficult to manipulate due to their radiotoxicity and may present rather complicated spectra. Therefore, the interpretation of experimental results is greatly helped by the use of theoretical models that provide detailed information on the electronic structure. This may, for instance, aid in deconvoluting the measured spectra in solution into the contributions of different species that may coexist in equilibrium.[1,2]

Theoretical (semi)empirical approaches based upon crystal or ligand-field theory[3–5] provide a simple physical picture based on effective Hamiltonians. This makes such methods the first choice for the interpretation of experimental results. However, as their accuracy depends on the validity of the simple model and the quality of the experimental data used in the parametrization, their predictive power is limited.

The computationally much more demanding *ab initio* electronic structure approach[6–8] based on wavefunction theory (WFT) or density functional theory (DFT) can, for small models, provide accurate non-empirical data for small model systems. This was demonstrated by a number of studies over the past two decades.[9–30] However, such studies also demonstrate that the accurate prediction of the energies of electronically excited states is a very demanding task. Often, one passes from a relatively simple, closed-shell ground state to excited states which have contributions from several close-lying open-shell configurations. This makes the balanced treatment of ground and excited states difficult and may give rise to substantial errors in the calculated transition energies. Fortunately, one often finds that differential correlation effects for excited states of similar character are smaller, so that their relative energies can be described with less effort than absolute ones.

These difficulties are illustrated in the case of the actinyl(VI) and (V) ions ($AnO_2^{2+/+}$, An = U, Np, Pu, Am), which are very stable species that are found both in the gas and in the condensed phases (in particular in solution).[31] For those systems, it is well-established[31] that the low-lying excited states arise from

[a] *Université de Lille 1, Laboratoire PhLAM, CNRS UMR 8523, CNRS FR 2416, Bât P5, F-59655 Villeneuve d'Ascq Cedex, France. E-mail: andre.gomes@univ-lille1.fr; Fax: +33 3 2033 7020; Tel: +33 3 2043 4980*
[b] *Karlsruhe Institute of Technology (KIT), Center for Functional Nanostructures and Institute of Physical Chemistry, Wolfgang-Gaede-Str. 1a, 76131 Karlsruhe, Germany. E-mail: christoph.jacob@kit.edu*
[c] *Amsterdam Center for Multiscale Modeling, Department of Theoretical Chemistry, Faculty of Sciences, Vrije Universiteit Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, Netherlands. E-mail: l.visscher@vu.nl*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3cp52090k

excitations to unoccupied non-bonding ($5f_\delta$, $5f_\phi$) actinide orbitals from the (i) actinyl bonding orbitals (for An = U(VI)), in internal uranyl ligand-to-metal charge-transfer (LMCT) excitations; and from the (ii) partially filled f (for An = U(V), Np, Pu, Am), in the so-called f–f transitions. Somewhat higher in energy one also finds (iii) LMCT transitions from ligands bound in the equatorial plane. The latter, even when not significantly contributing to excitations of types (i) and (ii), can affect their energies and oscillator strengths.[32,33] This makes the description of these nearest ligands essential in investigations on condensed phases, while also second-nearest neighbors such as outer solvation shells can still differentially affect electronic states (*via*, *e.g.*, hydrogen bonding or by aggregation[34]).

For the investigation of excitations of types (i)–(iii) with WFT methods employing many-electron model spaces (*e.g.*, CAS or RAS)[16,24] consideration of just the equatorial ligands directly bound to the actinyls is already at the limit of what is computationally feasible. A more approximate treatment of the excitations from the ligands[16,19,35] is therefore often attempted to reduce the computational cost. In addition, idealized geometries are used to make as much use of symmetry as possible. Time-dependent DFT (TD-DFT) approaches offer a computationally less demanding alternative, but are often considered as too inaccurate for actinyl systems due to the known tendency of most available exchange–correlation functionals to spuriously stabilize delocalized charges.[36–39] Recent studies[17,18] using the CAM-B3LYP[40] functional indicated that range-separated hybrid functionals may yield quantitatively correct spectra for uranyl complexes. Nevertheless, (TD-)DFT remains inapplicable to f–f spectra (*i.e.*, excitations of type (ii)) because of the multireference character of the ground states.

A way to overcome the size restriction for WFT approaches is to resort to embedding techniques[41] such as the frozen density embedding (FDE) method.[42,43] With FDE, a large system is partitioned into smaller subsystems interacting through a so-called embedding potential. This embedding potential is determined from the densities of the individual subsystems. The subsystems can then be treated exclusively with DFT (DFT-in-DFT),[44–49] or one (or a few) subsystems can be treated with WFT (WFT-in-DFT),[50–55] depending on the balance between computational cost and accuracy one wishes to achieve. An interesting aspect of FDE is that, by retaining a fully quantum mechanical description for all subsystems, both ground and excited state properties for the whole system remain in principle accessible.[56–59] This has led to the development of efficient, DFT-based protocols to investigate the coupling of localized electronic excitations in large systems.[60–62]

Some of us have previously employed WFT-in-DFT embedding to study the f–f spectrum [*i.e.*, excited states of type (ii)] of $NpO_2^{2+}$ as an impurity in a $Cs_2UO_2Cl_4$ crystal,[53] using FDE to construct a (frozen) model for the crystal environment surrounding a central neptunyl ($NpO_2^{2+}$) or neptunyl tetrachloride ($NpO_2Cl_4^{2-}$) unit. Subsequently, some of us have applied DFT-in-DFT and WFT-in-DFT embedding to investigate the low-lying excited states of the CUO molecule, which is isoelectronic to $UO_2^{2+}$, surrounded by noble gas atoms in its equatorial plane.[63] These excited states are of type (i), *i.e.*, internal LMCT-type excitations. While for these LMCT-type excitations in CUO it turned out that the limited accuracy of the (orbital-free) kinetic energy functionals prevented an accurate description of the noble-gas actinide species already for the ground state, we have found for the f–f spectrum of $NpO_2^{2+}$ that a simple embedding model where the chlorides are represented by an FDE embedding potential did yield accurate results. This is due to the intrinsically localized nature of f–f transitions where ground and low-lying excited states are dominated by molecular spinors with strong contributions from Np-centered 5f spinors. These previous findings raise the question to what extent actinyl spectra can be modeled with WFT-in-DFT embedding approaches, and what accuracy can be reached using the minimal model discussed above for excitations other than f–f excitations.

Thus, our aim in this paper is to further evaluate the performance of FDE-based embedding schemes by investigating the spectra of uranyl tetrachloride ($UO_2Cl_4^{2-}$). This species is known to play a key role in pyroreprocessing techniques of the spent nuclear fuels due to the use of high-temperature chloride melts.[64] There is a wealth of accurate experimental spectra available both for the $Cs_2UO_2Cl_4$ crystal[65–68] and for $UO_2Cl_4^{2-}$ in non-aqueous solvents[69,70] and in the gas phase.[71] We believe that calculating the spectra of $UO_2^{2+}$ in $Cs_2UO_2Cl_4$ represents an interesting test for subsystem models since the low-lying transitions are again of type (i), with potentially important contributions from the ligand to the occupied orbitals involved. Moreover, it presents an opportunity to further investigate the performance of different electronic structure methods in conditions that closely mimic those of experiments. Of all previous *ab initio* studies,[10,18–20,71,72] only the one by Matsika and Pitzer[10] calculated the spectrum with the inclusion of a model for the crystal environment.

## 2 Computational methods

All electronic spectra calculations were performed at the experimental X-ray structure[73] (see Section 3.1 for details) with a development version of the DIRAC electronic structure code[74] (revision ab65b36), employing Dyall's basis sets[75] of triple-zeta quality for uranium, and Dunning's aug-cc-pVTZ sets[76] for oxygen and chlorine, all of which are left uncontracted.

The Dirac–Coulomb (DC) Hamiltonian was used throughout, along with the usual approximation of the (SS|SS)-type integrals by a point charge model.[77] In (TD-)DFT calculations the CAM-B3LYP[40] functional was used, whereas the WFT approach employed here is the intermediate Hamiltonian Fock-space coupled cluster method (IHFSCCSD),[78–81] which allows for a proper description of a possible multiconfigurational nature of excited states.

In the IHFSCCSD calculations the excitation energies were obtained with the (1h,1p) sector of Fock-space, meaning that in the process electron attachment and ionization energies were also calculated *via* the (0h,1p) and (1h,0p) sectors of Fock-space, respectively. For cases in which we have not been able to obtain solutions for the (1h,1p) sector due to the presence of

intruder states, we have resorted to an approximate treatment of the CC amplitudes in this sector, which are determined in a manner akin to that of MP2 – in practice by performing a single CC iteration for the (1h,1p) sector after having converged the preceding sectors. We shall here denote this approach by the IHFSCC-112 acronym.

Due to computational constraints we are forced to truncate the virtual space in the WFT calculations. In order to verify the effect of this truncation, we have explored three different correlating ($Q$) spaces by considering all uranyl spinors with energies between (1) −6.0 and 5.0 a.u.; (2) −6.0 and 12.0 a.u.; and (3) −6.0 and 100.0 a.u. This way, the occupied 5d spinors are always correlated (yielding a total of 34 correlated electrons), with up to 446 virtual spinors. The IHFS model ($P$) spaces were the same for all correlation spaces. These are slightly modified compared to those employed in prior work,[17,27] and contain at least the 6d and 5f spinors. More details on the definition of the $P$ and $Q$ spaces can be found in Table 1. Information on the computational cost of these calculations can be found in Table S1 in the ESI.†

We note that we were not able to obtain a Fock-space reference spectrum for uranyl tetrachloride, due to difficulties in performing calculations with the large active spaces required to take into account the ligand (occupied and virtual) spinors that have energies in between the occupied and virtual spinors involved in the excitations.

DFT-in-DFT embedding calculations[42,44,47] were performed with the ADF[82] code *via* the PyADF scripting environment.[83] In the calculations, the spin-free (SF) ZORA[84,85] Hamiltonian was used along with the corresponding TZ2P basis sets[86] for uranium, oxygen, and chlorine. We have employed the SAOP[87–89] model potential for the active subsystems, whose densities were allowed to relax through the freeze–thaw procedure until convergence (reached within 20 iterations). The non-additive exchange–correlation and kinetic-energy contributions to the embedding potential were calculated with the PW91[90] exchange–correlation and PW91k[91] kinetic energy functionals, respectively. The integration accuracy parameter in ADF was set to 10. The DFT-in-DFT embedding potentials obtained with ADF and PyADF were subsequently used in Dirac calculations as effective one-electron operators according to the "static" embedding scheme outlined in ref. 53.

**Table 1** IHFSCCSD main model ($P_m$), intermediate model space ($P_i$) and correlation ($Q$) spaces employed for the different models for the environment surrounding the $UO_2^{2+}$ species, given in terms of number of spinor pairs of gerade (ungerade) symmetry in each subspace. The "h" and "p" superscripts denote "holes" and "particles", respectively

| Model | Environment | | $P_i^h$ | $P_m^h$ | $P_m^p$ | $P_i^p$ | $Q^h$ | $Q^p$ |
|---|---|---|---|---|---|---|---|---|
| (**a**) | None | $Q_2$ | 2 (4) | 3 (3) | 5 (10) | 6 (7) | 5 (0) | 101 (93) |
| (**b**) | $Cl_4^{4-}$ p.c. | $Q_2$ | 2 (4) | 3 (3) | 7 (13) | 9 (10) | 5 (0) | 101 (93) |
| (**c**) | $Cl_4^{4-}$ FDE | $Q_1$ | 2 (4) | 3 (3) | 7 (13) | 9 (10) | 5 (0) | 72 (58) |
| | | $Q_2$ | | | | | | 101 (93) |
| | | $Q_3$ | | | | | | 122 (128) |
| (**e**), (**f**) | $Cs_2UO_2Cl_4$ | $Q_1$ | 2 (4) | 3 (3) | 7 (13) | 9 (10) | 5 (0) | 72 (58) |

## 3 Results and discussion

### 3.1 Models for uranyl in $Cs_2UO_2Cl_4$

Our main goal in this work is to explore the construction of models of increasing complexity that can describe the absorption spectra of $Cs_2UO_2Cl_4$. In this crystal, whose structure has been accurately determined from X-ray diffraction studies,[73] the uranyl cation is surrounded by four chlorine atoms. These are oriented along the equatorial plane, but show a $C_{2h}$ site symmetry, as the O–U–O axis intersects the plane defined by the four chlorides with a slight deviation from 90°. The U–O and U–Cl distances are 1.774 Å and 2.671 Å, respectively. Further away from the uranium one finds a shell of cesium atoms at ≃4.6 Å from the central uranium, and the nearest uranium atom at ≃5.8 Å. Therefore, the crystal is made up of well-separated uranyl tetrachloride ($UO_2Cl_4^{2-}$) units interspersed with cesium ions.

The simplest models are (**a**) the bare $UO_2^{2+}$ ion and (**d**) the $UO_2Cl_4^{2-}$ unit, as shown in Fig. 1. In addition, one can consider intermediate models in which only the uranyl species is treated explicitly while the equatorial ligands are included in an approximate fashion. This could be achieved either (**b**) by a simple point-charge embedding[41] or (**c**) by using an FDE-based embedding potential.[53] For all four models, we use an idealized structure with $D_{4h}$ symmetry instead of the $C_{2h}$ point group corresponding to the crystal's site symmetry. This will simplify our analysis and allow for a direct comparison to calculations in the literature. In the calculations, only the Abelian point group $D_{2h}$ can be used (instead of $D_{4h}$) and, therefore, the irreducible representations of $D_{2h}$ are used to label the excited states in our tables.

More sophisticated models, shown in Fig. 2, extend those above and include effects arising from the long-range interactions with the crystal lattice. The structure of the crystal

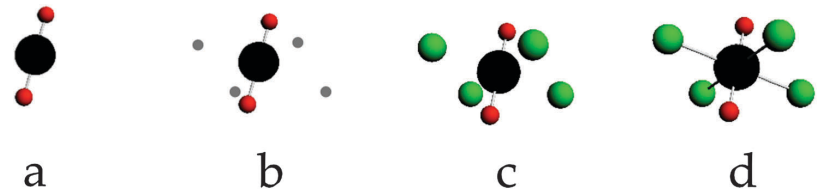



**Fig. 1** Models without the crystal environment: (**a**) the bare uranyl species; (**b**) uranyl with point-charge embedding; (**c**) uranyl with FDE embedding; and (**d**) the uranyl chloride species (uranium: black; oxygen: red; chlorine: green; point charges: grey).

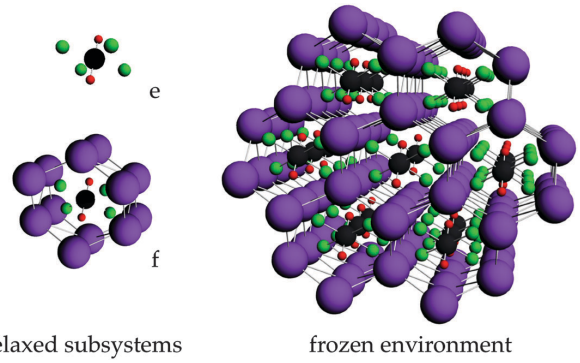



**Fig. 2** Uranyl FDE embedding models including the crystal environment (shown on the right), where one relaxes (**e**) only the nearest chlorides; (**f**) the nearest chlorides and 12 cesium ions (uranium: black; oxygen: red; cesium: purple; chlorine: green).

suggests a natural subdivision of $Cs_2UO_2Cl_4$ into three distinct regions – an inner one containing the core model(s) above, an intermediate one containing at least the shell of the nearest cesium atoms, and finally the remainder of the crystal. Our first model (**e**) is built using the same strategy as in ref. 53. The basic representation of the crystal environment is the combination of the intermediate region (comprising 20 $UO_2Cl_4^{2-}$ and 90 $Cs^+$ ions) around the central unit and an array of point charges is defined to represent the Madelung potential due to the rest of the crystal.[92] The electron density of the intermediate region, which is necessary to determine the corresponding FDE embedding potential, is obtained as the sum of the densities from DFT calculations (SAOP/TZP) on the isolated ($UO_2Cl_4^{2-}$ and $Cs^+$) species. In a more refined model (**f**) the density for the 12 cesium ions nearest to the uranyl species was allowed to relax through a DFT-in-DFT freeze-and-thaw procedure. Both models (**e**) and (**f**) are derived from the experimental crystal structure, in which the central uranyl unit only has $C_{2h}$ symmetry.

### 3.2 Assessing the models without a crystal environment

First, we consider the models without the crystal environment. To assess the accuracy that can be achieved with these models, we will compare the electronic spectrum of uranyl tetrachloride (model **d**) to those obtained with the approximate models (**a**), (**b**) and (**c**) (see Fig. 1).

**3.2.1 The electronic structure of $UO_2Cl_4^{2-}$.** Before discussing the approximate models, we recall some key findings from calculations on uranyl tetrachloride.[10,16,18–20,71] First, theoretical and experimental[31,65] works assign the spectrum below $\simeq$30 000 $cm^{-1}$ to excitations local to the uranyl species [*i.e.*, excitations of type (i) in the classification introduced above]. These involve the highest occupied *ungerade* orbital and part of the virtual uranium f manifold. LMCT from the chloride ligands [*i.e.*, excitations of type (iii)] occurs at somewhat higher energy and concerns excitations from essentially pure chloride ligand orbitals.[31,71] This reflects the fact that the U–Cl bonds, on the basis of experimental results and Mulliken population analyses from correlated wavefunctions,[19] are considered to be largely ionic and have only weak covalent character. Recent AIM studies[93,94] corroborate this picture, although the experimentally determined densities used in ref. 93 seem to indicate somewhat stronger covalency. One should keep in mind, however, that this is not an intrinsic characteristic of U–Cl bonds; for instance, in compounds not containing the actinyl group such as metallocene dichlorides, there is evidence that U–Cl covalency can be substantial.[95]

Rather good agreement is found[18] between WFT and DFT calculations with the CAM-B3LYP exchange–correlation functional, with a few notable exceptions: CAM-B3LYP reorders some states (in $D_{4h}$ notation) with respect to CASPT2[16,19] as well as to the experimental assignment. In particular, the first $B_{1g}$ and $B_{2g}$ states (both of $\sigma_{1/2u} \rightarrow f_\phi$ character) and the second $B_{1g}$ and $B_{2g}$ states (both of $\sigma_{1/2u} \rightarrow f_\delta$ character) are each interchanged. Apart from these discrepancies, there is also a crossing between the first $E_g$ (of $\sigma_{1/2u} \rightarrow f_\delta$ character) and $B_{1g}$ (for CAM-B3LYP)/$B_{2g}$ (CASPT2) states found at a U–O distance of $\simeq$1.83 Å for the former and of $\simeq$1.86 Å for the latter.

**3.2.2 Approximate models, DFT.** Proceeding now with an analysis of the simplest models (**a–c**), we take the CAM-B3LYP results for model (**d**), *i.e.*, the full uranyl tetrachloride, as the reference. All values are given in Table 2, where we order the electronic states according to the experimental classification.[31,65] We note that the experimental excitation energies for the twelve

**Table 2** CAM-B3LYP excitation energies (columns ''abs.'') in wavenumbers for different uranyl models (**a–c**) and uranyl chloride (model **d**), without the presence of the crystal environment ($r_{U-O}$ = 1.774 Å). The energies relative to the first excited state are also shown (columns ''rel.'')

| | | Model (**a**) | | Model (**b**) | | Model (**c**) | | Model (**d**) | |
|---|---|---|---|---|---|---|---|---|---|
| State | Label ($D_{2h}$) | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. |
| I, II | $B_{2g}$, $B_{3g}$ | 13 215 | | 17 265 | | 18 115 | | 19 018 | |
| III | $B_{1g}$ | 11 805 | −1410 | 16 341 | −924 | 18 321 | 206 | 19 934 | 917 |
| IV | $A_g$ | 11 805 | −1410 | 16 239 | −1026 | 17 981 | −134 | 19 288 | 270 |
| V, VI | $B_{2g}$, $B_{3g}$ | 15 135 | 1920 | 17 681 | 416 | 19 565 | 1451 | 20 970 | 1952 |
| VII | $A_g$ | 17 084 | 3869 | 19 394 | 2129 | 20 539 | 2424 | 21 745 | 2728 |
| VIII | $B_{1g}$ | 17 084 | 3869 | 19 660 | 2395 | 20 829 | 2714 | 21 592 | 2574 |
| IX, X | $B_{2g}$, $B_{3g}$ | 20 461 | 7246 | 23 424 | 6159 | 24 747 | 6633 | 25 531 | 6513 |
| XI | $B_{1g}$ | 18 896 | 5681 | 23 875 | 6610 | 26 137 | 8022 | 27 058 | 8040 |
| XII | $A_g$ | 18 896 | 5681 | 23 876 | 6611 | 26 140 | 8025 | 27 112 | 8094 |

lowest excited states discussed in this paper can be found in Table 5.

Model (**a**), the isolated uranyl unit, is obviously the least suitable model and places states III and IV below the doubly degenerate (I–II) first excited state of uranyl tetrachloride. In addition, the relative order of the highest states (IX, X *vs.* XI and XII) is different relative to the reference model (**d**). The strong red shift ($\simeq$7000 cm$^{-1}$) of all states with respect to model (**d**) can be easily understood: the isolated uranyl cation has a much shorter bond length so that at $r_{U-O}$ = 1.774 Å the ground and excited states—which have longer bond lengths—are calculated to have a very small energy gap. The great sensitivity of the vertical excitation energies to the bond length for this system is also evident if our results for model (**d**) are compared to those of Tecmer *et al.*[18] With the shorter bond length of $r_{U-O}$ = 1.764 Å used in this study, the excitation energies are 1000–3000 cm$^{-1}$ higher than the one found here.

These flaws of the simplest model are considerably reduced in the point-charge embedding model (**b**). The lowest four states still resemble the spectrum for the isolated species. In particular, the ordering of the states is at odds with that for uranyl tetrachloride and almost no splitting between the lowest non-degenerate (III–IV) states is found. On the other hand, the highest four states (IX–XII) now follow the ordering of the reference calculation for model (**d**), even though they remain too close to each other. The relative positioning of the intermediate (V–VIII) states among themselves also very much resembles that of the reference. We note that states VII and VIII now come in the same order as in WFT calculations and in experiments. By including the point charges the overall red shift of the spectrum is now about $\simeq$3500 cm$^{-1}$, i.e., only half of that for uranyl.

Model (**c**), FDE embedding, represents a significant improvement over the point charge approach. The energies of most states approach those of the reference calculation. The most important remaining discrepancy is the position of state IV, which still appears as the lowest excited state. The overall red shift is reduced to about 1000 cm$^{-1}$ and we believe this remaining discrepancies can be attributed to the still too short U–O ground state bond length predicted by this model. If we compare excitation energies relative to the first excited state (columns ''rel'' in Table 2) we find much better agreement with the reference calculation than for models (**a**) and (**b**).

