University of Groningen

Master Thesis

Relativistic coupled cluster ionisation potential calculations for HgCl, HgI, and TaN

Accurate molecular calculations for eEDM experiment candidates

Author
Aiden Orion Strijker

Supervisors
Anastasia Borschevsky (first examiner)
Steven Hoekstra (second examiner)
Agustín Aucar (daily supervisor)

August 4, 2025

rijksuniversiteit
groningen

Relativistic coupled cluster ionisation potential calculations for HgCl, HgI, and TaN

Accurate molecular calculations for eEDM experiment candidates

Aiden Orion Strijker

Abstract

The ionisation potentials of three molecules, HgCl, HgI, and TaN, have been calculated using the relativistic Dirac-Hartree-Fock and coupled cluster methods. The computational methods are explored to determine the accuracy of the theoretical predictions, leading to corrections for higher order effects and the determination of a conservative uncertainty estimate. This includes basis set effects, level of relativity, and electron correlation. With this, the ionisation potentials obtained are $IP = 9.411 \pm 0.033$ eV, $IP = 8.768 \pm 0.032$ eV, and $IP = 8.485 \pm 0.039$ eV for HgCl, HgI, and TaN, respectively. For easier comparison to experimental results, the vibrationally corrected ionisation potentials have also been determined as $IP_{vc} = 9.446 \pm 0.0035$ eV (HgCl), $IP_{vc} = 8.770 \pm 0.035$ eV (HgI), and $IP_{vc} = 8.485 \pm 0.042$ eV (TaN). No experimental measurements are available for comparison, although the calculated equilibrium bond lengths are in agreement with experimental and earlier theoretical values.

Acknowledgements

I could not have done this project alone, and I would like to sincerely thank Prof. Dr. Anastasia Borschevsky and Prof. Dr. Steven Hoekstra for allowing me to work on my own tempo. Your patience and feedback has been instrumental in the success of this thesis. I am especially grateful for the great supervision and support I received from Prof. Dr. Anastasia Borschevsky, who was very involved and always available for questions in a way that greatly boosted my confidence. I don't think that I could have asked for better guidance.

Special thanks goes out to Dr. Agustín Aucar, who patiently helped me develop a basic understanding of much of the theory and had great advise for the many computational difficulties I encountered. Your encouragement and in-depth explanations were a tremendous help. Additionally, every other person in the AIM research group has been kind and informative, and I have learned much from being included in group meetings and social events. Even though I am most comfortable as an observer, I always felt welcome to ask questions and join discussions, which means a lot to me.

Thank you all, and thank you to the reader for reading this.

Contents

1 Introduction - the bigger picture

Nearing the end of the 19th century, some held the belief that physics was a mature science, without much left to discover[1]. A century-and-a-half later, this idea is thoroughly thrown out the window, as more advanced theories regularly emerge and our collective understanding of the universe and its inner workings grows. Continually, researchers target structures within structures, aiming to understand the patterns of the universe in the overlap of physics with other fields. This thesis fits in that picture, combining chemistry and physics to create a better understanding of molecular structures to be used in the improvement of particle physics models.

The rest of this introduction will explain the bigger picture that this master project is part of, zooming in towards the focus of this thesis. This also includes the research questions the project aims to answer in subsection 1.5. The theory behind the calculations performed is described in section 2. Following that, section 3 explains the methodology. In 4.1 relevant earlier research is addressed. The calculation results are reported and discussed in section 4.2. Finally, everything is summarised in 5.

1.1 The Standard Model and CP violation

The Standard Model (SM) is the current best model of particle physics, describing all (known) elementary particles and their interactions. It explains three of the four fundamental forces in the universe from a quantum mechanical perspective. Through extensive research the SM has also become a very robust theory, explaining many phenomena in the universe and predicting particles such as the Higgs boson [2].

Tests of discrete symmetries have been instrumental throughout the last century in making the SM as well-structured as it is now [3]. These symmetries are described by the actions;

$$C\phi(t,x) = \overline{\phi}(t,x), \tag{1}$$

$$P\phi(t,x) = \phi(t,-x), \tag{2}$$

$$T\phi(t,x) = \phi(-t,x). \tag{3}$$

Here, C is the charge symmetry (particle into anti-particle exchange, roughly), P stands for parity (reflection in spatial coordinates), and T for time reversal (reversing the time coordinates) [4] [5]. Together they form the CPT symmetry, which is invariant at the level of known fundamental physics due to Lorentz symmetry and causality [1].