**3.2.3 Approximate models, WFT.** The above comparison with supermolecular DFT results is useful to assess errors in the embedding approach. However, also with the relatively well-performing CAM-B3LYP functional, TD-DFT cannot capture the subtle correlation effects determining the details of the spectrum, such as the relative order of states VII and VIII. This order is fortuitously corrected in our approximate models whereas the reference model (**d**) has them in the wrong order compared to the experimental assignment. We will now consider a more advanced treatment of electron correlation. To this end, we begin by addressing the different choices of correlation spaces for the IHFSCCSD method.

We present the electronic spectrum of the FDE embedding model (**c**) for different $Q$ spaces in Table 3, along with IHFSCC-112 results for the smallest $Q$ space. It is clear that as $Q$ becomes larger, absolute IHFSCCSD excitation energies as well as the spacing between adjacent electronic states become smaller, while there are no significant changes in the composition of the states (for further details see Table S7 in the ESI†). Because the smaller $Q_2$ active space yields results close to the largest space, $Q_3$, for both absolute energies and spacings, one can consider the former as sufficiently accurate for evaluating different structural models. Reducing $Q_2$ further to $Q_1$ gives errors of about 1200–1400 cm$^{-1}$ for absolute energies and up to 300 cm$^{-1}$ for the spacings of higher-lying states. The sensitivity of the outcome to the amount of electron correlation that is included is also visible in the differences between IHFSCCSD/$Q_1$ and IHFSCC-112/$Q_1$, where the spacings between levels and the composition of the states are rather similar but the absolute energies differ by about 1000 cm$^{-1}$.

Next, we turn to the results of IHFSCCSD/$Q_2$ calculations on the different approximate models, which can be found in Table 4. As we do not have IHFSCCSD results available for uranyl chloride [model (**d**)], we also list the vertical SO-CASPT2 excitation energies of ref. 19 in Table 4 in order to provide a

**Table 3** IHFSCCSD excitation energies (columns ''abs.'') in wavenumbers for the FDE embedding uranyl model (**c**) without the presence of the crystal environment ($r_{U-O}$ = 1.774 Å), employing different $Q$ spaces (see Table 1). The energies relative to the first excited state are also shown (columns ''rel.'')

| | | IHFSCCSD-112/$Q_1$ | | IHFSCCSD/$Q_1$ | | IHFSCCSD/$Q_2$ | | IHFSCCSD/$Q_3$ | |
|---|---|---|---|---|---|---|---|---|---|
| State | Label ($D_{2h}$) | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. |
| I, II | $B_{2g}$, $B_{3g}$ | 17 998 | 0 | 16 896 | 0 | 15 746 | 0 | 15 680 | 0 |
| III | $B_{1g}$ | 18 705 | 707 | 17 624 | 728 | 16 432 | 686 | 16 365 | 685 |
| IV | $A_g$ | 19 409 | 1411 | 18 400 | 1504 | 17 116 | 1370 | 17 043 | 1363 |
| V, VI | $B_{2g}$, $B_{3g}$ | 20 689 | 2691 | 19 696 | 2800 | 18 389 | 2643 | 18 318 | 2637 |
| VII | $A_g$ | 21 797 | 3800 | 20 834 | 3938 | 19 400 | 3654 | 19 323 | 3643 |
| VIII | $B_{1g}$ | 21 855 | 3858 | 20 915 | 4019 | 19 448 | 3702 | 19 370 | 3690 |
| IX, X | $B_{2g}$, $B_{3g}$ | 25 131 | 7134 | 24 108 | 7212 | 22 805 | 7059 | 22 746 | 7065 |
| XI | $B_{1g}$ | 27 602 | 9604 | 26 626 | 9730 | 25 218 | 9472 | 25 155 | 9475 |
| XII | $A_g$ | 27 603 | 9606 | 26 628 | 9732 | 25 220 | 9474 | 25 157 | 9477 |

**Table 4** IHFSCCSD/$Q_2$ excitation energies (columns ''abs.'') in wavenumbers for different uranyl models (**a–c**) without the presence of the crystal environment ($r_{U-O}$ = 1.774 Å). The energies relative to the first excited state are also shown (columns ''rel.''). As there are no IHFSCCSD results for uranyl chloride (model **d**), we include here the vertical SO-CASPT2 excitation energies of Pierloot and van Besien[19] for comparison

| | | Model (**a**) | | Model (**b**) | | Model (**c**) | | Model (**d**), [a]ref. 19 | |
|---|---|---|---|---|---|---|---|---|---|
| State | Label ($D_{2h}$) | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. |
| I, II | $B_{2g}$, $B_{3g}$ | 12 296 | | 14 757 | | 15 746 | | 21 024 | |
| III | $B_{1g}$ | 11 105 | −1191 | 15 132 | 375 | 16 432 | 686 | 21 273 | 249 |
| IV | $A_g$ | 11 105 | −1191 | 15 421 | 664 | 17 116 | 1370 | 22 125 | 1101 |
| V, VI | $B_{2g}$, $B_{3g}$ | 12 303 | 7 | 16 620 | 1863 | 18 389 | 2643 | 22 859 | 1835 |
| VII | $A_g$ | 14 426 | 2130 | 17 600 | 2843 | 19 400 | 3654 | 24 056 | 3032 |
| VIII | $B_{1g}$ | 14 426 | 2130 | 17 687 | 2930 | 19 448 | 3702 | 24 339 | 3315 |
| IX, X | $B_{2g}$, $B_{3g}$ | 17 593 | 5297 | 21 061 | 6304 | 22 805 | 7059 | 27 494 | 6470 |
| XI | $B_{1g}$ | 17 659 | 5363 | 22 829 | 8072 | 25 218 | 9472 | 29 842 | 8818 |
| XII | $A_g$ | 17 659 | 5363 | 22 829 | 8073 | 25 220 | 9474 | 29 849 | 8825 |

[a] Vertical SO-CASPT2 results[19] calculated for $r_{U-O}$ = 1.783 Å and $r_{U-Cl}$ = 2.712 Å.

comparison to a supermolecular WFT calculation. One should, however, focus more on comparing trends for the spectra rather than absolute values for two reasons. First, the SO-CASPT2 calculations were performed for a slightly different geometry ($r_{U-O}$ = 1.783 Å and $r_{U-Cl}$ = 2.712 Å). From the discussion above as well as from previous studies of uranyl[21,27] it is apparent that small changes in geometry may correspond to large changes in the vertical excitation energies. Second, there is evidence in the literature[17,21,27] that CASPT2 consistently places equivalent excitations at higher energies than IHFSCCSD for a given geometry, while spacings between excited levels are often in good agreement between the two methods.

As was the case for TD-DFT, we see a steady improvement in the agreement on the absolute excitation energies between the approximate models (**a–c**) and the reference (**d**). We observe differences ($\simeq$3000–4000 cm$^{-1}$) similar to those found for CAM-B3LYP between the bare (**a**) and point-charge embedded (**b**) models, and slightly smaller ($\simeq$1000–2000 cm$^{-1}$) ones between the latter and the FDE model (**c**). The isolated uranyl model (**a**) yields once more rather low values for the excitation energies. As for the TD-DFT calculations discussed above, we attributed these differences to the bond length employed here, which is nearly 0.1 Å larger than the CCSD equilibrium value[27] of 1.685 Å. We observe a qualitative agreement between the IHFSCCSD and CAM-B3LYP excitation energies, for instance with the first $E_g$ state higher than the $B_{1g}$ or $A_g$ states, but note that these are typically $\simeq$1000 cm$^{-1}$ lower for IHFSCCSD than for the corresponding CAM-B3LYP ones. This agrees with earlier observations.[17] For both the point-charge and FDE embedded models, excitation energies are still red-shifted compared to experiments. Nevertheless, in contrast to CAM-B3LYP, IHFSCCSD provides the correct order of the low-lying states.

Furthermore, the spacings between levels are generally larger than those for CAM-B3LYP and are in rather good agreement with the SO-CASPT2 results with the exception of states VII and VIII, which are much closer together for IHFSCCSD than for CASPT2. Another interesting difference with respect to CAM-B3LYP is that here models (**b**) and (**c**) exhibit differences of similar magnitude relative to the CASPT2 $E^{rel}$ values [$\Delta E^{rel}_{(b)} \simeq -315$ cm$^{-1}$ and $\Delta E^{rel}_{(c)} \simeq 552$ cm$^{-1}$, respectively], but with a standard deviation for $E^{rel}_{(c)}$ of about half of that for $E^{rel}_{(b)}$ [$\sigma^{rel}_{(b)} \simeq 306$ cm$^{-1}$ and $\sigma^{rel}_{(c)} \simeq 163$ cm$^{-1}$, respectively]. For this reason, we think model (**c**) indeed yields an improvement over (**b**). That said, the overall agreement with respect to spacings between models (**b–d**) can be related to a similar composition of the excited states' wavefunctions. To that end, it is instructive to compare the composition of the states (see Tables S6 and S7 in the ESI†) to their analogues in ref. 19. From this comparison, we observe rather similar ratios between the contributions of $\sigma^1\delta^1$ and $\sigma^1\phi^1$ character to the different states for CASPT2 and IHFSCCSD. However, for some excitations (*e.g.* IV, V–VI, IX–X) IHFSCCSD gives higher weights to the latter, as well as to $\pi^1\delta^1$ and $\pi^1\phi^1$ configurations.

### 3.3 Inclusion of the crystal environment

From the discussion above we believe we can consider the FDE embedded uranyl model (**c**) as a sufficiently accurate representation of the uranyl tetrachloride species (**d**), and, given the significant differences in computational costs between the two (see Table S1 in the ESI†) as well as the need to use a lower ($C_{2h}$) symmetry when considering the crystal environment, from now on we only consider FDE embedded $UO_2^{2+}$ models. Our results are presented in Table 5.

We observe that, when passing from an idealized geometry (model **c**) to the experimental one (model **c′**), there is little change in the excited states' energies. The only difference is the splitting of the doubly degenerate states in the former (I–II, V–VI and IX–X respectively). In the case of CAM-B3LYP calculations, already at this stage the magnitude of the splitting for the lowest two states matches quite well the experimentally observed one, while for higher states there is an underestimation. Adding the crystal environment brings about a nearly homogeneous stabilization of occupied and virtual orbitals with respect to the isolated uranyl chloride species. Therefore, one sees only relatively small changes in the electronic spectrum for the crystal models.

Model (**e**), in which the nearest cesium atoms are not relaxed, changes the excitation energies relative to those of model (**c′**) by

**Table 5** CAM-B3LYP and IHFSCC-112/$Q_1$ excitation energies (columns ''abs.''), in wavenumbers, at the experimental geometry ($C_{2h}$) for models (**c′**), (**e**) and (**f**). The energies relative to the first excited state are also shown (columns ''rel.'') for each case

| | | CAM-B3LYP | | | | | | IHFSCC-112/$Q_1$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Model (**c′**) | | Model (**e**) | | Model (**f**) | | Model (**c′**) | | Model (**f**) | | Ref. 10 | | Exp. ref. 31 and 65 | |
| State | Label ($D_{2h}$) | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. | abs. | rel. |
| I | $B_{2g}$ | 18 114 | | 18 134 | | 18 119 | | 18 151 | | 18 128 | | 20 364 | | 20095.7 | |
| II | $B_{3g}$ | 18 112 | −2 | 18 136 | 2 | 18 120 | 1 | 18 154 | 3 | 18 124 | −4 | 20 363 | −1 | 20097.3 | 1.6 |
| III | $B_{1g}$ | 17 975 | −139 | 17 938 | −196 | 17 913 | −206 | 18 874 | 723 | 18 816 | 688 | 21 013 | 649 | 20406.5 | 310.8 |
| IV | $A_g$ | 18 317 | 203 | 18 263 | 129 | 18 236 | 117 | 19 552 | 1401 | 19 492 | 1364 | 21 838 | 1474 | 21316 | 1220.3 |
| V | $B_{2g}$ | 19 568 | 1454 | 19 520 | 1386 | 19 494 | 1375 | 20 836 | 2685 | 20 760 | 2632 | 22 808 | 2444 | 22026.1 | 1930.4 |
| VI | $B_{3g}$ | 19 552 | 1438 | 19 501 | 1367 | 19 475 | 1356 | 20 843 | 2692 | 20 768 | 2640 | 22 830 | 2466 | 22076 | 1980.3 |
| VII | $A_g$ | 20 536 | 2422 | 20 514 | 2380 | 20 494 | 2375 | 21 944 | 3793 | 21 848 | 3720 | 24 618 | 4254 | 22406 | 2310.3 |
| VIII | $B_{1g}$ | 20 825 | 2711 | 20 826 | 2692 | 20 808 | 2689 | 22 005 | 3854 | 21 905 | 3777 | 24 780 | 4416 | 22750 | 2654.3 |
| IX | $B_{2g}$ | 24 749 | 6635 | 24 733 | 6599 | 24 711 | 6592 | 25 297 | 7146 | 25 185 | 7057 | 26 763 | 6399 | 26197.3 | 6101.6 |
| X | $B_{3g}$ | 24 738 | 6624 | 24 719 | 6585 | 24 698 | 6579 | 25 307 | 7156 | 25 201 | 7073 | 26 871 | 6507 | 26247.3 | 6151.6 |
| XI | $B_{1g}$ | 26 131 | 8017 | 26 045 | 7911 | 26 014 | 7895 | 27 779 | 9628 | 27 634 | 9506 | 29 169 | 8805 | 27719.6 | 7623.9 |
| XII | $A_g$ | 26 134 | 8020 | 26 048 | 7914 | 26 017 | 7898 | 27 781 | 9630 | 27 637 | 9509 | 29 145 | 8781 | 27757 | 7661.3 |

no more than $\simeq$90 cm$^{-1}$. Generally, it decreases the excitation energies and the largest (in magnitude) changes are observed for the highest states considered. Relaxing the nearest cesiums (model **f**) accentuates this tendency by an additional lowering of about $\simeq$23 cm$^{-1}$. Therefore, we conclude that, as expected, crystal contributions are an order of magnitude smaller than those of the equatorial ligands. The small differential effect from the crystal implies that CAM-B3LYP continues to underestimate the experimental excitation energies by about 1800–2000 cm$^{-1}$, as it was the case for model (**d**) discussed above. For the relative spacings between states these small effects captured by the models for the crystal are more significant.

Similar trends are found in the WFT-in-DFT results. We should note, however, that in these calculations it was difficult to converge the coupled cluster amplitudes for the (1h,1p) sector, so that we had to employ exclusively the IHFSCC-112 approximation discussed previously. Due to the increase in computational costs caused by the lower symmetry of the central uranyl unit, we could only employ the $Q_1$ space. Consequently, the absolute excitation energies in Table 5 are probably overestimating the calculation of IHFSCCSD/$Q_2$ quality by about 2000 cm$^{-1}$ for excitations and 100 cm$^{-1}$ for spacings.

The IHFSCC-112 excited states' composition for the uranyl embedded in the chlorides (shown in Table S8 of the ESI†) is essentially the same as that in the idealized structure discussed above. In the presence of the crystal environment, we observe that the lowest four states show more equivalent contributions from $\{\sigma^1\delta^1,\pi^1\delta^1\}$ and $\{\sigma^1\phi^1,\pi^1\phi^1\}$ configurations than the isolated models, whereas for higher states the same picture is found for all models.

The only other study which considers the effect of the crystal environment in detail is due to Matsika and Pitzer.[10] They combined a SO-MRCI description of the central uranyl tetrachloride with an embedded cluster model in which pseudopotentials are used to represent the six nearest-neighbor cesium ions, while the Madelung potential arising from the rest of the crystal was represented by an array of point charges. Compared to these results, the CAM-B3LYP calculations better describe the excited state spacings for all but the third and fourth excited states. On the other hand, the results obtained with Fock-space are generally of similar quality to the SO-MRCI ones, with better agreement with the experimental results for some of the lower states, but with a strong underestimation of the energy difference between states VII and VIII.

As we can establish from our models that interactions between the central uranyl chloride unit and further species are relatively small, the main source of errors in our calculations is then likely to be due to the treatment of electron correlation. There, we observe that neither CAM-B3LYP nor any of the *ab initio* approaches employed so far to investigate $Cs_2UO_2Cl_4$ or the bare uranyl chloride species are able to achieve "spectroscopic" accuracy for the absolute excitation energies of $Cs_2UO_2Cl_4$, by which we mean discrepancies between theory and experiment of $\simeq$50–100 cm$^{-1}$ for the low-lying transitions (due to the extremely good resolution of the experimental data). This underscores both the difficulty of determining such spectra from first principles and the need to investigate the performance of higher accuracy methods (*e.g.* those including triple or higher excitations explicitly) which, albeit too costly to be employed in routine calculations, might nevertheless provide insight into the factors controlling the accuracy of more widely applicable approaches (*e.g.* DFT, CASPT2 or IHFSCCSD) and help devise more efficient and accurate approximations.

### 3.4 A closer look at the occupied spinors

Apart from the analysis of the optical spectra above, it is also instructive to compare orbital energies between the models to gauge the accuracy of the embedding. Furthermore, the occupied orbital energies can be compared with experimental studies of ionization energies.

**3.4.1 Uranyl (chloride) in the gas phase.** We start with CAM-B3LYP, for which we show in Table 6 the orbital energies ($\varepsilon$, in eV) for the valence region ($-\varepsilon < 6$ eV) for models (**c**) and (**d**). Here we note that for the FDE-embedded model (**c**), we have orbital energies for both the $UO_2^{2+}$ and $Cl_4^{4-}$ subsystems because of the employed freeze-and-thaw procedure. We see a very good agreement for the outer orbitals between the two models. In particular, the HOMO energy for model (**d**) agrees closely with the one of the $Cl_4^{4-}$ fragment in model (**c**). This is understandable, since calculations on uranyl tetrachloride identify the HOMO essentially as a ligand orbital. Therefore, both models can yield approximations of similar quality to the molecule's first ionization potential,[96] with 1.78 eV and 1.72 eV for models (**c**) and (**d**) respectively. These are significantly smaller than the recent gas-phase experimental vertical electron detachment (ED) energy of 2.62 eV,[71] and reflect the fact that DFT calculations can strongly underestimate this quantity.[17]

Discrepancies between the calculations for the two models become larger for the region between $-2.5$ and 5.0 eV. In this region, one starts to see spinors with both chloride and uranyl contributions in model (**d**), whereas such direct mixing is

**Table 6** CAM-B3LYP orbital energies ($\varepsilon$, in eV) for the valence region for models (**c**) and (**d**) in the idealized ($D_{4h}$) structure and the difference between the two ($\Delta\varepsilon$)

| | Model (**c**) | | Model (**d**) | |
|---|---|---|---|---|
| Label | $-\varepsilon$ | Fragment | $-\varepsilon$ | $\Delta\varepsilon$ |
| $e_{1g}$ | 1.78 | $Cl_4^{4-}$ | 1.72 | 0.06 |
| $e_{1u}$ | 2.09 | $Cl_4^{4-}$ | 1.90 | 0.19 |
| $e_{1g}$ | 2.26 | $Cl_4^{4-}$ | 1.97 | 0.29 |
| $e_{1g}$ | 2.27 | $Cl_4^{4-}$ | 1.97 | 0.30 |
| $e_{1u}$ | 2.13 | $Cl_4^{4-}$ | 2.10 | 0.03 |
| $e_{1u}$ | 2.16 | $Cl_4^{4-}$ | 2.24 | −0.08 |
| $e_{1u}$ | 2.45 | $Cl_4^{4-}$ | 2.32 | 0.13 |
| $e_{1u}$ | 3.39 | $Cl_4^{4-}$ | 2.35 | 1.04 |
| $e_{1u}$ | 3.44 | $Cl_4^{4-}$ | 2.61 | 0.83 |
| $e_{1g}$ | 2.55 | $Cl_4^{4-}$ | 3.07 | −0.52 |
| $e_{1g}$ | 2.81 | $Cl_4^{4-}$ | 3.26 | −0.45 |
| $e_{1g}$ | 4.78 | $Cl_4^{4-}$ | 3.52 | 1.26 |
| $e_{1u}$ | 4.24 | $UO_2^{2+}$ | 3.70 | 0.54 |
| $e_{1u}$ | 4.93 | $UO_2^{2+}$ | 4.23 | 0.70 |
| $e_{1u}$ | 5.33 | $UO_2^{2+}$ | 4.49 | 0.84 |
| $e_{1g}$ | 5.38 | $UO_2^{2+}$ | 4.90 | 0.48 |
| $e_{1g}$ | 5.38 | $UO_2^{2+}$ | 5.01 | 0.37 |
| $e_{1g}$ | 5.49 | $UO_2^{2+}$ | 5.09 | 0.40 |

**Table 7** CAM-B3LYP (columns a), IHFSCCSD/$Q_2$ (columns b) and IHFSCCSD/$Q_1$ (columns c) absolute ionization energies (in eV) for different models for embedded uranyl

| | Model (**b**) | | Model (**c**) | | | Model (**f**) | |
|---|---|---|---|---|---|---|---|
| | (**a**) | (**b**) | (**a**) | (**b**) | (**c**) | (**a**) | (**c**) |
| $e_{1u}$ | 6.16 | 7.78 | 4.24 | 5.81 | 5.51 | 9.32 | 10.61 |
| $e_{1u}$ | 6.78 | 8.14 | 4.93 | 6.21 | 6.01 | 9.99 | 11.08 |
| $e_{1u}$ | 7.15 | 8.54 | 5.33 | 6.63 | 6.31 | 10.40 | 11.50 |
| $e_{1g}$ | 7.10 | 8.32 | 5.38 | 6.56 | 6.40 | 10.41 | 11.50 |
| $e_{1g}$ | 7.13 | 8.46 | 5.38 | 6.65 | 6.50 | 10.47 | 11.54 |
| $e_{1g}$ | 7.23 | 8.56 | 5.49 | 6.75 | 6.59 | 10.56 | 11.67 |

absent by construction in model (**c**). Nevertheless, we observe a good agreement between the supermolecule (3.7 eV) or FDE results (4.2 eV) for the highest uranyl-dominated orbital. Moreover, this is close to the attributed vertical ED energy for the experimental uranyl-dominated spinors[71] at about 5 eV. In contrast, the corresponding uranyl ionization energy derived with the point-charge embedding of 6.2 eV, shown in Table 7, is significantly higher.

The IHFSCCSD results for the uranyl ionization energies in the FDE and point-charge models (**b**) and (**c**), also shown in Table 7, follow a similar trend. There is a strong overestimation in the point-charge model (7.78 eV) compared to FDE (5.81 eV), which gives values in good agreement with experiments. The many-electron ED states are essentially dominated by single determinants so that one can associate those to the ionizations of the individual Hartree–Fock spinor stabilized by correlation effects of the order of 1 eV. The effect of the completeness of the *Q* space is again not negligible and amounts to about 0.3 eV.

**3.4.2 Uranyl in $Cs_2UO_2Cl_4$.** Our results for the crystal model in which we allow the nearest $Cs^+$ ions to be polarized (model **f**) are also included in Table 7. Due to the effect of the Madelung potential, we observe a marked increase (about 5 eV) in the ionization energies for the uranyl electrons, compared to the corresponding isolated uranyl chloride models. Furthermore, we now have the three $e_g$ levels in the valence more stable than the $e_u$ ones for both CAM-B3LYP and IHFSCCSD.

The results for the first ionization energy are in good agreement with an experimental estimate[65] which places the binding energy of the valence uranyl electrons at about 9.4 eV. Compared to that value, CAM-B3LYP shows only a slight underestimation, while IHFSCCSD would appear to overestimate it by about 1 eV. However, in our calculations the polarization of the surroundings that would occur after the removal of an electron – which would provide a net stabilization of the final state – is not taken into account. Therefore, one can expect that once these are included, CAM-B3LYP would yield too low ionization potentials, in line with its known tendency to underestimate the ionization energies,[17] whereas IHFSCCSD would approach the experimental values.

## 4 Conclusions

We have investigated the electronic structure and spectra in the UV-Visible range of the uranyl cation ($UO_2^{2+}$) in $Cs_2UO_2Cl_4$, employing subsystem embedding approaches (DFT-in-DFT and WFT-in-DFT), in order to construct models of increasing sophistication for the crystal environment.

We have found that with the FDE formalism one can construct models in which the equatorial ligands to the uranyl species are represented in an approximate fashion as an embedding potential. The electronic spectra of such approximate models are able to capture, without significant loss of accuracy, the spectral features (spacing between states, symmetry classification) of the uranyl tetrachloride molecule for states that do not exhibit LMCT character, as well as its first ionization potential.

These models were further applied in calculations taking into account the crystal environment beyond the chloride ligands. As found in our prior investigation of $NpO_2^{2+}$ as an impurity in $Cs_2UO_2Cl_4$, at the experimental geometry we see rather small contributions due to the frozen crystal environment, which are larger for higher-lying states than for lower-lying ones. The relaxation of the electron density for atoms in the immediate vicinity of the central uranyl tetrachloride species accentuates this tendency, and turns out to be significant for describing the states' relative positions with respect to experiments. As the most significant environment effects are due to the presence of the equatorial ligands, the common practice in the literature, which consists of considering the isolated uranyl tetrachloride species, is indeed justified and a very good model for the spectrum in condensed phase.

The overall good performance of our embedded uranyl model makes us confident in applying such models to investigate the spectrum of uranyl in other condensed media and in the presence of different ligands. Nevertheless, these approximate models yield spectra which are on the whole red shifted. From our results and those available in the literature, we have concluded that these shifts can be attributed to a tendency of the approximate models to yield U–O equilibrium bond lengths which are shorter than the experimental ones. Thus, our calculations would in fact be sampling a region of the potential energy curves where the ground and excited states are starting to coalesce. We plan to investigate this issue further. We also plan to investigate the extent to which one can employ subsystem approaches to define minimalistic models for other classes of actinide-containing molecules (*e.g.* not containing the actinyl species, and where ligand–actinide interactions are more covalent than those investigated here) which can still yield accurate electronic spectra in the optical range.

In addition to an assessment of physical models, our results also provide further evidence for the applicability of the CAM-B3LYP functional to describe the electronic structure of actinyl-containing species, while at the same time underscoring the difficulty of all *ab initio* approaches employed so far to obtain very accurate energies for the low-lying electronic states of $Cs_2UO_2Cl_4$.

## Acknowledgements

This study was supported by the EC-supported ACTINET-i3 Integrated Infrastructure Initiative (JRP-C3-10, JRP-C5-08). ASPG, FR and

VV acknowledge support from PhLAM (Laboratoire de Physique des Lasers, Atomes et Molécules, Unité Mixte de Recherche de l'Université du CNRS), as well as the use of HPC resources from GENCI-CCRT (Grants 2012-081859 and 2013-081859). LV acknowledges the financial support from The Netherlands Organization for Scientific Research (NWO) *via* the Vici and NCF (computer time) programs. C.R.J. acknowledges support from the DFG-Center for Functional Nanostructures.

## References

1 F. J. C. Rossotti and H. Rossoti, *The Determination of Stability Constants*, McGraw-Hill, New York, 1961.
2 V. Eliet, I. Grenthe and G. Bidoglio, *J. Chem. Soc., Faraday Trans.*, 1995, **91**, 2275–2285.
3 P. Lindqvist-Reis, C. Walther, R. Klenze and N. M. Edelstein, *J. Phys. Chem. C*, 2009, **113**, 449–458.
4 G. K. Liu, *J. Phys. Chem. A*, 2011, **115**, 12419–12425.
5 G. K. Liu, *J. Phys. Chem. A*, 2012, **116**, 7443.
6 G. Schreckenbach and G. A. Shamov, *Acc. Chem. Res.*, 2010, **43**, 19.
7 T. Fleig, *Chem. Phys.*, 2011, **124**, 104106.
8 M. Dolg and X. Cao, *Chem. Rev.*, 2012, **112**, 403.
9 S. Matsika, Z. Zhang, S. R. Brozell, J.-P. Blaudeau, Q. Wang and R. M. Pitzer, *J. Phys. Chem. A*, 2001, **105**, 3825.
10 S. Matsika and R. M. Pitzer, *J. Phys. Chem. A*, 2001, **105**, 637–645.
11 T. Yang, R. Tyagi, Z. Zhang and R. M. Pitzer, *Mol. Phys.*, 2009, **107**, 1193.
12 Z. Zhang and R. M. Pitzer, *J. Phys. Chem. A*, 1999, **103**, 6880.
13 T. Fleig, H. J. A. Jensen, J. Olsen and L. Visscher, *J. Chem. Phys.*, 2006, **124**, 104106.
14 T. Privalov, P. Macak, B. Schimmelpfennig, E. Fromager, I. Grenthe and U. Wahlgren, *J. Am. Chem. Soc.*, 2004, **126**, 9801.
15 E. Fromager, F. Real, P. Wahlin, U. Wahlgren and H. J. A. Jensen, *J. Chem. Phys.*, 2009, **131**, 054107.
16 F. Ruipérez and U. Wahlgren, *J. Phys. Chem. A*, 2010, **114**, 3615–3621.
17 P. Tecmer, A. S. P. Gomes, U. Ekström and L. Visscher, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6249.
18 P. Tecmer, R. Bast, K. Ruud and L. Visscher, *J. Phys. Chem. A*, 2012, **116**, 7397–7404.
19 K. Pierloot and E. van Besien, *J. Chem. Phys.*, 2005, **123**, 204309.
20 K. Pierloot, E. van Besien, E. van Lenthe and E. J. Baerends, *J. Chem. Phys.*, 2007, **126**, 194311.
21 F. Réal, V. Vallet, C. Marian and U. Wahlgren, *J. Chem. Phys.*, 2007, **127**, 214302.
22 I. Infante, M. Vilkas, I. Ishikawa, U. Kaldor and L. Visscher, *J. Chem. Phys.*, 2007, **127**, 124308.
23 I. Infante and L. Visscher, *J. Chem. Phys.*, 2004, **121**, 5783.
24 I. Infante, L. Andrews, X. Wang and L. Gagliardi, *Chem.–Eur. J.*, 2010, **16**, 12804.
25 L. Andrews, X. Wang, B. Liang, F. Ruipérez, I. Infante, A. D. Raw and J. A. Ibers, *Eur. J. Inorg. Chem.*, 2011, 4457–4463.
26 I. Infante, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2006, **125**, 074301.
27 F. Réal, A. S. P. Gomes, L. Visscher, V. Vallet and E. Eliav, *J. Phys. Chem. A*, 2009, **113**, 12504.
28 I. Infante, A. Kovacs, G. L. Macchia, A. R. M. Shahi, J. K. Gibson and L. Gagliardi, *J. Phys. Chem. A*, 2010, **114**, 6007.
29 G. L. Macchia, I. Infante, J. Raab, J. K. Gibson and L. Gagliardi, *Phys. Chem. Chem. Phys.*, 2008, **48**, 7278.
30 R. Bast, H. J. A. Jensen and T. Saue, *Int. J. Quantum Chem.*, 2009, **109**, 2091.
31 R. G. Denning, *J. Phys. Chem. A*, 2007, **111**, 4125–4143.
32 G. Geipel, *Coord. Chem. Rev.*, 2006, **250**, 844–854.
33 L. S. Natrajan, *Coord. Chem. Rev.*, 2012, **256**, 1583–1603.
34 K. E. Knope and L. Soderholm, *Chem. Rev.*, 2012, **113**, 944–994.
35 C. Danilo, V. Vallet, J.-P. Flament and U. Wahlgren, *Phys. Chem. Chem. Phys.*, 2010, **12**, 1116–1130.
36 D. J. Tozer, *J. Chem. Phys.*, 2003, **119**, 12697.
37 O. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2004, **121**, 655.
38 S. Kümmel and L. Kronin, *Rev. Mod. Phys.*, 2008, **80**, 3.
39 A. M. Teale, F. D. Proft and D. J. Tozer, *J. Chem. Phys.*, 2008, **129**, 044110.
40 T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51–57.
41 A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C*, 2012, **108**, 222–277.
42 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050.
43 T. A. Wesolowski and J. Weber, *Chem. Phys. Lett.*, 1996, **248**, 71.
44 Ch. R. Jacob, J. Neugebauer and L. Visscher, *J. Comput. Chem.*, 2008, **29**, 1011–1018.
45 C. R. Jacob and L. Visscher, *J. Chem. Phys.*, 2008, **128**, 155102.
46 S. Laricchia, E. Fabiano and F. Della Sala, *J. Chem. Phys.*, 2013, **138**, 124112.
47 J. Neugebauer, Ch. R. Jacob, T. A. Wesolowski and E. J. Baerends, *J. Phys. Chem. A*, 2005, **109**, 7805–7814.
48 J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski, *J. Chem. Phys.*, 2005, **122**, 094115.
49 J. Neugebauer and E. J. Baerends, *J. Phys. Chem. A*, 2006, **110**, 8786–8796.
50 N. Govind, Y. A. Wang, A. J. R. da Silva and E. Carter, *Chem. Phys. Lett.*, 1998, **295**, 129.
51 N. Govind, Y. A. Wang and E. A. Carter, *J. Chem. Phys.*, 1999, **110**, 7677.
52 P. Huang and E. A. Carter, *J. Chem. Phys.*, 2006, **125**, 084102.
53 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353.
54 S. Höfener and L. Visscher, *J. Chem. Phys.*, 2012, **137**, 204120.
55 C. Daday, C. König, O. Valsson, J. Neugebauer and C. Filippi, *J. Chem. Theory Comput.*, 2013, **9**, 2355–2367.
56 J. Neugebauer, *J. Chem. Phys.*, 2007, **126**, 134116.
57 J. Neugebauer, *J. Chem. Phys.*, 2009, **131**, 084104.
58 J. Neugebauer, *J. Phys. Chem. B*, 2008, **112**, 2207–2217.

59 S. Höfener, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2012, **136**, 044104.
60 J. Neugebauer, *Phys. Rep.*, 2010, **489**, 1–87.
61 C. König and J. Neugebauer, *Phys. Chem. Chem. Phys.*, 2011, **13**, 10475.
62 C. König and J. Neugebauer, *ChemPhysChem*, 2012, **13**, 386–425.
63 P. Tecmer, H. van Lingen, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2012, **137**, 084308.
64 K. Takao, T. J. Bell and Y. Ikeda, *Inorg. Chem.*, 2013, **53**, 3459–3472.
65 R. G. Denning, *Struct. Bonding*, 1992, **79**, 215–276.
66 C. D. Flint and P. A. Tanner, *J. Chem. Soc., Faraday Trans. 2*, 1978, **74**, 2210–2217.
67 P. A. Tanner, C. S. K. Mak, Z. W. Pei, Y. L. Liu and L. Jun, *J. Phys.: Condens. Matter*, 2001, **13**, 189–194.
68 C. Görller-Walrand and L. G. Vanquickenborne, *J. Chem. Phys.*, 1973, **2**, 360.
69 C. Görller-Walrand, S. De Houwer, L. Fluyt and K. Binnemans, *Phys. Chem. Chem. Phys.*, 2004, **6**, 3292–3298.
70 K. Servaes, C. Hennig, R. VanDeun and C. Görller-Walrand, *Inorg. Chem.*, 2005, **44**, 7705–7707.
71 P. D. Dau, J. Su, H.-T. Liu, D.-L. Huang, J. Li and L.-S. Wang, *J. Chem. Phys.*, 2012, **137**, 064315.
72 E. van Besien, K. Pierloot and C. Görller-Walrand, *Phys. Chem. Chem. Phys.*, 2006, **8**, 4311–4319.
73 D. J. Watkin, R. G. Denning and K. Prout, *Acta Crystallogr.*, 1991, **C47**, 2517.
74 DIRAC, a relativistic *ab initio* electronic structure program, Release DIRAC11 (2011), written by R. Bast, H. J. Aa. Jensen, T. Saue and L. Visscher, with contributions from V. Bakken, K. G. Dyall, S. Dubillard, U. Ekström, E. Eliav, T. Enevoldsen, T. Fleig, O. Fossgaard, A. S. P. Gomes, T. Helgaker, J. K. Lærdahl, J. Henriksson, M. Iliaš, Ch. R. Jacob, S. Knecht, C. V. Larsen, H. S. Nataraj, P. Norman, G. Olejniczak, J. Olsen, J. K. Pedersen, M. Pernpointner, K. Ruud, P. Sałek, B. Schimmelpfennig, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, S. Villaume, O. Visser, T. Winther and S. Yamamoto (see http://dirac.chem.vu.nl).
75 K. G. Dyall, *Theor. Chem. Acc.*, 2007, **491**, 483.
76 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.
77 L. Visscher, *Theor. Chem. Acc.*, 1997, **98**, 68.
78 L. Visscher, E. Eliav and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 9720.
79 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2000, **113**, 9905.
80 A. Landau, E. Eliav, Y. Ishikawa and U. Kaldor, *J. Chem. Phys.*, 2001, **115**, 6862.
81 A. Landau and E. Eliav, *J. Chem. Phys.*, 2000, **115**, 6862.
82 ADF2012.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.
83 C. R. Jacob, S. M. Beyhan, R. E. Bulo, A. S. P. Gomes, A. W. Götz, K. Kiewisch, J. Sikkema and L. Visscher, *J. Comput. Chem.*, 2011, **32**, 2328.
84 E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1993, **99**, 4597.
85 E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1994, **101**, 9783.
86 E. van Lenthe and E. J. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142.
87 P. R. T. Schipper, O. V. Gritsenko, S. J. A. van Gisbergen and E. J. Baerends, *J. Chem. Phys.*, 2000, **112**, 1344–1352.
88 O. V. Gritsenko, P. R. T. Schipper and E. J. Baerends, *Chem. Phys. Lett.*, 1999, **302**, 199–207.
89 O. V. Gritsenko, P. R. T. Schipper and E. J. Baerends, *Int. J. Quantum Chem.*, 2000, **76**, 407–419.
90 J. P. Perdew, in *Electronic Structure of Solids*, ed. P. Ziesche and H. Eschrig, Akademie Verlag, Berlin, 1991, pp. 11–20.
91 A. Lembarki and H. Chermette, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1994, **50**, 5328.
92 A. Gelle and M.-B. Lepetit, *J. Chem. Phys.*, 2008, **128**, 244716.
93 V. V. Zhurov, E. A. Zhurova, A. I. Stash and A. A. Pinkerton, *J. Phys. Chem. A*, 2011, **115**, 13016–13023.
94 V. Vallet, U. Wahlgren and I. Grenthe, *J. Phys. Chem. A*, 2012, **116**, 12373–12380.
95 S. A. Kozimor, P. Yang, E. R. Batista, K. S. Boland, C. J. Burns, D. L. Clark, S. D. Conradson, R. L. Martin, M. P. Wilkerson and L. E. Wolfsberg, *J. Am. Chem. Soc.*, 2009, **131**, 12125–12136.
96 D. P. Chong, O. V. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2002, **117**, 1760.

## D.7 Paper XIV

PCCP

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c6cp08561j

# On the calculation of second-order magnetic properties using subsystem approaches in a relativistic framework†

Małgorzata Olejniczak,[a] Radovan Bast[b] and André Severo Pereira Gomes*[a]

Received 15th December 2016,
Accepted 16th February 2017

DOI: 10.1039/c6cp08561j

rsc.li/pccp

We report an implementation of nuclear magnetic resonance (NMR) shielding ($\sigma$), isotope-independent indirect spin–spin coupling ($K$) and the magnetizability ($\xi$) tensors in a frozen density embedding scheme using the four-component (4c) relativistic Dirac–Coulomb (DC) Hamiltonian and non-collinear spin density functional theory. The formalism takes into account the magnetic balance between the large and the small components of molecular spinors and assures the gauge-origin independence of the NMR shielding and magnetizability results. This implementation has been applied to hydrogen-bonded HXH$\cdots$OH$_2$ complexes (X = Se, Te, Po) and compared with supermolecular calculations and with an approach based on the integration of the magnetically induced current density vector. A comparison with the approximate zeroth-order regular approximation (ZORA) Hamiltonian indicates non-negligible differences in $\sigma$ and $K$ in the HPoH$\cdots$OH$_2$ complex, and calls for a thorough comparison of ZORA and DC Hamiltonians in the description of environment effects on NMR parameters for molecular systems with heavy elements.



## 1 Introduction

The response to magnetic fields can be of great help in investigating molecular systems in complex environments. This is perhaps best illustrated by the widespread use of experimental techniques such as NMR spectroscopy to characterize compounds in a condensed phase, including disordered and amorphous solids. The great sensitivity of this technique to any changes in the electronic structure in the vicinity of a probed nucleus, triggered for instance by specific inter- and intramolecular interactions (*e.g.* hydrogen bonds) found in complex biological systems (*e.g.* proteins, peptides, amino acids),[1,2] catalysts,[3] paramagnetic systems[4,5] and radioactive compounds containing actinide atoms[6,7] is often unrivaled by other experimental techniques.

Due to the complexity of such systems, theoretical tools can be of extreme importance to aid experimentalists in interpreting their results. Most theoretical approaches hinge on the recognition that, in the weak magnetic field regime under which most experiments are carried out, the magnetic field can be treated as a perturbation, the energy of a molecule can be Taylor-expanded in terms of perturbation strengths and the effect of a perturbation on a given system can be evaluated *via* response theory.[8]

For closed-shell systems, a static magnetic field induces only even-order changes in the total energy:[9]

$$E(\varepsilon) = E_0 + \frac{1}{2}\frac{\mathrm{d}^2 E}{\mathrm{d}\varepsilon_1 \mathrm{d}\varepsilon_2}\varepsilon_1\varepsilon_2 + \frac{1}{4!}\frac{\mathrm{d}^4 E}{\mathrm{d}\varepsilon_1 \ldots \mathrm{d}\varepsilon_4}\varepsilon_1 \ldots \varepsilon_4 + \cdots, \quad (1)$$

where $E_0$ denotes the energy at zero field and $\{\varepsilon_n\}$ are the field strengths of the applied perturbations collected in vector $\varepsilon$. The coefficients of this expansion, taken in the zero-field limit, define molecular properties in the Born–Oppenheimer approximation. In this paper, we focus on three second-order magnetic properties arising from a perturbation of an external field $\vec{B}$ or the field of nuclear magnetic dipole moments, $\{\vec{m}_A\}$: the NMR shielding tensor of a nucleus $K$,

$$\sigma^K_{\alpha\beta} = \left.\frac{\mathrm{d}^2 E}{\mathrm{d}B_\alpha \mathrm{d}m_{K;\beta}}\right|_{\vec{B},\{\vec{m}_A\}=0}, \quad (2)$$

the reduced spin–spin coupling tensor of nuclei $K$ and $L$,

$$K^{KL}_{\alpha\beta} = \left.\frac{\mathrm{d}^2 E}{\mathrm{d}m_{K;\alpha}\mathrm{d}m_{L;\beta}}\right|_{\{\vec{m}_A\}=0}, \quad (3)$$

related to the indirect spin–spin coupling constants observed in the NMR experiment, $J^{KL} = (\hbar/2\pi)\gamma_K\gamma_L K^{KL}_{2\beta}$, with $\gamma_M$ denoting the

[a] Université de Lille, CNRS, UMR 8523 – PhLAM – Physique des Lasers, Atomes et Molécules, F-59000 Lille, France. E-mail: gosia.olejniczak@univ-lille1.fr, andre.gomes@univ-lille1.fr; Fax: +33-3-2033-7020; Tel: +33-3-2043-4163
[b] High Performance Computing Group, UiT The Arctic University of Norway, N-9037 Tromsø, Norway. E-mail: radovan.bast@uit.no; Tel: +47-776-44117

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp08561j (also available at http://dx.doi.org/10.5281/zenodo.291373)

gyromagnetic ratio of a given isotope of $M$, and the molecular magnetizability tensor,

$$\xi_{\alpha\beta} = -\frac{\mathrm{d}^2 E}{\mathrm{d}B_\alpha \mathrm{d}B_\beta}\bigg|_{\vec{B}=0}, \quad (4)$$

which, unlike the first two, is not a local but rather an extensive property as it does not depend on the magnetic moment of a given nucleus – though it may play an important role in NMR spectroscopy by inducing changes in the local magnetic fields which, in turn, will affect chemical shifts.[10–12]

NMR properties can nowadays be routinely calculated using density functional theory (DFT) for relatively large closed-shell systems and using methods based on wave function theory (WFT) for much smaller systems.[9,13,14] Calculating the magnetizability tensor is computationally more demanding due to the slow convergence with the basis set size. This can be partly alleviated by the use of London atomic orbitals (LAOs),[9,15] which also remedy the gauge-origin dependence problem for properties arising from an external magnetic field perturbation ($\xi$, $\sigma^K$), at the cost of more complex derivations and more involved implementation.

Furthermore, it is now recognized that relativistic effects[16,17] can be appreciable for magnetizabilities[18,19] and are essential for a proper description of NMR properties, including for the light elements neighbouring the heavy one(s) due to the so-called "heavy-atom on the light atom" (HALA) effect.[20] NMR properties obtained using approximate Hamiltonians such as the quasirelativistic two-component (2c) ZORA Hamiltonian can differ significantly from those obtained with a more rigorous treatment afforded by four-component approaches as in the Dirac–Coulomb Hamiltonian, even for relatively light elements such as those in the fourth row of the periodic table. The conventional wisdom has been that approximate 2c approaches, which are computationally much cheaper than the 4c ones, can be nevertheless reliable for relative quantities such as chemical shifts, due to error cancellation. Recent studies[21,22] paint a more nuanced picture, and seem to indicate that there may be significant differences between Hamiltonians for relative quantities as well, notably for heavier elements.