While the SM is the most exhaustive theory to describe the laws of physics of the universe at the moment, it is by no means complete. There are a few observed phenomena that fail to be explained by the model. These include the missing connection to the fourth fundamental force of gravity, the uneven distribution of matter and anti-matter, the expansion of the universe and the elusive dark matter and dark energy [7]. To explain the matter-antimatter asymmetry, scientists look for sources of CP violation.

Within the Standard Model CP violation is possible in interactions governed by the weak force, but many of these interactions are heavily suppressed. The first experimentally measured case of direct CP violation was done by analysing the decay of strange mesons into kaons and pions, where a significant asymmetry in the decay process was found [8], though indirect CP-violation was already measured in 1964 through P and T violation in the decay of Kaons [9]. While many other CP

[1] For a more extensive and relatively accessible explanation of CPT invariance and its implications, see ref. [6]

violating sources exist, none known currently are abundant or strong enough to explain the matter-antimatter asymmetry. Thus the SM does not lead to the matter-antimatter asymmetry we observe in the universe, and an explanation might be found in analysis of new CP violating sources.

1.2 The eEDM

One signature of CP violation is the eEDM; the electron electronic dipole moment. This is a permanent charge separation on the axis of the electron-spin [10]. It violates parity (P) and time reversal (T) due to the opposite behaviour [2] of $\vec{E}\cdot\vec{S}$ and $\vec{B}\cdot\vec{S}$ under both these symmetry actions [3], as can be seen in Figure 1. Consequently, assuming the CPT theorem, the combined CP symmetry is also violated.

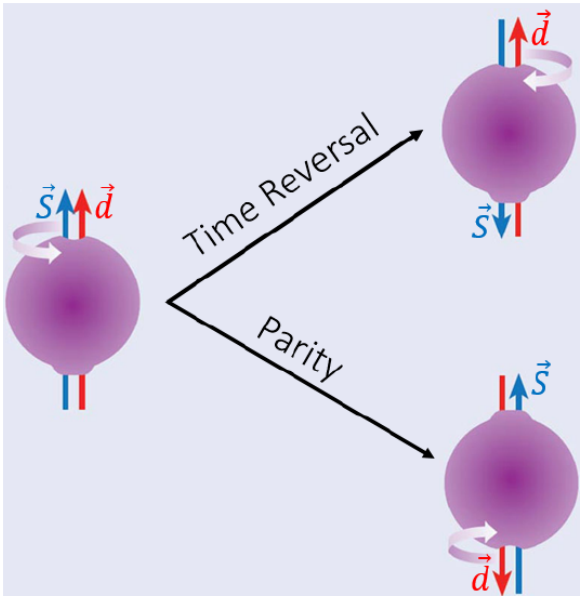
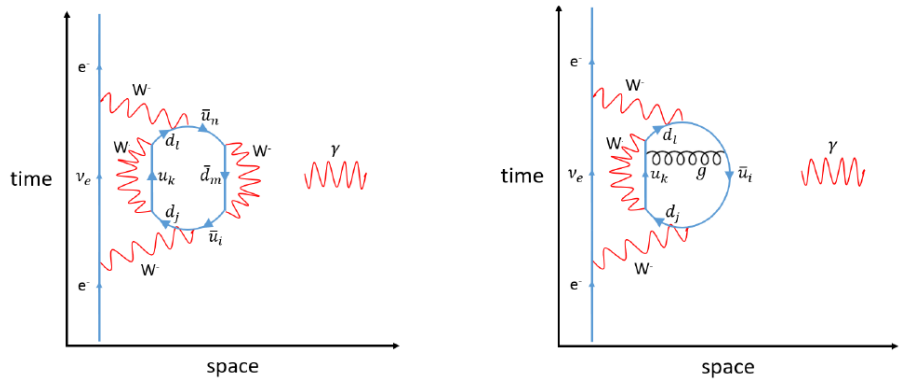


Figure 1: For a non-zero EDM $\vec{d}$, it must be parallel or antiparallel to spin $\vec{S}$. Time reversal would change the orientation of the spin, not the EDM, while parity inversion flips the orientation of the EDM, but not the spin. Whether we start with a parallel or antiparallel system, we end up with the opposite after either operation, meaning it violates the symmetries [11]

The eEDM does exist within the Standard Model, but is - as many other CP violating sources - heavily suppressed. In the language of Feynman diagrams, this is because it only occurs in higher loop order. The integrals can be described by a perturbative expansion have ever decreasing coefficients depending on the amount of loops in the diagram. [11]

[2] The electric $\vec{E}$ and magnetic $\vec{B}$ field dot-product with the spin $\vec{S}$



(a) SM eEDM, all weak interactions

(b) SM eEDM, one strong interaction

Figure 2: Four-loop Feynman diagrams with a non-zero eEDM [11]. The weak interaction (a) changes the quark flavour with W-bosons, and strong interaction (b) is characterised by the gluon g interaction between quarks.

As is visible in Figure 2, the eEDM appears at fourth order within the Standard Model. The prediction of the size is therefore also very small in the theory; about $d_e \sim 10^{-40}$ e cm [11], or having an upper limit around $|d_e| \leq 10^{-38}$ e cm [3]. Many other theories that try to include the matter-antimatter asymmetry predict a much higher value of the eEDM, that should be within measureable range of modern experiments [10]. With state-of-the-art high precision experiments and theoretical molecular knowledge, these predictions are in a measurable range.

1.3 eEDM in molecules

Motivated by the SM and Beyond the Standard Model (BSM) predictions, there are many experiments aimed at detecting the eEDM. Though the SM-upper limit is still quite a few orders of magnitude away, the current experimental limit on the eEDM lies at $|d_e| < 4.1 \cdot 10^{-30} e$cm (90% confidence interval) [12], from a measurement using HfF^+. This is an improvement on the previous upper limit - $|d_e| < 1.1 \cdot 10^{-29} e$cm from the ACME II measurement using ThO [13] - by a factor of about 2.4. The reason this range is reachable is because molecules experience an enhancement factor W_d that creates a measurable energy shift of $U = -\vec{d_e} \cdot (1/\Omega)\vec{W_d}$ [13][14][3]. While earlier experiments worked with atoms, current experiments use molecules, which give an enhancement of 3 to 4 orders of magnitude compared to atoms [7]. Specifically, W_d is dependent on the internal electric field for molecules and of the order of GV/cm, while in atoms the dependence is on an external electric field ∼kV/cm [16]. A sufficiently large W_d is required and the exact enhancement is different for each system. Because the factor is intrinsic for molecules, this needs to be theoretically calculated before the size of the eEDM can be extracted from experimental data.

The improvement of the most recent upper limit was also made possible due to the trapping of the ions [12]. In order to improve from here, theoretical knowledge about molecules is required to both propose good candidates for future experiments and provide W_d values for chosen candidates. Therefore, not only the large value of W_d is important, but relevant molecular aspects are found in

[3] In literature, the enhancement factor is also often called the effective electric field ($E_{eff} = \Omega W_d$, where the factor Ω is the projection of the angular momentum onto the internuclear axis [15]), but nowadays the use of W_d is more common

the statistical uncertainty relation of an eEDM experiment, given as

$$\sigma_d = \frac{\hbar}{e} \frac{1}{2|P|\Omega W_d \tau \sqrt{\dot{N} T}}, \tag{4}$$

where P is the polarisation, Ω the projection of the angular momentum onto the internuclear axis, τ the time that the molecules interact with the electric field, $\dot{N}$ the rate of detection, and T the time of measurement [17][10]. In the rest of this section, I will go over these factors, that all need to be as large as possible in order to lower the statistical uncertainty.

Starting with the enhancement factor; there is not a standard methodical way to find a system with a large enhancement factors, but there are a few hints to start the search. W_d is a relativistic property and it scales as Z^3 in atoms. The factor also benefits from higher Z in molecules [3][16], hence heavy systems are studied to find the eEDM [18]. In calculations on these molecules (or atoms), relativity needs to be taken into account due to the large mass differences between the nuclei and electrons and high velocities these electrons reach.

Polarisation is also a relevant factor for the eEDM measurement. The size of the σ_d is inversely proportional to the electric polarisation of a system, thus having the ability to polarise a molecule fully ($|P| \rightarrow 1$) is ideal to take full advantage of a large enhancement factor W_d [19]. The best systems are polarisable with highly controlled electric fields, in order to limit errors from electric and magnetic field noise [11]. Combined with the previous point, heavy polar molecules become most interesting. However, these systems come with increasing complexity as well; the bigger the molecule, the more computational resources required to make accurate calculations. The focus of this thesis will therefore be on diatomic molecules, though some larger systems have already been proposed as candidates (see for example [20], [21], or [22]).