One must also take into account the effect of the surroundings on the molecular properties, something preferably done by embedding approaches due to the steep increase in computation cost for the explicit inclusion of *e.g.* solvent molecules. As methods representing the environment in an implicit manner such as PCM[23] or COSMO[24–26] have difficulty describing specific interactions such as hydrogen bonds, one is better served by approaches such as frozen-density embedding (FDE),[27–29] in which the total system is partitioned into subsystems whose interaction is calculated with DFT, while the subsystems themselves can be treated with DFT (DFT-in-DFT embedding) or WFT (WFT-in-DFT embedding). FDE has been applied to the calculation of molecular properties arising from electric perturbations,[30–34] and in particular employing 4c Hamiltonians,[35,36] though there have been only a few studies of magnetic properties: NMR shieldings[37,38] and indirect spin–spin couplings,[39] the latter using the ZORA Hamiltonian.

The aim of this paper is therefore to bridge this gap and propose an FDE implementation that is capable of treating general second-order magnetic properties with the 4c Dirac–Coulomb (DC) Hamiltonian, by extending the general framework for response properties[32] in line with the 4c DFT simple magnetic balance (sMB) framework.[40] We also investigate the real-space determination of NMR shielding *via* the integration of magnetically induced currents and its use for understanding the effect of approximations introduced in practical DFT-in-DFT calculations. This will allow for investigating the suitability of approximate 2c approaches for the calculation of chemical shifts and provide a way to incorporate environment effects in the determination of shielding scales and the nuclear magnetic dipole moments,[41] a field which has received renewed interest in recent years.[42–45]

The paper is organized as follows: in the next section we present an overview of the theoretical formulation, followed by the presentation of proof-of-principle calculations on the $H_2X$–$H_2O$ (X = Se, Te, Po) family of compounds, where we also compare the description of environment effects with 2c and 4c Hamiltonians for NMR shieldings and indirect spin–spin couplings. Such a comparison of magnetizabilities is currently not possible as ours is, to the best of our knowledge, the first FDE implementation.

## 2 Theory

We begin by briefly summarizing the theory for NMR shieldings,[40,46,47] magnetizability[48] and NMR spin–spin couplings[46,49,50] in closed-shell molecules using the 4c relativistic DC Hamiltonian and mean-field methods and its implementation in the DIRAC[51] software, followed by the general FDE framework for molecular properties[32] and its extension to magnetic properties in the relativistic framework.

Throughout the text, $i$, $j$,... denote occupied molecular orbitals, $a$, $b$,... virtual orbitals and $p$, $q$,... orbitals in general. Greek indices are used for the three Cartesian components and Latin indices are used for the components of the four-component vector. The summation over repeated indices is assumed. The SI-based atomic units are employed ($\hbar = m_e = e = 1/(4\pi\varepsilon_0) = 1$).[52] As we restrict ourselves to closed-shell systems represented by a single Slater determinant, we employ the following parametrization in the second quantization for the unperturbed wavefunctions:

$$|\tilde{0}\rangle = \exp(-\hat{\kappa})|0\rangle; \quad \hat{\kappa} = \kappa_{ai}\hat{a}^\dagger\hat{i} - \kappa_{ai}^*\hat{i}^\dagger\hat{a}, \quad (5)$$

where $\hat{\kappa}$ is an anti-Hermitian operator represented by a matrix of orbital rotation amplitudes, which serve as variational parameters in the optimization of the ground state and its response to an external perturbation.

### 2.1 Molecules in magnetic fields in a 4c framework

**Molecular Hamiltonian.** Starting with the generic form of the Hamiltonian in the Born–Oppenheimer approximation,

$$\hat{H} = \sum_i \hat{h}_i + \sum_{i<j} \hat{g}_{ij} + V_{\mathrm{NN}}, \quad (6)$$

with $V_{\mathrm{NN}}$ denoting a classical repulsion potential of clamped nuclei, it is the choice of one- and two-electron operators – $\hat{h}_i$

and $\hat{g}_{ij}$, that determines whether and which relativistic effects are included.[53] The one-electron part corresponds to the Dirac operator, which in the presence of a uniform external magnetic field ($\vec{B}$) and the magnetic dipole moments of nuclei ($\{\vec{m}_K\}$) reads:

$$\begin{aligned}\hat{h} &= \hat{h}_0 + \vec{B}\cdot\hat{h}_{\mathrm{B}} + \sum_K \vec{m}_K\cdot\hat{h}_{m_K},\\ \hat{h}_0 &= \beta' mc^2 + c(\vec{\alpha}\cdot\vec{p}) + v_{\mathrm{nuc}},\end{aligned} \tag{7}$$

where $\alpha$ and $\beta' = \beta - 1_{4\times4}$ are $4 \times 4$ Dirac matrices in their standard representation, and $c$ is the speed of light.[54] The Zeeman operator, $\hat{h}_{\mathrm{B}}$, and the hyperfine operators, $\hat{h}_{m_K}$, are defined as:

$$\hat{h}_{\mathrm{B}} = \frac{1}{2}(\vec{r}_{\mathrm{G}}\times c\vec{\alpha}), \quad \hat{h}_{m_K} = \frac{1}{c^2}\frac{\vec{r}_K\times c\vec{\alpha}}{{r_K}^3}, \tag{8}$$

where $\vec{r}_{\mathrm{X}} = \vec{r} - \vec{R}_{\mathrm{X}}$ with an arbitrary gauge origin $\vec{R}_{\mathrm{G}}$ and the center of nucleus $K$ in $\vec{R}_K$. The two-electron part of the 4c DC Hamiltonian is restricted to the Coulomb potential, which in the relativistic regime also includes spin-same-orbit interaction.[53]

**One-electron basis.** The 4c molecular spinors, eigenfunctions of the DC Hamiltonian, are expanded in scalar finite basis sets.[54] The small component basis functions ($\chi^{\mathrm{S}}$) are generated from the large component functions ($\chi^{\mathrm{L}}$) by the restricted kinetic balance (RKB) or the restricted magnetic balance (RMB) prescription, which properly describes the relation between large and small components in the presence of external magnetic fields. The question of magnetic balance in 4c calculations was initially addressed by Aucar and coworkers[46] and Kutzelnigg,[55,56] though later investigations have shown that these yielded mixed results.[57] Later, Komorovsky and coworkers,[58–60] followed by Cheng[61] and Reynolds,[62] revisited the question formally – and computationally – establishing RMB.

For properties explicitly dependent on an external magnetic field ($\sigma$, $\xi$), the basis functions are replaced by London atomic orbitals,

$$\omega_\mu^K(\vec{r}) = \exp\left\{-\frac{\mathrm{i}}{2}\vec{B}\times\left(\vec{R}_K - \vec{R}_{\mathrm{G}}\right)\cdot\vec{r}\right\}\chi_\mu^K(\vec{r}), \tag{9}$$

which guarantee the gauge-origin invariance of results in a finite basis approximation. LAOs are also appealing when used in combination with RMB as they make the magnetic balance atomic and easy to handle by the simple scheme (sMB).[40] The orthogonality of molecular orbitals at all field strengths is ensured by connection matrices, $T$,[63] which couple the unmodified molecular orbitals (UMOs) to the orthonormalized set (OMOs, $\{\breve{\psi}\}$),

$$\psi_q^{\mathrm{UMO}}(\vec{B}) = \omega_\mu(\vec{B})c_{\mu q}(0), \quad \breve{\psi}_p(\vec{B}) = \psi_q^{\mathrm{UMO}}(\vec{B})T_{qp}(\vec{B}), \tag{10}$$

yet for the price of more complex equations, as the wave function is now dependent on a perturbation through $\omega_\mu(\vec{B})$ and $T(\vec{B})$ in addition to $\kappa_{pq}(\vec{B})$.

**Spin density functional theory.** The choice of spin-density functional theory (SDFT)[64–66] is a compromise between a desirable[67–69] yet so far unattainable[70–73] DFT formalism for molecules in magnetic fields involving current density, and the conventional charge-density-only approaches whose density functional approximations (DFAs) are developed to reproduce energy in the absence of magnetic perturbations.[74,75] Also, due to the complexity of the relativistic generalization of the DFT method,[65,76,77] non-relativistic functionals are used with relativistic densities.

In this work, the non-collinear SDFT is employed, with the spin density (calculated as a norm of the spin magnetization vector) and the charge density as basic variables, expressed together as a general density component:[40,58,78]

$$\rho_k(\vec{r},\vec{B}) = \breve{\Omega}_{pq;k}(\vec{r},\vec{B})\tilde{D}_{pq}(\kappa) \quad k\in\{0,x,y,z\}, \tag{11}$$

with the elements of the density matrix, $\tilde{D}_{pq} = \langle\tilde{0}|p^\dagger q|\tilde{0}\rangle$ and the generalized overlap distribution, $\breve{\Omega}$, expressed in the OMO basis whenever LAOs are used:

$$\breve{\Omega}_{pq;k} = \breve{\psi}_p^\dagger\Sigma_k\breve{\psi}_q; \quad \Sigma_0 = I_{4\times4}, \quad \Sigma_\mu = \begin{bmatrix}\sigma_\mu & 0_{2\times2}\\ 0_{2\times2} & \sigma_\mu\end{bmatrix}. \tag{12}$$

The ground state energy and the optimal $\rho_k$ are obtained by minimization of the energy functional, $E[\rho_k]$, which can be written in a Kohn–Sham (KS) manner as a sum of five terms:

$$\begin{aligned}E[\rho_k] = {}& T_{\mathrm{s}}[\rho_k] + E_{\mathrm{xc}}[\rho_k] + V_{\mathrm{NN}}\\ &+\int\left(\rho_0 v_{\mathrm{nuc}} + \sum_{\mu=x,y,z}\rho_\mu\cdot B_\mu\right)\mathrm{d}\vec{r} + \frac{1}{2}\iint\frac{\rho_0(\vec{r}_1)\rho_0(\vec{r}_2)}{|\vec{r}_1-\vec{r}_2|}\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2,\end{aligned} \tag{13}$$

where $T_{\mathrm{s}}$ is the kinetic energy of non-interacting electrons, $E_{\mathrm{xc}}$ is the exchange–correlation (xc) contribution and the two last terms describe the interaction of electrons with an electromagnetic potential and with other electrons, respectively. The minimization of eqn (13) with respect to $\rho_k$ (in the zero magnetic-field limit) yields the 4c KS equations for the DC Hamiltonian:

$$(\beta' mc^2 + c(\vec{\alpha}\cdot\vec{p}) + v_{\mathrm{KS}}[\rho_k])\psi = \varepsilon\psi \tag{14}$$

with the effective KS potential:

$$v_{\mathrm{KS}}[\rho_k] = -\left(v_{\mathrm{nuc}} + \int\frac{\rho_0(\vec{r}')}{|\vec{r}-\vec{r}'|}\mathrm{d}\vec{r}' + \frac{\delta E_{\mathrm{xc}}}{\delta\rho_k}\right). \tag{15}$$

**Linear response (LR) at the 4c SDFT level.** Considering now the case of time-independent perturbations with strengths $\varepsilon_1$ and $\varepsilon_2$, the second-order molecular property can be written as:

$$\left.\frac{\mathrm{d}^2E}{\mathrm{d}\varepsilon_1\mathrm{d}\varepsilon_2}\right|_{\varepsilon=0} = \left.\frac{\partial^2E}{\partial\kappa_{pq}\partial\varepsilon_2}\frac{\partial\kappa_{pq}}{\partial\varepsilon_1}\right|_{\varepsilon=0} + \left.\frac{\partial^2E}{\partial\varepsilon_1\partial\varepsilon_2}\right|_{\varepsilon=0}, \tag{16}$$

assuming that the energy is optimized with respect to variational parameters at all field strengths, $\partial E/\partial\kappa = 0$. The first contribution is determined perturbatively, with the first-order orbital rotation amplitudes, $\partial\kappa/\partial\varepsilon$, obtained from the LR equations:

$$0 = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon_1}\left(\frac{\partial E}{\partial\kappa_{pq}}\right)\right|_{\varepsilon=0} = \left.\left(\frac{\partial^2E}{\partial\kappa_{pq}\partial\varepsilon_1} + \frac{\partial^2E}{\partial\kappa_{pq}\partial\kappa_{\mathrm{rs}}}\frac{\partial\kappa_{\mathrm{rs}}}{\partial\varepsilon_1}\right)\right|_{\varepsilon=0}, \tag{17}$$

which can be recast in a matrix form as:[79]

$$0 = \mathbf{E}_{\varepsilon_1}^{[1]} + \mathbf{E}_0^{[2]}\mathbf{X}_{\varepsilon_1}. \tag{18}$$

Here, $\mathbf{E}_0^{[2]}$ is the electronic Hessian, $\mathbf{E}_{\varepsilon_1}^{[1]}$ the property gradient and $\mathbf{X}_{\varepsilon_1}$ the solution vector yielding $\{\kappa_{\mathrm{rs}}^{\varepsilon_1}\}$. While the Hessian is independent of a perturbation, the property gradient is calculated as the first-order derivative of the KS matrix with respect to the field strength of the applied perturbation,

$$\mathbf{E}_{\varepsilon_1}^{[1]} = \begin{bmatrix} g^{\varepsilon_1} \\ g^{*\varepsilon_1} \end{bmatrix}, \quad g_{ai}^{\varepsilon_1} = \frac{\partial E_{\varepsilon_1}}{\partial \kappa_{ai}^*}\bigg|_0 = \left\langle 0 \middle| \left[-\hat{a}_i^\dagger \hat{a}_a, \hat{h}_{\varepsilon_1}\right] \middle| 0 \right\rangle = -\tilde{F}_{ai}^{\varepsilon_1}. \tag{19}$$

In particular, if $\varepsilon_1 = \vec{B}$, the property gradient is calculated in the OMO basis and requires additional contributions involving derivatives of LAOs and of matrices $T$.[40,47] Once $\mathbf{X}_{\varepsilon_1}$ has been determined, one can construct the static linear response function:

$$\langle\langle \varepsilon_1; \varepsilon_2 \rangle\rangle_0 = E_{\varepsilon_1}^{[1]\dagger} X_{\varepsilon_2} = -E_{\varepsilon_1}^{[1]\dagger} \left(E_0^{[2]}\right)^{-1} E_{\varepsilon_2}^{[1]}, \tag{20}$$

which constitutes the response contribution to the molecular property expressed by the first term of eqn (16). The second term of eqn (16) can be thought of as an expectation value, which due to the linearity of the DC Hamiltonian in the applied perturbations (eqn (7)) is non-zero only in perturbation-dependent basis sets.

This brings about the final form of the properties of interest in this paper:

$$K_{\alpha\beta}^{KL} = \langle\langle m_{K;\alpha}; m_{L;\beta} \rangle\rangle_0 \tag{21}$$

$$\sigma_{\alpha\beta}^{K} = \langle\langle m_{K;\alpha}; B_\beta \rangle\rangle_0 \tag{22}$$

$$\xi_{\alpha\beta} = -\left( \langle\langle B_\alpha; B_\beta \rangle\rangle_0 + \frac{\partial^2 E}{\partial B_\alpha \partial B_\beta}\bigg|_0 \right) \tag{23}$$

with the LAO basis used for the last two.

### 2.2 Frozen density embedding

In FDE, the total system is partitioned into interacting subsystems (for simplicity here we shall consider only two, the one of interest ($I$) and the other ($II$) representing the environment) implying a partition of the total density and energy.[27] We can further consider the case of spin-density FDE[80,81] and partition the generalized density component and the energy,

$$\rho_k^{\mathrm{tot}}(\vec{r}) = \rho_k^I(\vec{r}) + \rho_k^{II}(\vec{r}) \tag{24}$$

$$E_{\mathrm{tot}}[\rho_k^{\mathrm{tot}}] = E_I[\rho_k^I] + E_{II}[\rho_k^{II}] + E_{\mathrm{int}}[\rho_k^I, \rho_k^{II}], \tag{25}$$

where $E_{\mathrm{M}}[\rho_k^{\mathcal{M}}]$ is the energy of an isolated subsystem ($\mathcal{M} = I, II$) calculated from eqn (13) and $E_{\mathrm{int}}$ is the interaction energy dependent on the densities of both subsystems,

$$\begin{aligned} E_{\mathrm{int}}\left[\rho_k^I, \rho_k^{II}\right] &= E_{\mathrm{tot}}\left[\rho_k^{\mathrm{tot}}\right] - E_I\left[\rho_k^I\right] - E_{II}\left[\rho_k^{\mathrm{II}}\right] \\ &= \int \left[\rho_0^I(\vec{r}) v_{\mathrm{nuc}}^{II}(\vec{r}) + \rho_0^{II}(\vec{r}) v_{\mathrm{nuc}}^I(\vec{r})\right] \mathrm{d}\vec{r} + E_{\mathrm{nuc}}^{I,II} \\ &\quad + \iint \frac{\rho_0^I(\vec{r}_1)\rho_0^{II}(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} \mathrm{d}\vec{r}_1 \mathrm{d}\vec{r}_2 + E_{\mathrm{xc}}^{\mathrm{nadd}} + T_{\mathrm{s}}^{\mathrm{nadd}}. \end{aligned} \tag{26}$$

where $E_{\mathrm{nuc}}^{I,II}$ is the nuclear repulsion energy between subsystems, and $E_{\mathrm{xc}}^{\mathrm{nadd}}$ and $T_{\mathrm{s}}^{\mathrm{nadd}}$ are the non-additive contributions defined as:[32]

$$X^{\mathrm{nadd}} \equiv X^{\mathrm{nadd}}[\rho_k^I, \rho_k^{II}] = X[\rho_k^{\mathrm{tot}}] - X[\rho_k^I] - X[\rho_k^{II}]. \tag{27}$$

In order to determine $\rho_k^I$ in the presence of other subsystem(s) with a given generalized density $\rho_k^{II}$ one solves the 4c KS equations for a constrained electron density (KSCED)[82] which, in the limit of zero magnetic field have the form

$$(\beta' mc^2 + c(\vec{\alpha}\cdot\vec{p}) + v_{\mathrm{KS}}[\rho_k^I] + v_{\mathrm{emb}}^I[\rho_k^I, \rho_k^{II}])\psi^I(\vec{r}) = \varepsilon^I \psi^I(\vec{r}), \tag{28}$$

where an effective KS potential of eqn (15) is augmented by the embedding potential,

$$v_{\mathrm{emb};k}^I(\vec{r}) = \frac{\delta E_{\mathrm{int}}}{\delta \rho_k^I(\vec{r})} = \frac{\delta E_{\mathrm{xc}}^{\mathrm{nadd}}}{\delta \rho_k^I} + \frac{\delta T_{\mathrm{s}}^{\mathrm{nadd}}}{\delta \rho_k^I} + v_{\mathrm{nuc}}^{II}(\vec{r}) + \int \frac{\rho_0^{II}(\vec{r}')}{|\vec{r} - \vec{r}'|} \mathrm{d}\vec{r}', \tag{29}$$

representing the interaction of subsystem $I$ with other subsystem(s). One can also relax the constraints on $\rho_k^{II}$ by interchanging it with $\rho_k^I$ and solving the analogous KSCED equations in an iterative manner in the so-called freeze–thaw[83] procedure.

FDE is formally exact in the limit of exact functionals describing the non-additive exchange–correlation and kinetic energies, but for computational efficiency both are generally obtained using approximate density functionals and grouped into a single term,

$$E_{\mathrm{xck}}^{\mathrm{nadd}}[\rho_k^I, \rho_k^{II}] = E_{\mathrm{xc}}^{\mathrm{nadd}}[\rho_k^I, \rho_k^{II}] + T_{\mathrm{s}}^{\mathrm{nadd}}[\rho_k^I, \rho_k^{II}]. \tag{30}$$

It should be noted that the currently available kinetic energy functionals have a limited accuracy,[27,28] and while sufficient for relatively weak interactions (*e.g.* hydrogen bonds),[84] practical difficulties may emerge for stronger ones requiring the kinetic energy density functionals to be replaced by other approaches.[85,86]

**2.2.1 FDE molecular properties.** In what follows we shall discuss the contributions to second-order molecular properties presented in eqn (16) in a subsystem manner. We use separate sets of externally orthogonal orbitals for subsystems $I$ and $II$,[32,87] implying separate sets of orbital rotation coefficients, $\kappa_{p_M q_N} = \delta_{MN}\kappa_{p_M q_M}$ for $M, N \in \{I, II\}$, and the parametrization of the total density (eqn (24)):

$$\rho_k^{\mathrm{tot}}(\vec{r}, \kappa^I, \kappa^{II}) = \rho_k^I(\vec{r}, \kappa^I) + \rho_k^{II}(\vec{r}, \kappa^{II}), \tag{31}$$

with $\kappa^{\mathcal{M}} = \{\kappa_{p_M q_M}\}$ for $M \in \{I, II\}$.

*2.2.1.1 Linear response functions.* The electronic Hessian and property gradient are now subdivided into isolated subsystem and interaction contributions[32]

$$\mathbf{E}_0^{[2]} = \begin{bmatrix} \mathbf{E}_0^{[2];M,M} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_0^{[2];N,N} \end{bmatrix} + \begin{bmatrix} \mathbf{E}_{0,\mathrm{int}}^{[2];M,M} & \mathbf{E}_{0;\mathrm{int}}^{[2];M,N} \\ \mathbf{E}_{0,\mathrm{int}}^{[2];N,M} & \mathbf{E}_{0;\mathrm{int}}^{[2];N,N} \end{bmatrix}, \tag{32}$$

$$\mathbf{E}_{\varepsilon_1}^{[1]} = \begin{bmatrix} \mathbf{E}_{\varepsilon_1}^{[1];M} & \mathbf{E}_{\varepsilon_1}^{[1];N} \end{bmatrix}^\dagger + \begin{bmatrix} \mathbf{E}_{\varepsilon_1;\mathrm{int}}^{[1];M} & \mathbf{E}_{\varepsilon_1;\mathrm{int}}^{[1];N} \end{bmatrix}^\dagger, \tag{33}$$

with $M \neq N$, which leads to a system of LR equations

$$\left(\mathbf{E}_0^{[2];M,M} + \mathbf{E}_{0;\mathrm{int}}^{[2];M,M}\right)\mathbf{X}_{\varepsilon_1}^{M} + \mathbf{E}_{0;\mathrm{int}}^{[2];M,N}\mathbf{X}_{\varepsilon_1}^{N} = -\left(\mathbf{E}_{\varepsilon_1}^{[1];M} + \mathbf{E}_{\varepsilon_1;\mathrm{int}}^{[1];M}\right), \tag{34}$$

$$\mathbf{E}_{0;\mathrm{int}}^{[2];N,M}\mathbf{X}_{\varepsilon_1}^{M} + \left(\mathbf{E}_0^{[2];N,N} + \mathbf{E}_{0;\mathrm{int}}^{[2];N,N}\right)\mathbf{X}_{\varepsilon_1}^{N} = -\left(\mathbf{E}_{\varepsilon_1}^{[1];N} + \mathbf{E}_{\varepsilon_1;\mathrm{int}}^{[1];N}\right), \tag{35}$$

where the response vector has also been split into blocks pertaining to each subsystem, $\mathbf{X}_{\varepsilon_1} = \left[\mathbf{X}_{\varepsilon_1}^{M}\mathbf{X}_{\varepsilon_1}^{N}\right]^{\dagger}$. The matrix elements of each sub-block have the form

$$\mathbf{E}_0^{[2];M,M} = \frac{\partial^2 E_M}{\partial\kappa_{pq}^{M}\partial\kappa_{\mathrm{rs}}^{M}}; \quad \mathbf{E}_{0;\mathrm{int}}^{[2];M,N} = \frac{\partial^2 E_{\mathrm{int}}}{\partial\kappa_{pq}^{M}\partial\kappa_{\mathrm{rs}}^{N}} \tag{36}$$

for the Hessian and

$$\mathbf{E}_{\varepsilon_1}^{[1];M} = \frac{\partial^2 E_M}{\partial\kappa_{ai}^{M}\partial\varepsilon_1}; \quad \mathbf{E}_{\varepsilon_1;\mathrm{int}}^{[1];M} = \frac{\partial^2 E_{\mathrm{int}}}{\partial\kappa_{ai}^{M}\partial\varepsilon_1} \tag{37}$$

for the property gradient.

While the subsystem contributions to the electronic Hessian and property gradient are the same as in (S)DFT, the interaction contributions are calculated from the chain rule, employing the parametrization of eqn (31), which allows for a straightforward separation of the contributions from perturbed densities of both subsystems ($M, N \in \{I, II\}$):

$$\left.\frac{\partial^2 E_{\mathrm{int}}}{\partial\kappa_{pq}^{M}\partial\kappa_{\mathrm{rs}}^{N}}\right|_0 = \delta_{MN}\int\left.\frac{\delta E_{\mathrm{int}}}{\delta\rho_k^{M}}\frac{\partial^2\rho_k^{M}}{\partial\kappa_{pq}^{M}\partial\kappa_{\mathrm{rs}}^{N}}\right|_0 + \iint\left.\frac{\delta^2 E_{\mathrm{int}}}{\delta\rho_k^{M}\delta\rho_{k'}^{N}}\frac{\partial\rho_k^{M}}{\partial\kappa_{pq}^{M}}\right|_0\left.\frac{\partial\rho_{k'}^{N}}{\partial\kappa_{\mathrm{rs}}^{N}}\right|_0 \tag{38}$$

$$\left.\frac{\partial^2 E_{\mathrm{int}}}{\partial\kappa_{pq}^{M}\partial\varepsilon_1}\right|_0 = \int\left.\frac{\delta E_{\mathrm{int}}}{\delta\rho_k^{M}}\frac{\partial^2\rho_k^{M}}{\partial\kappa_{pq}^{M}\partial\varepsilon_1}\right|_0 + \iint\left.\frac{\delta^2 E_{\mathrm{int}}}{\delta\rho_k^{M}\delta\rho_{k'}^{N}}\frac{\partial\rho_k^{M}}{\partial\kappa_{pq}^{M}}\right|_0\left.\frac{\partial\rho_{k'}^{N}}{\partial\varepsilon_1}\right|_0 \tag{39}$$

The functional derivatives of the interaction energy with respect to the general density components are the embedding potential $v_{\mathrm{emb};k}^{M}$ of eqn (29) and the embedding kernel ($M, N \in \{I, II\}$):

$$w_{\mathrm{emb};k,k'}^{M,N}(\vec{r}_1,\vec{r}_2) = \frac{\delta^2 E_{\mathrm{int}}}{\delta\rho_k^{M}(\vec{r}_1)\delta\rho_{k'}^{N}(\vec{r}_2)} = (1-\delta_{MN})\frac{1}{|\vec{r}_1-\vec{r}_2|} + \frac{\delta^2 E_{\mathrm{xck}}}{\delta\rho_k^{\mathrm{tot}}(\vec{r}_1)\delta\rho_{k'}^{\mathrm{tot}}(\vec{r}_2)} - \delta_{MN}\frac{\delta^2 E_{\mathrm{xck}}}{\delta\rho_k^{M}(\vec{r}_1)\delta\rho_{k'}^{M}(\vec{r}_2)} \tag{40}$$

We recall that in both cases the functional derivatives are calculated with the ground-state densities.

*Electronic Hessian.* The interaction energy contributions to the electronic Hessian (eqn (38)) can be further rewritten as ($M \neq N$):

$$\left.\frac{\partial^2 E_{\mathrm{int}}}{\partial\kappa_{pq}^{M}\partial\kappa_{\mathrm{rs}}^{N}}\right|_0 = \delta_{MN}\int v_{\mathrm{emb};k}^{M}\left.\frac{\partial^2\rho_k^{M}}{\partial\kappa_{pq}^{M}\partial\kappa_{\mathrm{rs}}^{N}}\right|_0 + \iint w_{\mathrm{emb};k,k'}^{M,M}\left.\frac{\partial\rho_k^{M}}{\partial\kappa_{pq}^{M}}\right|_0\left.\frac{\partial\rho_{k'}^{M}}{\partial\kappa_{\mathrm{rs}}^{M}}\right|_0 \tag{41}$$

$$+\iint w_{\mathrm{emb};k,k'}^{M,N}\left.\frac{\partial\rho_k^{M}}{\partial\kappa_{pq}^{M}}\right|_0\left.\frac{\partial\rho_{k'}^{N}}{\partial\kappa_{\mathrm{rs}}^{N}}\right|_0, \tag{42}$$

discerning the embedding potential as well as the uncoupled and coupled embedding kernel terms. In the current DIRAC implementation[32] only the terms from eqn (41) are included, so that the coupling contributions of eqn (42) are neglected. As it is usually the case, the Hessian is not explicitly constructed but rather its eigenvectors and eigenvalues are obtained by iterative approaches.[79]

*Property gradient.* When there is no dependence of $\rho_k$ on the perturbation, eqn (39) is identically zero and the property gradient contains only contributions from the isolated subsystems. As the terms of eqn (39) are non-zero when LAOs are used, from now on they will be referred to as FDE-LAO contributions to the property gradient. Eqn (39) can be rewritten in a more explicit form ($M \neq N$),

$$\left.\frac{\partial}{\partial B_\alpha}\frac{\partial E_{\mathrm{int}}}{\partial\kappa_{ai}^{M}}\right|_0 = -\int v_{\mathrm{emb};k}^{M}(\vec{r})\breve{\Omega}_{ia;k}^{B_\alpha;M}\mathrm{d}\vec{r} \tag{43}$$

$$-\iint w_{\mathrm{emb};k,k'}^{M,M}(\vec{r}_1,\vec{r}_2)\Omega_{ia;k}^{M}(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\alpha;M}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 \tag{44}$$

$$-\iint w_{\mathrm{emb};k,k'}^{M,N}(\vec{r}_1,\vec{r}_2)\Omega_{ia;k}^{M}(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\alpha;N}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2. \tag{45}$$

employing the notation for the embedding potential (eqn (29)), the embedding kernel (eqn (40)) and the derivatives of orbital overlap distributions (summarized in Table 6). Detailed working expressions used for the practical implementation of eqn (43)–(45) are presented in the Appendix (eqn (71)).

Terms dependent on one subsystem, eqn (43) and (44), are in effect analogous to the XC contributions to the property gradient in the OMO basis,[40] only with derivatives of the interaction energy replacing derivatives of the XC energy. In the LR algorithm[79] the property gradient is calculated once and is not updated in the iterative procedure, therefore the computational cost of including FDE-LAO terms does not significantly increase the overall cost of the calculations, unless the coupling terms (eqn (45)) are considered, which will be briefly discussed in the following section.

*Coupling kernel contributions to the linear response function.* As terms dependent on the embedding kernel (eqn (40)) may involve perturbed densities of two different subsystems ($M \neq N$), they will introduce a coupling between these subsystems through Coulomb and non-additive terms.

The coupling kernel contributions to the electronic Hessian (eqn (42)) have been discussed at length in the context of excitation energies[30,32,86,88,89] or electric polarizabilities,[31] and are shown to be important for extensive properties or when excitations cannot be considered (to good accuracy) as dominated by local components, but can often be neglected otherwise.[35–38,90–92]

The coupling kernel contributions to the property gradient have received less attention so far. As all FDE-LAO contributions to the property gradient result from using LAOs, which shift the gauge origin from an arbitrary point to the center of nuclei, this coupling term can be regarded as a small correction due to the shift of the origin – *e.g.* in the center of mass of subsystem *I* – to the centers of nuclei in subsystem *II*. While we still lack the

coupling contributions to the Hessian, in the following section we shall nevertheless investigate the relative importance of this term in the property gradient.

The presence of coupling terms increases the complexity and cost of calculations. As LR equations are solved for one subsystem at a time, the necessary derivatives of the density of subsystem *II* have to be calculated and stored before the response equations for subsystem *I* are invoked. Eqn (45) involves the non-additive $E_{\mathrm{xc}k}$ term (calculated analogously as the uncoupled $E_{\mathrm{xc}k}$ part), as well as the Coulomb contribution,

$$\int w_{\mathrm{emb};0}^{I,II;\mathrm{Coulomb}} \breve{\Omega}_{jj;0}^{B_\alpha;II}(\vec{r}_1)\Omega_{ia;0}^{I}(\vec{r})\mathrm{d}\vec{r}_1, \tag{46}$$

which in our implementation is calculated *via* numerical integration of the expression

$$\int\left[\sum_{\mu\nu} c_{\mu i}{}^{*} c_{\nu a}\int\frac{\chi_\mu^{I,\dagger}(\vec{r}_1)\Sigma_0\chi_\nu^{I}(\vec{r}_1)}{|\vec{r}-\vec{r}_1|}\mathrm{d}\vec{r}_1\right]\breve{\Omega}_{jj,0}^{B_\alpha,II}(\vec{r}s)\mathrm{d}\vec{r}. \tag{47}$$

*2.2.1.2 Expectation values.* In a manner similar to the total energy in eqn (25), the expectation value term (second term in eqn (16)) can be subdivided into subsystem and interaction parts:

$$\left.\frac{\partial^2 E}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 = \left.\frac{\partial^2 E_I}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 + \left.\frac{\partial^2 E_{II}}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 + \left.\frac{\partial^2 E_{\mathrm{int}}}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0. \tag{48}$$

The subsystem contributions will have the same form as discussed elsewhere, whereas the interaction term is

$$\left.\frac{\partial^2 E_{\mathrm{int}}}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 = \sum_{M=I,II}\int\left.\frac{\delta E_{\mathrm{int}}}{\delta\rho_k^M}\frac{\partial^2\rho_k^M}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 \tag{49}$$

$$+\sum_{M,N=I,II}\iint\left.\frac{\delta^2 E_{\mathrm{int}}}{\delta\rho_k^M\delta\rho_{k'}^N}\frac{\partial\rho_k^M}{\partial\varepsilon_1}\right|_0\left.\frac{\partial\rho_{k'}^N}{\partial\varepsilon_2}\right|_0. \tag{50}$$

From this expression one can, as in the linear response case, distinguish interaction contributions to each of the subsystems,

$$\begin{aligned}\left.\frac{\partial^2 E_{\mathrm{int}}^M}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 &= \int\left.\frac{\delta E_{\mathrm{int}}}{\delta\rho_k^M}\frac{\partial^2\rho_k^M}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 + \iint\left.\frac{\delta^2 E_{\mathrm{int}}}{\delta\rho_k^M\delta\rho_{k'}^M}\frac{\partial\rho_k^M}{\partial\varepsilon_1}\right|_0\left.\frac{\partial\rho_{k'}^M}{\partial\varepsilon_2}\right|_0 \\ &= \int v_{\mathrm{emb};k}^{M}(\vec{r})\breve{\Omega}_{ii;k}^{B_\alpha B_\beta;M}\mathrm{d}\vec{r}\end{aligned} \tag{51}$$

$$+\iint w_{\mathrm{emb};k,k'}^{M,M}(\vec{r}_1,\vec{r}_2)\breve{\Omega}_{ii;k}^{B_\alpha;M}(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\beta;M}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2, \tag{52}$$

containing embedding potential and kernel contributions, and those which depend on both subsystems ($M \neq N$)

$$\begin{aligned}\left.\frac{\partial^2 E_{\mathrm{int}}^{MN}}{\partial\varepsilon_1\partial\varepsilon_2}\right|_0 &= \iint\left.\frac{\delta^2 E_{\mathrm{int}}}{\delta\rho_k^M\delta\rho_{k'}^N}\frac{\partial\rho_k^M}{\partial\varepsilon_1}\right|_0\left.\frac{\partial\rho_{k'}^N}{\partial\varepsilon_2}\right|_0 \\ &= \iint w_{\mathrm{emb};k,k'}^{M,N}(\vec{r}_1,\vec{r}_2)\breve{\Omega}_{ii;k}^{B_\alpha;M}(\vec{r}_1)\breve{\Omega}_{jj;k'}^{B_\beta;N}(\vec{r}_2)\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2,\end{aligned} \tag{53}$$

made-up exclusively of a coupling kernel term. Working equations are presented in the Appendix (eqn (72)).

*Coupling kernel contributions to the expectation value.* As in the case of the property gradient, all terms in eqn (49) and (50) would be zero in perturbation-independent basis sets, thus from now on they will be referred to as FDE-LAO contributions to the expectation value. Although eqn (53) depends on the coupling kernel, it does not involve the relaxation of the subsystem densities as in the electronic Hessian but rather the static correction to the choice of response parameters. Here, we note that the first-quantization and second-quantization formulations of second-order magnetic properties in the LAO basis define expectation value contributions differently.[47]

### 2.3 Tensor expressions for the molecular properties and their representation in terms of magnetically induced currents

The theory discussed above is sufficient to determine the properties of interest in the subsystem approach. However, these properties can also be presented in a different mathematical form using the linearity of the 4c DC Hamiltonian in the applied perturbations, complemented by the formulation involving magnetically induced current densities, which more directly conveys the physical characteristics of each property.

**2.3.1 NMR shielding and indirect spin–spin coupling tensors.** The NMR shielding or the NMR indirect spin–spin coupling tensors in eqn (21) and (22) can be recast in a computationally advantageous form[47] in terms of expectation values involving the hyperfine operator for a nucleus $L$, the unperturbed spinors $|\psi_i\rangle$ and the first-order perturbed spinors[40,47] $\left|\tilde{\psi}_i^{\varepsilon_i}\right\rangle$, yielding the general expression

$$M_{\alpha\beta}^{\varepsilon;L} = \sum_i\left\{\left\langle\tilde{\psi}_i^{\varepsilon_\alpha}\middle|\hat{h}_{m_{L;\beta}}\middle|\psi_i\right\rangle + \left\langle\psi_i\middle|\hat{h}_{m_{L;\beta}}\middle|\tilde{\psi}_i^{\varepsilon_\alpha}\right\rangle\right\}. \tag{54}$$

The expression for the shielding tensor $\sigma_{\alpha\beta}^{L}$ is therefore obtained from eqn (54) by employing the spinors perturbed by the external magnetic field ($\varepsilon = \vec{B}$), $\left|\tilde{\psi}_i^{B_\alpha}\right\rangle$, and by the same token the spin–spin coupling tensor $K_{\alpha\beta}^{KL}$ is obtained by employing the spinors perturbed by the nuclear magnetic dipole ($\varepsilon = \vec{m}_K$), $\left|\psi_i^{m_{K;\alpha}}\right\rangle$.

In the FDE case, as each subsystem is described by its own set of externally-orthogonal orbitals, we can rewrite the expression in eqn (54) as

$$M_{\alpha\beta}^{\varepsilon;L} = \sum_{i\in I}\left\{\left\langle\tilde{\psi}_i^{\varepsilon_\alpha}\middle|\hat{h}_{m_{L;\beta}}\middle|\psi_i\right\rangle + \left\langle\psi_i\middle|\hat{h}_{m_{L;\beta}}\middle|\tilde{\psi}_i^{\varepsilon_\alpha}\right\rangle\right\} \tag{55}$$

$$+\sum_{j\in II}\left\{\left\langle\tilde{\psi}_j^{\varepsilon_\alpha}\middle|\hat{h}_{m_{L;\beta}}\middle|\psi_j\right\rangle + \left\langle\psi_j\middle|\hat{h}_{m_{L;\beta}}\middle|\tilde{\psi}_j^{\varepsilon_\alpha}\right\rangle\right\}. \tag{56}$$

The FDE expression for $\sigma_{\alpha\beta}^{L}$ or $K_{\alpha\beta}^{KL}$ can be further approximated by neglecting the terms arising from eqn (56). In the case of NMR shieldings, assuming that nucleus $L$ belongs to subsystem $I$, this approximation should be sufficient, especially if the overlap between the two subsystems is small, but whatever the case we can estimate this missing contribution using the magnetically-induced current density formulation outlined in Section 2.3.3. For the spin–spin tensor, this approximation should also be good due to the local nature of the hyperfine

operator, if both $K$ and $L$ belong to subsystem $I$ (a restriction in our current implementation).

**2.3.2 Magnetizability tensor.** Contrary to NMR properties, the magnetizability tensor is not a local property as the Zeeman operator (eqn (8)) affects the whole system. It can be expressed in terms of the sum of (interacting) intra-subsystem and inter-subsystem contributions

$$\xi_{\alpha\beta} = \xi_{\alpha\beta}^{I,(II)} + \xi_{\alpha\beta}^{II,(I)}, \tag{57}$$

where

$$\xi_{\alpha\beta}^{I,(II)} = \left[\frac{\partial^2 E_I}{\partial B_\alpha \partial B_\beta}\bigg|_0 + \frac{\partial^2 E_{\text{int}}^{M=I}}{\partial B_\alpha \partial B_\beta}\bigg|_0\right] \tag{58}$$

$$+\left[\frac{\partial^2 E_I}{\partial \kappa_{pq}^I \partial B_\beta} + \frac{\partial^2 E_{\text{int}}}{\partial \kappa_{pq}^I \partial B_\beta}\right]\frac{\partial \kappa_{pq}^I}{\partial B_\alpha}. \tag{59}$$

The terms in eqn (59) are calculated by solving the linear response equations for subsystem $I$ with FDE-LAO contributions to the property gradient (eqn (43)–(45)). The term involving $E_{\text{int}}^{M=I}$ in eqn (58) will contain intra-subsystem contributions of eqn (51) and (52) with the summation restricted to subsystem $I$ and the inter-subsystem contribution from eqn (53), where the summation is constrained to have $M = I$. The second term in eqn (59), $\xi_{\alpha\beta}^{II,(I)}$, is obtained by permutation of indices $I$ and $II$.

**2.3.3 Tensors in terms of induced currents.** The relativistic current density vector and its first-order derivatives with respect to perturbations,[93]

$$\vec{j}(\vec{r}) = -e\sum_i \psi_i^\dagger c\vec{\alpha}\psi_i \tag{60}$$

$$\vec{j}^{\varepsilon_1}(\vec{r}) = -e\sum_i \left\{\tilde{\psi}_i^{\varepsilon_1;\dagger} c\vec{\alpha}\psi_i + \psi_i^\dagger c\vec{\alpha}\tilde{\psi}_i^{\varepsilon_1}\right\}, \tag{61}$$

allow property densities to be constructed,[94] which may be visualized on a grid and integrated – giving the value of the corresponding property. Thus, the properties studied in this paper can be written as:

$$\sigma_{\alpha\beta}^K = -\frac{1}{c^2}\int \frac{1}{r_K^3}\left(\vec{r}_K \times \vec{j}^{B_\alpha}\right)_\beta \mathrm{d}\vec{r}, \tag{62}$$

$$K_{\alpha\beta}^{KL} = -\frac{1}{c^2}\int \frac{1}{r_K^3}\left(\vec{r}_K \times \vec{j}^{m_{L;\alpha}}\right)_\beta \mathrm{d}\vec{r}. \tag{63}$$

$$\xi_{\alpha\beta} = -\frac{1}{2}\int \left(\vec{r}_G \times \vec{j}^{B_\alpha}\right)_\beta \mathrm{d}\vec{r}. \tag{64}$$

where the first-order current density perturbed by an external magnetic field is calculated with LAOs.[95]

An advantage of using the induced current density is that while evaluated for one subsystem, it can be contracted with the position vector pointing to the other subsystem allowing the evaluation of contributions, for example from eqn (56), for the NMR shielding tensor as:

$$\sigma_{\alpha\beta}^{K;II} = -\frac{1}{c^2}\int \frac{1}{\left(\vec{r}_i^{II} - \vec{R}_K^I\right)^3}\left(\left(\vec{r}_i^{II} - \vec{R}_K^I\right) \times \vec{j}^{B_\alpha;II}\right)_\beta \mathrm{d}\vec{r}, \tag{65}$$

where the superscripts $I$ and $II$ denoting the subsystems are written explicitly for each vector. Eqn (65) is analogous to nucleus independent chemical shift (NICS) calculations outlined for FDE by Jacob and Visscher.[37]

# 3 Computational details

We have investigated three hydrogen-bonded HXH–$OH_2$ complexes, where X = Se, Te, or Po. Their structures were optimized in ADF software,[96] using the scalar version of the zeroth-order regular approximation (ZORA)[97,98] Hamiltonian, the B3LYP[99] functional and basis sets of the triple-zeta quality (TZ2P).[100] The optimized structures are included in the ESI.† The structures for the subsystems are taken from supermolecules without any further reoptimization, so that calculation on isolated fragments can be thought of as equivalent to QM/MM embedding where only mechanical (“ME”) coupling between the subsystems is taken into account.[27]

The wave function optimization and magnetic property calculations performed in a development version of the DIRAC code[51] employed the DC Hamiltonian and the PBE[101,102] functional. In the FDE calculations, the non-additive exchange–correlation and kinetic energy contributions were calculated using the PBE and PW91k[103] functionals, respectively. In response calculations, we have used the full derivatives of the PBE and PW91k functionals provided by the XCFun library.[104] The basis sets were of augmented triple-zeta quality: aug-cc-pVTZ[105] for H and O and dyall.acv3z[106,107] for X.