Historically, the focus of research has been on molecules in a Σ state, as these tend to be the simplest systems, but they limit Ω. Δ states are being considered as well, especially in cases where the doublet states can help eliminate systematic errors due to a chance in sign for uncertainty causing effects [23]. Naively, it would be better to look at states with higher Ω, but the electronic structure gets much more complex and the electronic and vibrational energy levels become a saturated web that is hard to separate into different states and work with from both an experimental and theoretical perspective.

On top of that, there are logistical considerations to look out for in preparing future eEDM experiments, which could lead to an increase of interaction times, measurement times, and detection rates. One such logistical example is laser coolability; in order to trap molecules for measurements, they need to be slowed down, but suitable electronic transitions need to be available and understood well for laser cooling to be possible [24]. Being able to trap molecules leads to higher measurement times and longer lifetimes, and can also positively impact the detection rate.

In this thesis, my focus is on calculations of ionisation potentials (IPs) of different diatomic molecules. These could potentially be optically trapped at crygenic temperatures with a method proposed by Singh *et al* [25], making precise measurements with better τ, $\dot{N}$ and T possible. The ionisation potentials and other molecular properties need to be in a specific range in order for the molecule to be suitable for the trap. The calculated ionisation potentials can also be easily compared to experimentally measured IPs and other theoretical predictions in the future, in order to test their agreement and accuracy.

1.4 Molecules of interest

There are many molecules being used in eEDM research, such as BaF in the NL-eEDM collaboration [17], HfF^+ [12], and ThO [13]. For future research many more are being considered. Specifically for the optical trap of Singh *et al*, the most ideal molecules will have a high ionisation potential [25], but the IP of many diatomic systems is unknown. The molecules chosen for this project are mercury chloride (HgCl), mercury iodine (HgI), and tantalum nitride (TaN). I will introduce why these in particular are of interest in the search for the eEDM and for trapping experiments below.

Mercury atoms have been used in theoretical eEDM and other CP violating effects studies before [26][27][28][29], and mercury containing molecules were already proposed for molecular parity violating experiments before the turn of the century [30]. The first was mercury fluoride HgF (among other fluoride diatomics), which has been identified as a candidate for future research and the enhancement factor has been calculated [31][32][33]. It turned out W_d of HgF was much larger than that of other candidates like YbF and ThO, respectively about 5 times and 1.5 times larger [34][16], in spite of mercury being lighter than thorium. This is due to the enhanced relativistic effects in mercury which don't follow the typical behaviour of scaling as $\sim Z^2$ [35]. From HgF, the natural step was to look for diatomics with similar properties, namely mercury halides in general. They all have a $^2\Sigma$ ground state, can be produced easily and measured at high coherence times, and dissociate similarly leading to efficient detection [36][16]. This group includes HgF, HgCl, HgBr [4], and HgI. HgF and HgBr are left out of this project, but will be reconsidered in the future.

The connection of tantalum nitride with CP violation research is not as easily traceable to its metal atom like with the mercury halides, though both TaO^+ [38][39] and TaN [23][40] have been proposed as eEDM candidates recently. Along with HgF and a few other eEDM candidates, TaN has a nuclear magnetic quadrupole moment (MQM) due to its deformed nuclei and large nuclear spin I ($> 1/2$), which makes it interesting for CP violation research in general [41]. Specifically, the first excited $^3\Delta_1$ state of TaN is similar to that of ThO [41], which was used to set the previous eEDM limit [13]. On top of that, TaN is valence-isoelectric to both ThO and HfF^+ [40], the latter being the molecule used to set the current eEDM upper limit [12]. Thus, along the same lines of logic used for HgCl and HgI, tantalum nitride is a candidate for future eEDM research through interest in similar molecules.

With the knowledge that these three molecules are potential eEDM experimental candidates rises the need to understand them and their structures better. The ionisation potential is crucial for that, as it provides detailed information of the electronic structure [42]. The IP has not been determined for most molecules, while it is known to high precision for most atoms, only excepting the heaviest not naturally-occurring elements [43]. HgCl, HgI, and TaN are no exception, with barely any available data regarding their ionisation potentials. What is known from previous research will be discussed in section 4.1, but first, I will elaborate on the objective of my thesis.

1.5 Research focus

The goal of this thesis is to calculate the ionisation potentials of HgCl, HgI, and TaN, in order to determine if these molecules can be trapped in future experimental set-ups. This would be a small puzzle piece in the search for the eEDM and more generally in Standard Model-testing research. Besides that, ionisation potentials are a window into a deeper understanding of the electronic structure of molecules. The theoretical determination of IPs can provide a stepping stone for measuring

[4] HgBr has been looked into for eEDM research as well [36][37]

them in the future. The knowledge could also provide a helping hand to better understand trends in similar molecules and make predictions about heavier and more complex systems.