The calculations of NMR properties with the spin–orbit ZORA Hamiltonian (ZORA-SO) were performed in ADF using the TZ2P basis set and the PBE functional. For FDE calculations in ADF, we have also employed the PBE and PW91k functionals for the non-additive contributions.

In both DIRAC and ADF, the Gaussian model of nuclear charge distribution[108] was used and in the case of DIRAC, the (SS|SS) class of two-electron integrals was replaced by a standard correction.[109] Also, in both cases we performed two sets of FDE calculations, one using densities obtained for the isolated subsystems as frozen densities (hereafter referred to as “FDE(0)”) and another where we optimized both subsystem densities by exchanging their role as frozen/active densities in the “freeze–thaw” procedure, which was stopped after 4 iterations with both densities fully optimized (hereafter referred to as “FDE(4)”). In tables in this paper, we present only the latter results, while full tables are available in the ESI.†

We note that the choice of PBE was motivated by minimizing the differences in the computational setup between supermolecular and FDE calculations, so the only additional approximation in the FDE case comes from the kinetic energy functional. Since our aim is to compare the supermolecular and FDE results, a thorough study of the performance of different functionals (exchange–correlation and/or kinetic) is beyond the scope of this paper.

We use the definitions of Mason[110] for the isotropic and the anisotropic parts of a tensor $\Omega$ in the principal axis system, where $\Omega_{33} \geq \Omega_{22} \geq \Omega_{11}$,

$$\Omega_{\text{iso}} = 1/3(\Omega_{11} + \Omega_{22} + \Omega_{33}) \tag{66}$$

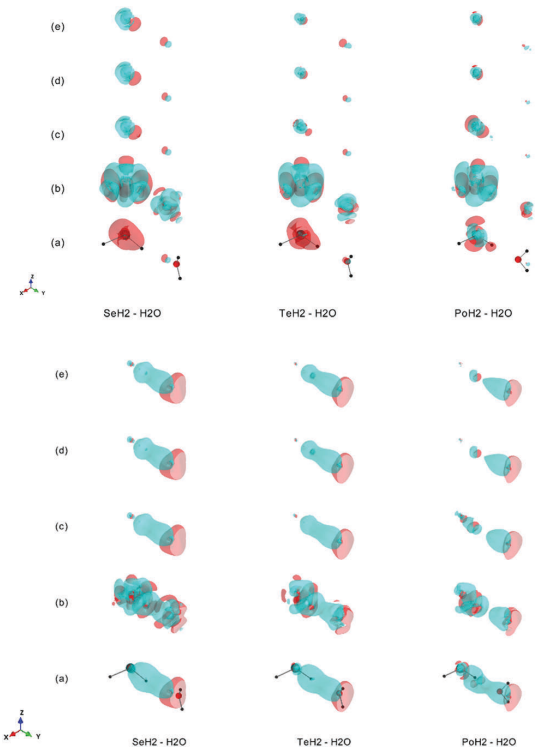



Fig. 1 Differential isotropic shielding density isosurfaces[119] (isovalues at +0.53 ppm (red) and −0.53 ppm (blue)) for $XH_2$–$H_2O$ systems with X = Se, Te, Po (top) and $H_b$ (bottom), calculated as a difference between supermolecule shielding densities and the sum of subsystem shielding densities approximated as: (a) ME, (b) FDE(4)[0], (c) FDE(4)[$v$], (d) FDE(4)[$v + w_u$] and (e) FDE(4)[$v + w_{all}$]. Color of atoms: X (blue), O (red), H (grey).

$$\Omega_{aniso} = \Omega_{33} - 1/2(\Omega_{11} + \Omega_{22}). \quad (67)$$

We have also calculated and plotted NMR shielding densities and their differences. In this paper, we present the differential NMR shielding densities calculated by subtracting the shielding density of nucleus X and $H_b$ in Fig. 1 arising from the induced current in both subsystems (employing the NICS method for the frozen subsystem) from the corresponding shielding density in a supermolecule. The shielding density is plotted on a 80 × 80 × 80 grid of points generated for supermolecules. The plots were made using Mayavi – a library for interactive scientific data visualization and 3D plotting in Python.[111] More plots and relevant data can be found in the ESI.†

## 4 Results and discussion

In what follows we shall present all embedding results relative to the supermolecular ones. Thus, for a given molecular property $P$ we define absolute ($\Delta P^f$) and relative ($\delta P^f$) property shifts in a general manner as

$$\Delta P^f = P^{super} - (P^{I,f} + P^{II,f}) \quad (68)$$

$$\delta P^f = \Delta P^f / P^{super}, \quad (69)$$

where $P^{super}$ corresponds to the supermolecular value of property $P$ and $P^{i,f}$ denotes the contribution to the property from the subsystem $i$. The latter is obtained for the isolated subsystem ($f$ = ME) or using FDE ($f$ = FDE($n$), with $n$ = 0, 4). In the case of NMR shieldings and indirect spin–spin couplings, which are essentially local to one of the subsystems, $\Delta P^f$ is well approximated by neglecting $P^{II,f}$ in eqn (68). In the case of FDE calculations, we introduce an additional notation in order to discern FDE-LAO contributions to the property gradient (NMR shielding and magnetizability) and to the expectation value (magnetizability), which can be either neglected (FDE($n$)[0]), limited to the embedding potential (FDE($n$)[$v$], eqn (43) and (51)) or to the embedding potential and the uncoupled kernel (FDE($n$)[$v + w_u$], eqn (44) and (52)) or also incorporated with the coupling kernel ($w_c$) terms (FDE($n$)[$v + w_{all}$], eqn (45) and 53).

### 4.1 NMR shielding tensor

The DC calculations of isotropic and anisotropic parts of the NMR shielding tensor are summarized in Table 1. We present only the results for nuclei of the active subsystem ($H_2X$) and we distinguish between the hydrogen involved in the hydrogen bond ($H_b$) and the other pointing away from the water molecule.

**4.1.1 Isolated subsystems.** From the isolated ("ME") calculation, we observe that the hydrogen-bonded water strongly affects the isotropic and anisotropic parts of the NMR shielding tensors of nuclei of the active subsystems, leading to the shielding of the heavy centers and the deshielding of $H_b$, in agreement with established observations on hydrogen-bonded systems.[112]

The values of $\Delta\sigma_{iso}^{ME}$ become progressively larger with the increase of the atomic number of X for all nuclei of an active subsystem: for the X nuclei they range from 38 ppm for $SeH_2$ to 138 ppm for $PoH_2$, for the non H-bonded hydrogen nuclei – from 0.4 ppm for $SeH_2$ to 0.9 ppm for $PoH_2$ and for $H_b$ nuclei – from −2 ppm for $SeH_2$ to −0.6 ppm for $PoH_2$, which for hydrogen nuclei are significant since $^1$H NMR shielding is between 10 ppm and 30 ppm in most applications.[14]

While these $\Delta\sigma$ values are relatively small in comparison to the absolute shieldings, they can nevertheless be significant in NMR experiments – for instance, both $^{77}$Se and $^{125}$Te nuclei are known to be very sensitive to the environment (*e.g.* solvent, its concentration and temperature[113]) and even though they span wide chemical shift ranges (6000 ppm for $^{77}$Se[114,115] and 7000 ppm[113] for $^{125}$Te), shifts of around 30 ppm (Se) or 60 ppm (Te) are fingerprints of a specific solvent.[113,116–118]

For the anisotropies, the isolated ("ME") results are usually very different from the supermolecular ones for $H_b$, with $\Delta\sigma_{aniso}^{ME}$ of 7 ppm for $SeH_2$, −1 ppm for $TeH_2$ and −6 ppm for $PoH_2$, which represent deviations of 33%, 9% and 6%, respectively. This is not unexpected, since in order to properly capture the directionality of the hydrogen bond, electronic effects must be taken into account. Interestingly, these discrepancies are also seen for the Po nucleus in $PoH_2$, whose shielding anisotropy differs from the supermolecular value by 464 ppm (or 8% difference

Table 1 Absolute DC isotropic and anisotropic shielding values ($\sigma_{iso}^{super}$ and $\sigma_{aniso}^{super}$, in ppm) of nuclei in the $H_2X$ subsystems (X = Se, Te, Po) in $H_2X$–$H_2O$, and absolute shifts ($\Delta\sigma$, in ppm) for the isolated ("ME") and embedded ("FDE(4)") $H_2X$ molecules in the presence of $H_2O$. For FDE, the values for different approximations in the FDE-LAO treatment (*a*: [0], *b*: [$v$], *c*: [$v + w_u$], *d*: [$v + w_{all}$]) are shown

| Atom | $\sigma_{iso}^{super}$ | $\Delta\sigma_{iso[a]}^{FDE(4)}$ | $\Delta\sigma_{iso[b]}^{FDE(4)}$ | $\Delta\sigma_{iso[c]}^{FDE(4)}$ | $\Delta\sigma_{iso[d]}^{FDE(4)}$ | $\Delta\sigma_{iso}^{ME}$ |
|---|---|---|---|---|---|---|
| Se | 2378.03 | −100.75 | −12.54 | −12.55 | −12.33 | 38.25 |
| $H_b$ | 30.88 | −0.39 | −0.83 | −0.67 | −0.62 | −2.09 |
| H | 33.42 | 0.24 | 0.02 | 0.03 | 0.03 | 0.43 |
| Te | 4667.85 | −142.39 | −4.62 | −9.16 | −8.88 | 67.48 |
| $H_b$ | 35.62 | −0.29 | −0.68 | −0.44 | −0.42 | −1.70 |
| H | 37.85 | −0.02 | 0.01 | 0.02 | 0.02 | 0.54 |
| Po | 13985.80 | −224.54 | 18.28 | −3.52 | −3.13 | 137.84 |
| $H_b$ | 40.80 | 0.50 | 0.07 | −0.09 | −0.09 | −0.57 |
| H | 42.29 | −0.22 | −0.03 | 0.00 | 0.00 | 0.89 |

| Atom | $\sigma_{aniso}^{super}$ | $\Delta\sigma_{aniso[a]}^{FDE(4)}$ | $\Delta\sigma_{aniso[b]}^{FDE(4)}$ | $\Delta\sigma_{aniso[c]}^{FDE(4)}$ | $\Delta\sigma_{aniso[d]}^{FDE(4)}$ | $\Delta\sigma_{aniso}^{ME}$ |
|---|---|---|---|---|---|---|
| Se | 609.27 | 21.80 | −0.13 | −0.43 | −0.19 | 0.17 |
| $H_b$ | 22.19 | 5.05 | 6.00 | 5.76 | 5.70 | 7.29 |
| H | 15.13 | −1.17 | −0.22 | −0.22 | −0.21 | 0.09 |
| Te | 1189.67 | 29.78 | 0.56 | −1.22 | −1.21 | 3.60 |
| $H_b$ | 14.59 | 0.50 | 0.02 | 0.24 | 0.28 | −1.31 |
| H | 15.24 | 0.54 | 0.09 | 0.08 | 0.08 | −0.64 |
| Po | 5556.67 | 61.88 | −15.12 | −18.36 | −18.61 | −463.51 |
| $H_b$ | 105.25 | 0.98 | −1.28 | −1.67 | −1.67 | −6.01 |
| H | 107.80 | 1.93 | 0.16 | 0.08 | 0.09 | −4.03 |

from the supermolecular value), whereas no significant deviations are seen for Se or Te.

These tendencies can be better seen in the plots of the differential isotropic shielding density in Fig. 1(a) for the heavy centers and $H_b$. This figure exhibits positive (pink) and negative (blue) isosurfaces, which, respectively, depict more shielded and more deshielded areas in a supermolecule than in the embedded subsystems, and which upon integration give the corresponding values of $\Delta\sigma_{iso}^{ME}$.

We observe from Fig. 1(a) that the plots are rather similar for Se and Te nuclei, exhibiting small negative isosurfaces centered on a heavy nucleus, surrounded by a much larger positive isosurface elongated on the X–$H_b$ bond. In the case of Po the differential shielding density is represented by much more complex isosurfaces around the heavy center, as the negative isosurface centered on Po is larger than that observed for Se and Te and surrounded by many well-separated positive lobes. This indicates that even though the property shift $\Delta\sigma_{iso}^{ME}$(Po) turns out to be relatively modest compared to the value of the absolute shielding, it is a result of shielded and deshielded areas cancelling out upon integration, reflecting the intricate interplay between environmental and relativistic effects, which are quite different for Te and Se.

Towards the heavier neighbouring center, hydrogen nuclei experience larger HALA effects[120] (reflected by increasing absolute values of $\sigma_{iso}^{super}(H_b)$) contributing to the shielding of $H_b$ nuclei, and competing with the deshielding effect caused by the interaction with water molecules. We also observe the difference in the non-hydrogen bonded hydrogen shielding between $PoH_2$ and the other species, which could also be a consequence of the HALA effect.

**4.1.2 Frozen density embedding.** In what follows we discuss the relative importance of the FDE-LAO contributions ($v$, $w_u$, $w_c$), both in terms of the values of isotropic and anisotropic $\Delta\sigma^{FDE(4)}$ indices as well as from the plots of the differential isotropic shielding density in Fig. 1(b)–(e). We also calculated contributions to the shieldings of X and $H_b$ from the frozen subsystem using the induced current density formulation from eqn (65), but since for all approximations these were found to be smaller than the assumed accuracy (0.01 ppm), they are not shown separately.

Starting with the calculation where no FDE-LAO terms are added to the property gradient, the results of isotropic shieldings are much worse than the "ME" values for all nuclei considered. The same conclusion is drawn for the anisotropic shieldings of Se and Te nuclei, but not for the Po nucleus, for which the $\Delta\sigma_{aniso}^{FDE(4)}$ value remains smaller than the one incorporating only the mechanical coupling.

The inclusion of FDE-LAO[$v$] contributions yields a significant improvement over the ME values. For the heavy elements, $\Delta\sigma_{iso}^{FDE(4)}$ increases from the very negative values obtained without FDE-LAO terms to still significant negative values for Se (−13 ppm), less negative for Te (−4 ppm) and positive for Po (18 ppm). For $H_b$, the inclusion of the potential acts in the opposite direction and we observe a relatively small decrease of $\Delta\sigma_{iso}^{FDE(4)}$ values, whereas for the other hydrogen there is little change.

Uncoupled kernel contributions ($v + w_u$) further improve the results as they partially offset the $v$ contribution and reduce the $\Delta\sigma_{iso}^{FDE(4)}$ to rather small values. The $w_u$ correction is much more significant for Po than for Te, and for Se only little improvement is seen. The reason for this difference among elements is that the kernel terms introduce contributions from the response of the spin-density which becomes more significant as the elements become heavier. This is also the reason why $w_u$ contributions affect the shielding of $H_b$ atoms (spin–orbit mechanism of HALA effect). The uncoupled kernel also accentuates the trend seen for $\Delta\sigma_{aniso}^{FDE(4)}$(X), making it more negative and therefore overestimating the shielding anisotropy of X. This overestimation is also observed for the $\sigma_{aniso}^{FDE(4)}(H_b)$ except in $SeH_2$.

The coupling terms ($w_c$), on the other hand, act in general to offset the uncoupled kernel terms but their magnitude is, as expected from the local nature of the NMR shielding, much smaller than the latter for all atoms so it plays no significant role in either the isotropic shielding value or the shielding anisotropy.

All of the above let us conclude that, while FDE isotropic shieldings are rather good and relatively much better as the systems become heavier, there are still significant shortcomings in the description of the anisotropies in these systems. The key to further improving the results is in ameliorating the leading FDE-LAO contribution ($v$), and a possible way to do so is *via* the use of accurate approximations to the non-additive kinetic energy contributions.[85,86]

### 4.2 Magnetizability

The magnetizability tensors calculated using the DC Hamiltonian are summarized in Table 2, where we present the isotropic ($\xi_{iso}$)

and first anisotropy ($\xi_{aniso1}$) values. The results for the second anisotropy, $\xi_{aniso2}$, can be found in the ESI.†

The magnetizability is an extensive property and therefore in order to compare to the supermolecular values one should obtain the tensors for both subsystems. Another of its features, demonstrated by numerous studies, is that its value in a molecule can be very well approximated by the sum of contributing atomic susceptibilities,[121–123] with only few exceptions, *e.g.* aromatic hydrocarbons,[121] small molecules containing fluorine[124] or metal clusters.[19] This curious additivity of the magnetizability tensor – known as Pascal's rule[125,126] – has been attributed to the local diamagnetic nature of atoms in molecules[124,127] and the breakdown of this rule to the "long-range circulation of electrons" not accounted for in the atomic picture.[127]

While Pascal's rule has been defined in terms of atomic susceptibilities, one could consider the additivity of magnetizability of molecular assemblies in terms of its constituent molecules. Recent studies[128] have shown that Pascal's rule is particularly useful when analysed in terms of the magnetically induced current density, as the interaction of induced currents in neighbouring molecules and the increase of paramagnetic component of magnetizability tensor can be connected to the breakdown of the additivity rule.

Table 2 DC isotropic and first anisotropic magnetizabilities ($\xi_{iso}$ and $\xi_{aniso1}$, in SI units) for the $H_2X$–$H_2O$ systems (X = Se, Te, Po) as well as for the $H_2X$ and $H_2O$ subsystems, with the latter as isolated ("ME") and embedded ("FDE(4)") calculations. In the case of subsystem calculations, the total $\xi^{tot}_{iso}$ and $\xi^{tot}_{aniso1}$ are given as the sum of the subsystem values. For FDE, the values for the different approximations in the FDE-LAO treatment (*a*: [0], *b*: [$\nu$], *c*: [$\nu + w_u$], *d*: [$\nu + w_{all}$]) are shown

| System | $\xi^{super}_{iso}$ | $\xi^{FDE(4)}_{iso[a]}$ | $\xi^{FDE(4)}_{iso[b]}$ | $\xi^{FDE(4)}_{iso[c]}$ | $\xi^{FDE(4)}_{iso[d]}$ | $\xi^{ME}_{iso}$ |
|---|---|---|---|---|---|---|
| $SeH_2$ | — | −183.07 | −608.99 | −606.54 | −606.54 | −602.19 |
| $H_2O$ | — | 781.33 | −233.79 | −233.78 | −233.76 | −234.19 |
| $\xi^{tot}$ | −836.26 | 598.26 | −842.77 | −840.32 | −840.30 | −836.31 |
| $\Delta\xi$ | 0.0 | −1434.52 | 6.51 | 4.06 | 4.04 | 0.05 |
| $TeH_2$ | — | −630.07 | −858.94 | −848.68 | −848.74 | −842.57 |
| $H_2O$ | — | 235.10 | −233.74 | −233.74 | −233.60 | −233.83 |
| $\xi^{tot}$ | −1080.67 | −394.97 | −1092.69 | −1082.42 | −1082.33 | −1076.39 |
| $\Delta\xi$ | 0.0 | −685.71 | 12.01 | 1.74 | 1.66 | −4.28 |
| $PoH_2$ | — | −895.55 | −1030.19 | −949.80 | −949.71 | −940.09 |
| $H_2O$ | — | −169.52 | −234.11 | −234.10 | −233.18 | −234.02 |
| $\xi^{tot}$ | −1184.04 | −1065.08 | −1264.30 | −1183.91 | −1182.89 | −1174.11 |
| $\Delta\xi$ | 0.0 | −118.96 | 80.26 | −0.13 | −1.15 | −9.92 |

| System | $\xi^{super}_{aniso1}$ | $\xi^{FDE(4)}_{aniso1[a]}$ | $\xi^{FDE(4)}_{aniso1[b]}$ | $\xi^{FDE(4)}_{aniso1[c]}$ | $\xi^{FDE(4)}_{aniso1[d]}$ | $\xi^{ME}_{aniso1}$ |
|---|---|---|---|---|---|---|
| $SeH_2$ | — | −358.80 | −47.12 | −45.88 | −45.88 | −45.97 |
| $H_2O$ | — | −1396.29 | −5.76 | −5.76 | −5.78 | −6.51 |
| $\xi^{tot}$ | −57.94 | −1755.08 | −52.88 | −51.64 | −51.66 | −52.48 |
| $\Delta\xi$ | 0.0 | 1697.14 | −5.07 | −6.31 | −6.28 | −5.47 |
| $TeH_2$ | — | −227.24 | −86.32 | −80.29 | −80.32 | −80.30 |
| $H_2O$ | — | −650.52 | −5.53 | −5.53 | −5.70 | −6.06 |
| $\xi^{tot}$ | −81.63 | −877.77 | −91.85 | −85.82 | −86.02 | −86.36 |
| $\Delta\xi$ | 0.0 | 796.14 | 10.23 | 4.19 | 4.39 | 4.73 |
| $PoH_2$ | — | −290.25 | −261.88 | −91.58 | −91.56 | −91.79 |
| $H_2O$ | — | −93.62 | −6.45 | −6.45 | −7.47 | −6.73 |
| $\xi^{tot}$ | −89.81 | −383.87 | −268.34 | −98.03 | −99.02 | −98.52 |
| $\Delta\xi$ | 0.0 | 294.06 | 178.53 | 8.22 | 9.21 | 8.71 |

**4.2.1 Isolated subsystems.** As can be seen in Table 2, the $\Delta\xi^{ME}_{iso}$ values are rather small in absolute ($<0.1$ for Se, −4.3 for Te and −9.9 for Po, in SI units) as well as in relative terms. From these results, one would be justified in considering that Pascal's rule holds rather well for these systems.

The property shifts are larger for $\xi_{aniso1}$, with $\Delta\xi^{ME}_{aniso1}$ that tend to increase with an increase of the atomic number of X (−5.5 for Se, 4.7 for Te and 8.7 for Po, in SI units) and an inverse trend for $\Delta\xi^{ME}_{aniso2}$. It is more interesting, however, that these values amount overall to much larger relative differences (9.4%, −5.8%, −9.7% for Se, Te and Po for $\Delta\xi^{ME}_{aniso1}$, respectively), and the change of sign along the series again indicates a complex interplay between relativistic and environmental effects in this sort of embedding.

**4.2.2 Frozen density embedding.** For magnetizability, FDE-LAO contributions are also present in the property gradient and in the expectation value part, thus it is again useful to consider the relative importance of each of them. We shall here once again focus on the FDE(4) results, noting that the FDE(0) results follow these closely but show a slightly worse agreement with supermolecular values than FDE(4). The $\Delta\xi^{FDE(4)}_{iso}$ values calculated with no FDE-LAO terms are much larger in magnitude than those for the ME calculations. Unlike for $\sigma$, even the inclusion of FDE-LAO[$\nu$] terms yields $\Delta\xi^{FDE(4)}_{iso}$ values which are significantly larger in magnitude (6.5 for Se, 12 for Te and 80.3 for Po, in SI units) than those for the isolated ("ME") calculations.

The uncoupled kernel contribution ($\nu + w_u$) brings about little changes in $\Delta\xi^{FDE(4)}_{iso}$ for Se but a significant improvement for Te and, especially, for Po. As for the NMR shielding, the coupling kernel ($\nu + w_{all}$) term has a small effect, further leading to a decrease in the $\Delta\xi^{FDE(4)}_{iso}$ values.

These results indicate that the additivity of $\xi_{iso}$ resulting from FDE calculations with all second-order terms ($\nu + w_{all}$) is significantly better than from the isolated ("ME") calculations as the subsystem becomes heavier, and may suggest an inflexion point between Se and Te where electronic effects would become important enough for the results to start deviating from Pascal's rule.

A similar trend is found for the $\Delta\xi^{FDE(4)}_{aniso1}$ values, with FDE-LAO[$\nu$] calculations underperforming the isolated ("ME") ones and the kernel contributions being important for yielding a good agreement with reference values. As for the NMR shielding anisotropies, significant discrepancies with respect to the supermolecular results remain. The performance of FDE for the second anisotropy is slightly better but follows the same trends.

### 4.3 NMR spin–spin coupling tensor

The indirect reduced spin–spin coupling tensor calculated using the DC Hamiltonian for the H–$H_b$, X–H and X–$H_b$ pairs of nuclei corresponding to the $XH_2$ species are found in Table 3.

**4.3.1 Isolated subsystems.** The absolute values of isotropic one-bond spin–spin coupling constants (SSCCs) involving the heavy nuclei increase significantly and, due to relativity,[129] for $PoH_2$ are around 30 times larger than those in $SeH_2$. In relative terms, however, X–H SSCCs hardly change for the isolated ("ME") subsystems in relation to supermolecular values, with a slight increase as the systems become heavier (from about 2 for Se or Te

to 5 for Po, SI units), whereas for X–$H_b$ SSCCs an opposite trend is found (from 7–9 for Se/Te to −1 for Po, SI units).

The observation that environment effects are more important in the case of Se–$H_b$ and Te–$H_b$ than Se–H and Te–H, respectively, is intuitive considering that most of the studied one-bond spin–spin couplings are governed by Fermi contact interactions,[17] which probe the spin density at the coupled nuclei – expected to be perturbed more on $H_b$ than on the other H nucleus upon the formation of the hydrogen bond. That said, the $\Delta K_{iso}^{ME}$(Po–$H_b$) value is interesting in being the only one which is negative and smaller than the $\Delta K_{iso}^{ME}$(Po–H) value, something that may indicate that the spin–spin coupling mechanism in the $PoH_2$ molecule is more complex and may be dominated by other interactions, for instance spin–orbit-induced as in heavier interhalogen diatomics,[130] which may further be differently affected by environmental effects. Two-bond SSCCs (H–$H_b$) are in general very small for all systems, though there is a small increase in absolute terms as the systems become heavier.

For anisotropies, the values of $\Delta K_{aniso}^{ME}$ strongly increase for X–$H_b$ as X becomes heavier, though they remain very small for X–H, and for H–$H_b$ the isolated (''ME'') results are largely the same as the supermolecular ones.

**4.3.2 Frozen density embedding.** Unlike the other two properties discussed above, FDE contributions to the response only enter here in the electronic Hessian, which greatly simplifies the implementation. Here, we shall discuss the SSCC FDE(4) results which are, like for magnetizabilities, better than the FDE(0) results. We observe that FDE performs better than mechanical embedding, with $\Delta K_{iso}^{FDE(4)}$ values being consistently around 1–2 (SI units) for X–$H_b$ and smaller than 1 (SI units) for X–H or H–$H_b$.

Trends similar to those for $\Delta K_{iso}^{FDE(4)}$ are seen for $\Delta K_{aniso}^{FDE(4)}$, with the latter being generally small and slightly negative for all but $\Delta K_{aniso}^{FDE(4)}$(Po–$H_b$), which is of about 3 (SI units).

### 4.4 A Comparison of DC and ZORA Hamiltonians

Although SO-ZORA is known to yield rather different results from the DC ones, it is often sufficient in the determination of chemical shifts due to cancellation of errors.[21,44] Nevertheless when environmental effects are calculated as differences in absolute shieldings, the cancellation of errors is not always guaranteed.

In recent years, the differences in performance between the two Hamiltonians has gained attention, with a number of studies reporting significant discrepancies between SO-ZORA and DC values of NMR shielding tensors of heavy nuclei,[21,22,44,131] which were explained by a poor description of core orbitals of heavy elements by the ZORA Hamiltonian.[21]

It is therefore interesting to see how FDE and ME perform for the two Hamiltonians. The results of our calculations of NMR shieldings and SSCCs using the SO-ZORA Hamiltonian are shown in Tables 4 and 5, respectively.

A comparison of our SO-ZORA and DC results indicates that such error cancellation occurs for $SeH_2$, since the results and trends of shieldings and SSCCs are essentially the same for both Hamiltonians, for FDE and ME calculations. For $TeH_2$, both Hamiltonians yield largely similar results, but some quantitative differences start to appear in $\Delta\sigma_{aniso}$ and rather small differences in $\Delta\sigma_{iso}^{ME}$. For $PoH_2$, on the other hand, the differences are numerous: $\Delta\sigma_{iso}^{ME}$(Po) for SO-ZORA already differs from the DC value by 38 ppm, such a difference in $\Delta\sigma_{iso}^{FDE(4)}$ is about −16 ppm (DC using the ($\nu$ + $w_{all}$) FDE-LAO terms), whereas for $\Delta\sigma_{iso}^{ME}$($H_b$) and $\Delta\sigma_{iso}^{FDE}$($H_b$) these discrepancies are of the order of a ppm. The differences between Hamiltonians are also quite marked for anisotropies, where they amount to about 160 ppm for $\Delta\sigma_{aniso}^{ME}$(Po) and 30 ppm for $\Delta\sigma_{aniso}^{FDE}$(Po).

Large discrepancies between Hamiltonians are also seen for $PoH_2$ SSCCs, but much more marked for anisotropies than for isotropic values (about 1 SI unit for FDE or ME calculations).

While our dataset is rather small for drawing more general conclusions, it strengthens the case for a more thorough assessment of approximate Hamiltonians such as ZORA for calculating the NMR parameters of 6p molecules.

Here, one may also wonder the extent to which the quality of the basis sets used in the ZORA calculation may affect the results. While for valence properties one would expect the ZORA basis to be roughly comparable to the triple-zeta Gaussian basis used in the DC calculations, by construction the latter includes rather tight s and p functions, and that may make them more accurate for properties depending on the core region.

To that end, we have performed the same calculations as above using the TZP and QZ4P basis sets, with results presented in the ESI.† From these, we see that for the absolute SSCCs values

Table 3 DC isotropic and anisotropic reduced indirect spin–spin couplings ($K_{iso}^{super}$ and $K_{aniso}^{super}$, in SI units) for the $H_2X$ subsystems in $H_2X$–$H_2O$, and absolute shifts ($\Delta K$, in SI units) for the isolated (''ME'') and embedded (''FDE(4)'') $H_2X$ molecules in the presence of $H_2O$

| Nuclei | | $K_{iso}^{super}$ | $\Delta K_{iso}^{FDE(4)}$ | $\Delta K_{iso}^{ME}$ | $K_{aniso}^{super}$ | $\Delta K_{aniso}^{FDE(4)}$ | $\Delta K_{aniso}^{ME}$ |
|---|---|---|---|---|---|---|---|
| $H_b$ | Se | −11.22 | 1.52 | 7.25 | 113.79 | −1.23 | 1.25 |
| H | Se | −16.04 | 0.07 | 1.51 | 110.68 | −0.33 | −1.01 |
| $H_b$ | H | −0.77 | 0.00 | −0.02 | 0.89 | 0.00 | −0.01 |
| $H_b$ | Te | −53.11 | 1.69 | 9.08 | 208.54 | −0.42 | 8.07 |
| H | Te | −59.98 | 0.16 | 2.18 | 198.09 | −0.44 | −2.37 |
| $H_b$ | H | −0.75 | 0.00 | −0.03 | 0.42 | 0.00 | 0.00 |
| $H_b$ | Po | −442.55 | 1.37 | −1.25 | 429.17 | 3.10 | 41.96 |
| H | Po | −437.24 | 0.45 | 5.04 | 388.29 | −0.18 | 2.24 |
| $H_b$ | H | −0.61 | 0.00 | −0.04 | 0.69 | 0.00 | 1.32 |

Table 4 Absolute SO-ZORA isotropic and anisotropic shielding values ($\sigma_{iso}^{super}$ and $\sigma_{aniso}^{super}$, in ppm) of nuclei in the $H_2X$ subsystems (X = Se, Te, Po) in $H_2X$–$H_2O$, and absolute shifts ($\Delta\sigma$, in ppm) for the isolated (''ME'') and embedded (''FDE(4)'') $H_2X$ molecules in the presence of $H_2O$

| Atom | $\sigma_{iso}^{super}$ | $\Delta\sigma_{iso}^{FDE(4)}$ | $\Delta\sigma_{iso}^{ME}$ | $\sigma_{aniso}^{super}$ | $\Delta\sigma_{aniso}^{FDE(4)}$ | $\Delta\sigma_{aniso}^{ME}$ |
|---|---|---|---|---|---|---|
| Se | 2261.59 | −11.30 | 34.89 | 628.15 | 4.43 | 4.89 |
| $H_b$ | 30.10 | −0.77 | −2.07 | 23.95 | 5.83 | 7.22 |
| H | 32.55 | −0.03 | 0.35 | 16.97 | −0.24 | 0.11 |
| Te | 4251.23 | −9.61 | 64.66 | 1219.57 | 0.55 | 0.80 |
| $H_b$ | 33.43 | −0.46 | −1.59 | 18.09 | 3.87 | 4.90 |
| H | 35.47 | 0.00 | 0.45 | 13.09 | −0.33 | −0.09 |
| Po | 11168.82 | −20.93 | 101.35 | 3138.04 | −45.36 | −304.56 |
| $H_b$ | 37.46 | −0.21 | −0.82 | 62.41 | −1.32 | −4.11 |
| H | 38.99 | 0.00 | 0.67 | 64.80 | 0.25 | −2.05 |

Table 5 SO-ZORA isotropic and anisotropic reduced indirect spin–spin couplings ($K_{iso}^{super}$ and $K_{aniso}^{super}$, in SI units) for the $H_2X$ subsystems in $H_2X$–$H_2O$, and absolute shifts ($\Delta K$, in SI units) for the isolated ("ME") and embedded ("FDE(4)") $H_2X$ molecules in the presence of $H_2O$

| Nuclei | | $K_{iso}^{super}$ | $\Delta K_{iso}^{FDE(4)}$ | $\Delta K_{iso}^{ME}$ | $K_{aniso}^{super}$ | $\Delta K_{aniso}^{FDE(4)}$ | $\Delta K_{aniso}^{ME}$ |
|---|---|---|---|---|---|---|---|
| $H_b$ | Se | −12.63 | 1.08 | 6.58 | 130.61 | −2.01 | 1.15 |
| H | Se | −16.41 | 0.87 | 1.88 | 127.57 | −0.24 | −1.03 |
| $H_b$ | H | −1.11 | −0.03 | −0.07 | 0.51 | 0.01 | −0.01 |
| $H_b$ | Te | −55.69 | 1.84 | 9.82 | 226.45 | −1.00 | 8.95 |
| H | Te | −64.05 | −0.02 | 1.44 | 214.87 | −0.53 | −2.61 |
| $H_b$ | H | −0.94 | −0.02 | −0.06 | 0.20 | 0.00 | 0.00 |
| $H_b$ | Po | −439.33 | 1.02 | −0.87 | 343.72 | 1.40 | 34.68 |
| H | Po | −435.41 | 0.40 | 4.12 | 308.77 | −0.49 | 0.87 |
| $H_b$ | H | −0.80 | −0.02 | −0.07 | 0.78 | 0.00 | 0.08 |

the change in basis set from triple to quadruple zeta does not alter significantly the results for Se and Te and, even though there are more pronounced changes in isotropic and anisotropic constants involving Po, we observe the same trends and roughly the same values for $\Delta K^{FDE}$ and $\Delta K^{ME}$ as discussed above.

For shieldings a similar situation is found, with small variations in absolute (isotropic and anisotropic) shieldings between both triple-zeta and the QZ4P basis sets for Se and Te systems and significant changes for the Po ones. The consequence of these changes is that, whereas for Se and Te $\Delta\sigma^{FDE(4)}$ and $\Delta\sigma^{ME}$ change little, for Po the QZ4P set changes $\Delta\sigma_{iso}^{FDE(4)}$ and $\Delta\sigma_{aniso}^{FDE(4)}$ by 10 and 30 ppm, respectively, and $\Delta\sigma_{iso}^{ME}$ and $\Delta\sigma_{aniso}^{ME}$ by 30 and 60 ppm, respectively.

At this point, we cannot establish whether these strong variations in the shieldings only arise from deficiencies in the triple zeta sets or due to other factors. Doing so would require a thorough investigation that, we believe, is beyond the scope of this paper.

## 5 Conclusions and outlook

In this paper, we have described the implementation of frozen density embedding contributions in a response theory framework, in combination with the four-component DC Hamiltonian, to NMR indirect spin–spin couplings, NMR shieldings, and magnetizabilities for mean-field approaches (DFT-in-DFT or HF-in-DFT).

Due to the use of LAOs, which introduce the dependence of the electron- and spin-density on the external magnetic field in the case of NMR shieldings and magnetizabilities, additional embedding contributions to the property gradient (both properties) and expectation value (magnetizability only) arise, both for the individual subsystems as well as introducing a coupling between these.

By performing DFT calculations on $H_2X$–$H_2O$ (X = Se, Te, Po) model systems, we have been able to show the relative importance of these additional contributions to the properties in question, while at the same time confirming the findings of other studies that frozen density embedding is well suited to the calculation of NMR indirect spin–spin couplings and NMR shieldings.

We have observed that the inclusion of the embedding potential in the FDE-LAO property gradient contributions accounts for the bulk of the environment effects, and that the heavier the center the more intra-subsystem FDE-LAO kernel contributions are important for both NMR shieldings and magnetizabilities, due to the increasing importance of spin-density contributions. Coupling kernel LAO contributions, by contrast, are in general rather small.

We have exploited the use of magnetically induced currents to obtain the NMR shielding tensor *via* a real-space approach as well as to analyse, for the first time, the differences between supermolecular and embedded calculations in complement to the analysis of the electron density employed so far. We consider that the property density plots provide a much clearer picture of where in space the deficiencies in the FDE treatment manifest themselves compared to the scalar values of property shifts or unperturbed electron density plots as done prior to this work.

We present for the first time the FDE contributions to magnetizabilities. Unlike the case of the electric polarizability and in line with Pascal's rule, it appears that one can reconstruct with rather good accuracy the tensor for the supermolecular system from the tensors of the individual subsystems, obtained without the FDE coupling terms in the response equations. This may potentially make FDE a more reliable route to obtaining molecular magnetizabilities than other embedding approaches, since the whole system is treated quantum-mechanically.

We have also compared our results to those obtained using the spin–orbit ZORA Hamiltonian. Although the latter performs well, for the $PoH_2$–$H_2O$ system we have observed significantly different results between the Hamiltonians in the description of environment effects on the NMR shieldings and spin–spin coupling constants. Though part of these differences appear to be due to basis set effects, they nevertheless contrast with the common expectation that relative shieldings are largely insensitive to changes in Hamiltonians and therefore it would be worthwhile to verify whether that is indeed the case for other systems containing the heaviest elements.

## Appendix: working expressions for FDE-LAO contributions

In the DIRAC software, the quaternion algebra[132] is employed. Thus, $\Omega_{pq;0}$ and $i\Omega_{pq,\mu}$ are calculated from the real and imaginary parts, respectively, of the generally quaternion overlap distribution matrix (eqn (12)), which makes it easy to discern the charge- and spin-density contributions to the KS matrix and its derivatives.

Here, we present the working formulas for FDE-LAO contributions to the property gradient (eqn (43)–(45)) and to the expectation value part of the magnetizability tensor (eqn (49) and (50)) derived for a closed-shell reference. They were obtained by first separating the number-density ($n$: $\rho_0 = -en$) and spin-density ($s = \sqrt{\rho_\mu \cdot \rho_\mu}$, $\mu \in \{x, y, z\}$) contributions, then by applying the local ansatz, in which XC and kinetic energies are approximated by functions of local density variables:[78]

$$E_{xck} = \int \varepsilon_{xck}\left(Q^I \cup Q^{II}\right)\mathrm{d}\vec{r}, \tag{70}$$

with $Q^M = \{n^M, s^M, (\nabla n{\cdot}\nabla n)^M, (\nabla n{\cdot}\nabla s)^M, (\nabla s{\cdot}\nabla s)^M\}$ for $M \in \{I, II\}$. This allows us to express the FDE-LAO contributions to the property gradient of subsystem $I$ (the expression for subsystem $II$ can be obtained by exchanging the labels $I$ and $II$) and to the expectation value part of the magnetizability tensor in terms of scalar and vector pre-factors ($a_0$, $\vec{b}_0$, $c_0^{M,N}$, $c_\mu^{M,N}$, $\vec{d}_0^{M,N}$, $\vec{d}_\mu^{M,N}$ for $M, N \in \{I, II\}$), summarized in Tables 7 and 8, while various perturbed densities are outlined in Table 6 and discussed at length in the literature.[40,78,133]

$$\begin{aligned}\frac{\partial}{\partial B_\alpha}\frac{\partial E_{\mathrm{int}}}{\partial \kappa_{ai}^I}\bigg|_0 = &-\int\left[a_0^I \breve{\Omega}_{ia;0}^{B_\alpha;I} + \vec{b}_0^I\cdot\nabla\breve{\Omega}_{ia;0}^{B_\alpha;I}\right]\mathrm{d}\vec{r}\\ &-\iint\Bigg[c_0^{I,I}\Omega_{ia;0}^I + \vec{d}_0^{I,I}\cdot\nabla\Omega_{ia;0}^I + c_\mu^{I,I}\Omega_{ia;\mu}^I\cdot\\ &+\sum_{\mu=x,y,z}\vec{d}_\mu^{I,I}\cdot\nabla\Omega_{ia;\mu}^I\Bigg]\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2\\ &-\iint\Bigg[c_0^{I,II}\Omega_{ia;0}^I + \vec{d}_0^{I,II}\cdot\nabla\Omega_{ia;0}^I\cdot\\ &+c_\mu^{I,II}\Omega_{ia;\mu}^I + \sum_{\mu=x,y,z}\vec{d}_\mu^{I,II}\cdot\nabla\Omega_{ia;\mu}^I\Bigg]\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2\end{aligned} \tag{71}$$

$$\begin{aligned}\frac{\partial^2 E_{\mathrm{int}}}{\partial B_\alpha\partial B_\beta}\bigg|_0 = &\sum_{M=I,II}\Bigg\{\int\left[a_0^M\breve{\Omega}_{ii;0}^{B_\alpha B_\beta;M} + \vec{b}_0^M\cdot\nabla\breve{\Omega}_{ii;0}^{B_\alpha B_\beta;M}\right]\mathrm{d}\vec{r}\\ &+\iint\Bigg[c_0^{M,M}\breve{\Omega}_{ii;0}^{B_\beta;M} + \vec{d}_0^{M,M}\cdot\nabla\breve{\Omega}_{ii;0}^{B_\beta;M} + c_\mu^{M,M}\breve{\Omega}_{ii;\mu}^{B_\beta;M}\cdot\\ &+\sum_{\mu=x,y,z}\vec{d}_\mu^{M,M}\cdot\nabla\breve{\Omega}_{ii;\mu}^{B_\beta;M}\Bigg]\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2\Bigg\}\\ &+\sum_{\substack{M,N=I,II\\M\neq N}}\iint\Bigg[c_0^{M,N}\breve{\Omega}_{ii;0}^{B_\beta;M} + \vec{d}_0^{M,N}\cdot\nabla\breve{\Omega}_{ii;0}^{B_\beta;M}\cdot\\ &+c_\mu^{M,N}\breve{\Omega}_{ii;\mu}^{B_\beta;M} + \sum_{\mu=x,y,z}\vec{d}_\mu^{M,N}\cdot\nabla\breve{\Omega}_{ii;\mu}^{B_\beta;M}\Bigg]\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2.\end{aligned} \tag{72}$$

Table 6 The derivatives of the general density component, $\rho_k$, in the OMO basis. $\mathscr{P}_{\alpha\beta}$ denotes the permutation over indices $\alpha$ and $\beta$, $k \in \{0, x, y, z\}$. Basis functions labeled by $\mu$ and $\nu$ are centered on nuclei $K$ and $L$, respectively. Subsystem indices are skipped. Second-order derivatives of overlap distribution involve the derivatives of LAO overlaps ($\Omega_{ii;k}^{B_\alpha}$ and $\Omega_{ii;k}^{B_\alpha B_\beta}$) and the brace notation of Helgaker and Jørgensen[134] and are discussed in detail elsewhere[40]

$$\frac{\partial\rho_k}{\partial\kappa_{pq}}\bigg|_{\kappa=0} = -\Omega_{qp;k} \qquad \frac{\partial\rho_k}{\partial B_\alpha}\bigg|_{\vec{B}=0} = \breve{\Omega}_{ji;k}^{B_\alpha}$$

$$\frac{\partial^2\rho_k}{\partial\kappa_{pq}\partial B_\alpha}\bigg|_{\vec{B}=0} = -\breve{\Omega}_{qp;k}^{B_\alpha} \qquad \frac{\partial^2\rho_k}{\partial B_\alpha\partial B_\beta}\bigg|_{\vec{B}=0} = \breve{\Omega}_{ji;k}^{B_\alpha B_\beta}$$

$$\breve{\Omega}_{pq;k}^{B_\alpha} = \frac{\mathrm{i}e}{2}\left(\vec{R}_{KL}\times\vec{r}\right)_\alpha c_{\mu p}{}^* c_{\nu q}\Omega_{\mu\nu;k} + \Omega_{\mu\nu;k}\left\{c_{\mu t}{}^* T_{tp}^{B_\alpha *} c_{\nu q} + c_{\mu p}{}^* c_{\nu t} T_{tq}^{B_\alpha}\right\}$$

$$\breve{\Omega}_{ii;k}^{B_\alpha B_\beta} = \Omega_{ii;k}^{B_\alpha B_\beta} + \mathscr{P}_{\alpha\beta}\left\{T^\alpha, \Omega^{B_\beta}\right\}_{ii} + \left\{T^{B_\alpha B_\beta}, \Omega\right\}_{ii} + \frac{1}{2}\mathscr{P}_{\alpha\beta}\left(\left\{T^{B_\alpha}, \left\{T^{B_\beta}, \Omega\right\}\right\}_{ii} - \left\{T^{B_\beta}T^{B_\alpha}, \Omega\right\}_{ii;k}\right)$$

Table 7 Scalar prefactors derived for a closed-shell reference $a_0^I = v_{\mathrm{emb};0}^I$

$$\vec{b}_0^I = 2\left(\frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}}\nabla n^{\mathrm{tot}} - \frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)}\bigg|_I\nabla n^I\right)$$

$$\begin{aligned}c_0^{I,I} = &\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n^2}\bigg|_{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n^2}\bigg|_I\right)\breve{\Omega}_{jj;0}^{I;B_\alpha}\\ &+2\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n\partial(\nabla n\cdot\nabla n)}\bigg|_I\cdot\nabla n^I\right)\cdot\nabla\breve{\Omega}_{jj;0}^{I;B_\alpha}\end{aligned}$$

$$\begin{aligned}\vec{d}_0^{I,I} = &2\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n\partial(\nabla n\cdot\nabla n)}\bigg|_I\cdot\nabla n^I\right)\cdot\breve{\Omega}_{jj;0}^{I;B_\alpha}\\ &+4\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)^2}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)^2}\bigg|_I\cdot\nabla n^I\cdot\nabla n^I\right)\cdot\nabla\breve{\Omega}_{jj;0}^{I;B_\alpha}\\ &+2\left(\frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}} - \frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)}\bigg|_I\right)\cdot\nabla\breve{\Omega}_{jj;0}^{I;B_\alpha}\end{aligned}$$

$$\begin{aligned}c_0^{I,II} = &\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n^2}\bigg|_{\mathrm{tot}}\breve{\Omega}_{jj;0}^{II;B_\alpha} + \int\frac{1}{|\vec{r}-\vec{r}_1|}\breve{\Omega}_{jj;0}^{II;B_\alpha}(\vec{r}_1)\mathrm{d}\vec{r}_1\\ &+2\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla\breve{\Omega}_{jj;0}^{II;B_\alpha}\end{aligned}$$

$$\begin{aligned}\vec{d}_0^{I,II} = &2\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial n\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\breve{\Omega}_{jj;0}^{II;B_\alpha}\\ &+4\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)^2}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla\breve{\Omega}_{jj;0}^{II;B_\alpha} + 2\frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla n)}\bigg|_{\mathrm{tot}}\cdot\nabla\breve{\Omega}_{jj;0}^{II;B_\alpha}\end{aligned}$$

Table 8 Vector prefactors derived for a closed-shell reference

$$\begin{aligned}c_\mu^{I,I} = &\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s^2}\bigg|_{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s^2}\bigg|_I\right)\breve{\Omega}_{jj;\mu}^{I;B_\alpha}\\ &+\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s\partial(\nabla n\cdot\nabla s)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s\partial(\nabla n\cdot\nabla s)}\bigg|_I\cdot\nabla n^I\right)\cdot\nabla\breve{\Omega}_{jj;\mu}^{I;B_\alpha}\end{aligned}$$

$$\begin{aligned}\vec{d}_\mu^{I,I} = &\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s\partial(\nabla n\cdot\nabla s)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s\partial(\nabla n\cdot\nabla s)}\bigg|_I\cdot\nabla n^I\right)\cdot\breve{\Omega}_{jj;\mu}^{I;B_\alpha}\\ &+\left(\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla s)^2}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}} - \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla s)^2}\bigg|_I\cdot\nabla n^I\cdot\nabla n^I\right)\cdot\nabla\breve{\Omega}_{jj;\mu}^{I;B_\alpha}\\ &+2\left(\frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla s\cdot\nabla s)}\bigg|_{\mathrm{tot}} - \frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla s\cdot\nabla s)}\bigg|_I\right)\cdot\nabla\breve{\Omega}_{jj;\mu}^{I;B_\alpha}\end{aligned}$$

$$\begin{aligned}c_\mu^{I,II} = &\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s^2}\bigg|_{\mathrm{tot}}\breve{\Omega}_{jj;\mu}^{II;B_\alpha} + \frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s\partial(\nabla n\cdot\nabla s)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla\breve{\Omega}_{jj;\mu}^{II;B_\alpha}\end{aligned}$$

$$\begin{aligned}\vec{d}_\mu^{I,II} = &\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial s\partial(\nabla n\cdot\nabla s)}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\breve{\Omega}_{jj;\mu}^{II;B_\alpha}\\ &+\frac{\partial^2\varepsilon_{\mathrm{xck}}}{\partial(\nabla n\cdot\nabla s)^2}\bigg|_{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla n^{\mathrm{tot}}\cdot\nabla\breve{\Omega}_{jj;\mu}^{II;B_\alpha} + 2\frac{\partial\varepsilon_{\mathrm{xck}}}{\partial(\nabla s\cdot\nabla s)}\bigg|_{\mathrm{tot}}\cdot\nabla\breve{\Omega}_{jj;\mu}^{II;B_\alpha}\end{aligned}$$

## Acknowledgements

We would like to acknowledge inspiring discussions with Kenneth Ruud on the calculation of magnetizabilities, and with Christoph Jacob on the coupled response in FDE-NMR. We also acknowledge illuminating discussions with Michal Repisky on the question of magnetic balance in four-component calculations. The members of the PhLAM laboratory acknowledge support from the

CaPPA project (Chemical and Physical Properties of the Atmosphere), funded by the French National Research Agency (ANR) through the PIA (Programme d'Investissement d'Avenir) under contract ''ANR-11-LABX-0005-01'' as well as by the Ministry of Higher Education and Research, Hauts de France council and European Regional Development Fund (ERDF) through the Contrat de Projets Etat-Region (CPER) CLIMBIO (Changement climatique, dynamique de l'atmosphère, impacts sur la biodiversité et la santé humaine). Furthermore, ASPG acknowledges funding from the CNRS Institute of Physics (INP) *via* the PICS program (grant 6386), and computational time provided by the French national supercomputing facilities (grants DARI t2015081859, x2016081859).



## References

1 E. Oldfield, *Annu. Rev. Phys. Chem.*, 2002, **53**, 349.
2 A. Frank, H. M. Möller and T. E. Exner, *J. Chem. Theory Comput.*, 2012, **8**, 1480.
3 J. Bargon and L. T. Kuhn, *In situ NMR Methods in Catalysis*, Springer-Verlag Berlin Heidelberg, 2007, vol. 276.
4 G. Pintacuda and G. Kervern, in *Modern NMR Methodology*, ed. H. Heise and S. Matthews, Springer Berlin Heidelberg, 2013, vol. 335, p. 157.
5 J. Autschbach, in *Annu. Rep. Comput. Chem.*, ed. D. A. Dixon, Elsevier, 2015, vol. 11, p. 3.
6 H. Yasuoka, G. Koutroulakis, H. Chudo, S. Richmond, D. K. Veirs, A. I. Smith, E. D. Bauer, J. D. Thompson, G. D. Jarvinen and D. L. Clark, *Science*, 2012, **336**, 901.
7 L. Martel, J. Somers, C. Berkmann, F. Koepp, A. Rothermel, O. Pauvert, C. Selfslag and I. Farnan, *Rev. Sci. Instrum.*, 2013, **84**, 055112.
8 T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen and K. Ruud, *Chem. Rev.*, 2012, **112**, 543.
9 T. Helgaker, M. Jaszuński and K. Ruud, *Chem. Rev.*, 1999, **99**, 293.
10 H. M. McConnell, *J. Chem. Phys.*, 1957, **27**, 226.
11 D. Sitkoff and D. A. Case, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1998, **32**, 165.
12 J. C. Facelli, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 176.
13 K. Jackowski and M. Jaszuński, *Gas Phase NMR*, The Royal Society of Chemistry, 2016.
14 M. Kaupp, M. Buhl and V. G. Malkin, *Calculation of NMR and EPR Parameters: Theory and Applications*, Wiley-VCH, 2004.
15 F. London, *J. Phys. Radium*, 1937, **8**, 397.
16 J. Autschbach, *Philos. Trans. R. Soc., A*, 2014, **372**, 20120489.
17 M. Repisky, S. Komorovsky, R. Bast and K. Ruud, *Gas Phase NMR*, The Royal Society of Chemistry, 2016, p. 267.
18 T. Yoshizawa and M. Hada, *J. Comput. Chem.*, 2009, **30**, 2550.
19 P. Schwerdtfeger, B. Assadollahzadeh, U. Rohrmann, R. Schäfer and J. R. Cheeseman, *J. Chem. Phys.*, 2011, **134**, 204102.
20 P. Pyykkö, A. Görling and N. Rösch, *Mol. Phys.*, 1987, **61**, 195.
21 J. Autschbach, *Mol. Phys.*, 2013, **111**, 2544.
22 J. Vicha, J. Novotny, M. Straka, M. Repisky, K. Ruud, S. Komorovsky and R. Marek, *Phys. Chem. Chem. Phys.*, 2015, **17**, 24944.
23 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117.
24 A. Klamt and G. Schuurmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799.
25 C. C. Pye and T. Ziegler, *Theor. Chem. Acc.*, 1999, **101**, 396.
26 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669.
27 A. S. P. Gomes and C. R. Jacob, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2012, **108**, 222.
28 C. R. Jacob and J. Neugebauer, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 325.
29 T. A. Wesolowski, S. Shedge and X. Zhou, *Chem. Rev.*, 2015, **115**, 5891.
30 J. Neugebauer, *J. Chem. Phys.*, 2007, **126**, 134116.
31 J. Neugebauer, *J. Chem. Phys.*, 2009, **131**, 084104.
32 S. Höfener, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2012, **136**, 044104.
33 S. Höfener, A. S. P. Gomes and L. Visscher, *J. Chem. Phys.*, 2013, **139**, 104106.
34 S. Höfener and L. Visscher, *J. Chem. Phys.*, 2012, **137**, 204120.
35 A. S. P. Gomes, C. R. Jacob and L. Visscher, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5353.
36 A. S. P. Gomes, C. R. Jacob, F. Real, L. Visscher and V. Vallet, *Phys. Chem. Chem. Phys.*, 2013, **15**, 15153.
37 C. R. Jacob and L. Visscher, *J. Chem. Phys.*, 2006, **125**, 194104.
38 R. E. Bulo, C. R. Jacob and L. Visscher, *J. Phys. Chem. A*, 2008, **112**, 2640.
39 A. W. Götz, J. Autschbach and L. Visscher, *J. Chem. Phys.*, 2014, **140**, 104107.
40 M. Olejniczak, R. Bast, T. Saue and M. Pecul, *J. Chem. Phys.*, 2012, **136**, 014108.
41 A. Antusek and M. Sulka, *Chem. Phys. Lett.*, 2016, **660**, 127.
42 P. Lantto, K. Jackowski, W. Makulski, M. Olejniczak and M. Jaszunski, *J. Phys. Chem. A*, 2011, **115**, 10617.
43 S. Komorovsky, M. Repisky, E. Malkin, K. Ruud and J. Gauss, *J. Chem. Phys.*, 2015, **142**, 091102.
44 E. Malkin, S. Komorovsky, M. Repisky, T. B. Demissie and K. Ruud, *J. Phys. Chem. Lett.*, 2013, **4**, 459.
45 T. B. Demissie, M. Jaszunski, S. Komorovsky, M. Repisky and K. Ruud, *J. Chem. Phys.*, 2015, **143**, 164311.
46 G. A. Aucar, T. Saue, L. Visscher and H. J. A. Jensen, *J. Chem. Phys.*, 1999, **110**, 6208.
47 M. Iliaš, T. Saue, T. Enevoldsen and H. J. A. Jensen, *J. Chem. Phys.*, 2009, **131**, 124119.
48 M. Iliaš, H. J. A. Jensen, R. Bast and T. Saue, *Mol. Phys.*, 2013, **111**, 1373.
49 L. Visscher, T. Enevoldsen, T. Saue, H. J. A. Jensen and J. Oddershede, *J. Comput. Chem.*, 1999, **20**, 1262.
50 T. Enevoldsen, L. Visscher, T. Saue, H. J. A. Jensen and J. Oddershede, *J. Chem. Phys.*, 2000, **112**, 3493.
51 R. Bast, T. Saue, L. Visscher and H. J. Aa. Jensen, with contributions from V. Bakken, K. G. Dyall, S. Dubillard, U. Ekstroem, E. Eliav, T. Enevoldsen, E. Fasshauer, T. Fleig,

O. Fossgaard, A. S. P. Gomes, T. Helgaker, J. Henriksson, M. Ilias, Ch. R. Jacob, S. Knecht, S. Komorovsky, O. Kullie, J. K. Laerdahl, C. V. Larsen, Y. S. Lee, H. S. Nataraj, M. K. Nayak, P. Norman, G. Olejniczak, J. Olsen, Y. C. Park, J. K. Pedersen, M. Pernpointner, R. Di Remigio, K. Ruud, P. Salek, B. Schimmelpfennig, J. Sikkema, A. J. Thorvaldsen, J. Thyssen, J. van Stralen, S. Villaume, O. Visser, T. Winther and S. Yamamoto, *DIRAC, a relativistic ab initio electronic structure program, Release DIRAC15*, 2015, see http://www.diracprogram.org.

52 D. H. Whiffen, *Pure Appl. Chem.*, 1978, **50**, 75.
53 T. Saue, *ChemPhysChem*, 2011, **12**, 3077.
54 K. G. Dyall and K. Faegri Jr, *Introduction to Relativistic Quantum Chemistry*, Oxford University Press, 2007.
55 W. Kutzelnigg, *J. Comput. Chem.*, 1999, **20**, 1199.
56 W. Kutzelnigg, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2003, **67**, 032109.
57 L. Visscher, *Adv. Quantum Chem.*, 2005, **48**, 369.
58 S. Komorovský, M. Repiský, O. L. Malkina, V. G. Malkin, I. M. Ondk and M. Kaupp, *J. Chem. Phys.*, 2008, **128**, 104101.
59 M. Repisky, S. Komorovsky, O. L. Malkina and V. G. Malkin, *Chem. Phys.*, 2009, **356**, 236.
60 S. Komorovsky, M. Repisky, O. L. Malkina and V. G. Malkin, *J. Chem. Phys.*, 2010, **132**, 154101.
61 L. Cheng, Y. Xiao and W. Liu, *J. Chem. Phys.*, 2009, **131**, 244113.
62 R. D. Reynolds and T. Shiozaki, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14280.
63 J. Olsen, K. L. Bak, K. Ruud, T. Helgaker and P. Jørgensen, *Theor. Chem. Acc.*, 1995, **90**, 421.
64 U. von Barth and L. Hedin, *J. Phys. C: Solid State Phys.*, 1972, **5**, 1629.
65 A. K. Rajagopal and J. Callaway, *Phys. Rev. B: Solid State*, 1973, **7**, 1912.
66 C. R. Jacob and M. Reiher, *Int. J. Quantum Chem.*, 2012, **112**, 3661.
67 G. Vignale and M. Rasolt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 10685.
68 G. Vignale, M. Rasolt and D. J. W. Geldart, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 2502.
69 G. Vignale and M. Rasolt, *Phys. Rev. Lett.*, 1987, **59**, 2360.
70 A. M. Lee, N. C. Handy and S. M. Colwell, *J. Chem. Phys.*, 1995, **103**, 10095.
71 A. M. Lee, S. M. Colwell and N. C. Handy, *Chem. Phys. Lett.*, 1994, **229**, 225.
72 E. I. Tellgren, A. M. Teale, J. W. Furness, K. K. Lange, U. Ekström and T. Helgaker, *J. Chem. Phys.*, 2014, **140**, 034101.
73 J. W. Furness, J. Verbeke, E. I. Tellgren, S. Stopkowicz, U. Ekström, T. Helgaker and A. M. Teale, *J. Chem. Theory Comput.*, 2015, **11**, 4169.
74 O. B. Lutnaes, A. M. Teale, T. Helgaker, D. J. Tozer, K. Ruud and J. Gauss, *J. Chem. Phys.*, 2009, **131**, 144104.
75 A. M. Teale, O. B. Lutnaes, T. Helgaker, D. J. Tozer and J. Gauss, *J. Chem. Phys.*, 2013, **138**, 024111.
76 E. Engel, in *Relativistic Electronic Structure Theory*, ed. P. Schwerdtfeger, Elsevier, 2002, vol. 11, p. 523.
77 E. Engel and R. M. Dreizler, *Density Functional Theory. An Advanced Course*, Springer Berlin Heidelberg, 2011.
78 R. Bast, H. J. A. Jensen and T. Saue, *Int. J. Quantum Chem.*, 2009, **109**, 2091.
79 T. Saue and H. J. A. Jensen, *J. Chem. Phys.*, 2003, **118**, 522.
80 T. A. Wesolowski and J. Weber, *Recent Advances in Density Functional Methods*, World Scientific, 2002, ch. 24, p. 371.
81 A. Solovyeva, M. Pavanello and J. Neugebauer, *J. Chem. Phys.*, 2012, **136**, 194104.
82 T. A. Wesolowski and A. Warshel, *J. Phys. Chem.*, 1993, **97**, 8050.
83 C. R. Jacob, J. Neugebauer and L. Visscher, *J. Comput. Chem.*, 2008, **29**, 1011.
84 D. Schlüns, K. Klahr, C. Mück-Lichtenfeld, L. Visscher and J. Neugebauer, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14323.
85 S. Fux, C. R. Jacob, J. Neugebauer, L. Visscher and M. Reiher, *J. Chem. Phys.*, 2010, **132**, 164101.
86 D. G. Artiukhin, C. R. Jacob and J. Neugebauer, *J. Chem. Phys.*, 2015, **142**, 234101.
87 J. P. Unsleber, J. Neugebauer and C. R. Jacob, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21001.
88 S. Höfener and L. Visscher, *J. Chem. Theory Comput.*, 2016, **12**, 549.
89 M. Pavanello, *J. Chem. Phys.*, 2013, **138**, 204118.
90 J. Neugebauer, C. R. Jacob, T. A. Wesolowski and E. J. Baerends, *J. Phys. Chem. A*, 2005, **109**, 7805.
91 J. Neugebauer and E. J. Baerends, *J. Phys. Chem. A*, 2006, **110**, 8786.
92 J. Neugebauer, M. J. Louwerse, E. J. Baerends and T. A. Wesolowski, *J. Chem. Phys.*, 2005, **122**, 094115.
93 R. Bast, J. Jusélius and T. Saue, *Chem. Phys.*, 2009, **356**, 187–194.
94 C. J. Jameson and A. D. Buckingham, *J. Chem. Phys.*, 1980, **73**, 5684.
95 D. Sulzer, M. Olejniczak, R. Bast and T. Saue, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20682.
96 ADF2014, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, http://www.scm.com.
97 E. V. Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1993, **99**, 4597.
98 E. van Lenthe, E. J. Baerends and J. G. Snijders, *J. Chem. Phys.*, 1994, **101**, 9783.
99 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
100 E. Van Lenthe and E. J. Baerends, *J. Comput. Chem.*, 2003, **24**, 1142.
101 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
102 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1997, **78**, 1396.
103 A. Lembarki and H. Chermette, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1994, **50**, 5328.
104 U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen and K. Ruud, *J. Chem. Theory Comput.*, 2010, **6**, 1971.
105 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.
106 G. K. Dyall, *Theor. Chem. Acc.*, 2002, **108**, 335.
107 K. G. Dyall, *Theor. Chem. Acc.*, 2006, **115**, 441.
108 L. Visscher and K. G. Dyall, *At. Data Nucl. Data Tables*, 1997, **67**, 207.

109 L. Visscher, *Theor. Chem. Acc.*, 1997, **98**, 68.
110 J. Mason, *Solid State Nucl. Magn. Reson.*, 1993, **2**, 285.
111 P. Ramachandran and G. Varoquaux, *Comput. Sci. Eng.*, 2011, **13**, 40.
112 E. Arunan, G. Desiraju, R. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci and D. J. Nesbitt, *Pure Appl. Chem.*, 2011, **83**, 1637.
113 N. P. Luthra and J. D. Odom, *Nuclear magnetic resonance and electron spin resonance studies of organic selenium and tellurium compounds*, John Wiley & Sons, Inc., 2010, p. 189.
114 T. Kemp, A. Wong, M. Smith, P. Bishop and N. Carthey, *Solid State Nucl. Magn. Reson.*, 2008, **34**, 224.
115 H. Duddeck, $^{77}$Se NMR Spectroscopy and Its Applications in Chemistry, *Annu. Rep. NMR Spectrosc.*, Academic Press, 2004, vol. 52, p. 105.
116 Y. Y. Rusakov, I. L. Rusakova and L. B. Krivdin, *Magn. Reson. Chem.*, 2015, **53**, 485.
117 Y. Y. Rusakov and L. B. Krivdin, *J. Comput. Chem.*, 2015, **36**, 1756.
118 S. Hayashi, K. Matsuiwa and W. Nakanishi, *RSC Adv.*, 2014, **4**, 44795.
119 http://dx.doi.org/10.5281/zenodo.179667, figures available under a CC-BY 4.0 license.
120 J. Autschbach, in *High Resolution NMR Spectroscopy. Understanding Molecules and their Electronic Structures*, ed. R. H. Contreras, Elsevier, 2013, vol. 3, p. 69.
121 K. Ruud, H. Skaane, T. Helgaker, K. L. Bak and P. Jørgensen, *J. Am. Chem. Soc.*, 1994, **116**, 10135.
122 P.-O. Astrand and K. V. Mikkelsen, *Magn. Reson. Chem.*, 1998, **36**, 92.
123 R. F. W. Bader and T. A. Keith, *J. Chem. Phys.*, 1993, **99**, 3683.
124 K. Ruud, P.-O. Astrand and P. R. Taylor, *J. Phys. Chem. A*, 2001, **105**, 9926.
125 P. Pascal, *Ann. Chim. Phys.*, 1910, **19**, 5.
126 P. Pascal, *Rev. Sci. Instrum.*, 1948, **86**, 38.
127 W. H. Flygare, *Chem. Rev.*, 1974, **74**, 653.
128 K. Ruud, private communication.
129 J. Autschbach and S. Zheng, in *Annu. Rep. NMR Spectrosc.*, ed. G. A. Webb, Academic Press, 2009, vol. 67, p. 1.
130 D. L. Bryce, R. E. Wasylishen, J. Autschbach and T. Ziegler, *J. Am. Chem. Soc.*, 2002, **124**, 4894.
131 W. Makulski, *THEOCHEM*, 2012, **1017**, 45.
132 T. Saue and H. J. A. Jensen, *J. Chem. Phys.*, 1999, **111**, 6211.
133 P. Sałek, T. Helgaker and T. Saue, *Chem. Phys.*, 2005, **311**, 187.
134 T. Helgaker and P. Jørgensen, *J. Chem. Phys.*, 1991, **95**, 2595.

# List of Figures

# List of Tables