Now, having placed the goals of this thesis in the large perspective of fundamental physics research goals, I will continue this report by addressing the practical questions of my thesis. These are mostly aimed at addressing the accuracy of the theoretical calculations, including the level of theory and computational limits that can be taken into account or later corrected for, and the available information about the three chosen molecules.

2 Theory

In this section, a few of the main theories behind the molecular calculations are highlighted (section 2.1), including an explanation of the computational methods that were utilised (section 2.2). The equations in the rest of this thesis are expressed in atomic units [44], meaning $\hbar = m_e = e = 4\pi\epsilon_0 = 1$, unless otherwise specified.

2.1 Solving the Schrödinger equation

The theoretical understanding and description of atoms and molecules rests on the famous Schrödinger equation, describing the components of the (molecular) system and how they interact. In this section I will start with the general description of the Schrödinger equation and its relativistic counterpart, the Dirac equation (see section 2.1.1), after which I will introduce their respective Hamiltonians and the first approximation (section 2.1.2 and 2.1.3), and conclude with the first computationally motivated approximation (see section 2.1.4), leading into the next section 2.2.

2.1.1 Schrödinger and DiraC

Describing molecules with quantum chemistry must start with the time-independent Schrödinger equation [45];

$$H\psi = E\psi, \tag{5}$$

which is a second order differential equation, where H is a scalar Hamiltonian, ψ the wavefunction, and E the energy. The Hamiltonian will be discussed in detail in section 2.2, but for now its complexity depends on the number of electrons in the system. In the case of the hydrogen atom, analytical solutions can be found and there is a simple and well-known expression for the energy [45]. But, as soon as we introduce a second electron or nucleus, we are dealing with a many-body problem, which becomes complex and quickly (when adding even more electrons and nuclei) impossible to solve analytically. On top of that, for heavier systems, the Schrödinger equation is not the most accurate to work with, as relativistic effects become increasingly relevant. Instead, we switch to the Dirac equation, [45]

$$\mathbf{H}\Psi = E\Psi, \tag{6}$$

which is written similarly to the Schrödinger equation, but with the key difference that it is no longer scalar. Instead, $\mathbf{H}$ is a 4x4 matrix operator and Ψ is a vector. The 4 comes from quantum electrodynamics (QED), where spin-1/2 particles are described using 4x4 Dirac matrices [46]

$$\alpha = \begin{bmatrix} 0_2 & \vec{\sigma} \\ \vec{\sigma} & 0_2 \end{bmatrix}, \tag{7}$$

with $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ the traditional 2x2 Pauli spin matrices that are explicit components of the relativistic Hamiltonian. We also get a set of 2-spinor components [46] for the wavefunction

$$\Psi = \begin{pmatrix} \Psi^L \\ \Psi^s \end{pmatrix}, \tag{8}$$

where Ψ^L is upper block describing positive energy solutions (electrons), and Ψ^s is lower block [5] forming the negative energy solutions (positrons). However, before diving deeper into the representation of the wavefunction, let me first describe the Hamiltonian(s) in more detail.

2.1.2 The Born-Oppenheimer approximation

The full hamiltonian of a system in the Schrödinger equation (eq. 5) is

$$H = T_e(\vec{r}) + T_N(\vec{R}) + V_{eN}(\vec{r}, \vec{R}) + V_{ee}(\vec{r}) + V_{NN}(\vec{R}), \tag{9}$$

where T are the kinetic energies of the electrons e and nuclei N, and V are the potential energies including Coulombic repulsion of nuclei V_{NN} or electrons V_{ee}, and Coulombic attraction of electrons to the nuclei V_{eN} [47]. They are described by the sums;

$$T_e(\vec{r}) = -\frac{1}{2}\sum_i \nabla_i^2, \qquad T_N(\vec{R}) = -\sum_A \frac{1}{2M_A}\nabla_A^2,$$
$$V_{eN}(\vec{r}, \vec{R}) = -\sum_{A,i} \frac{Z_A}{r_{Ai}}, \quad V_{ee}(\vec{r}) = \sum_{i<j} \frac{1}{r_{ij}}, \quad V_{NN}(\vec{R}) = \sum_{A>B} \frac{Z_A Z_B}{R_{AB}}.$$

Here, the capital letters refer to the nuclei (A, B) and their respective masses M, atomic numbers Z and internuclear distances R, while all lowercase letters refer to the variables for the electrons [47]. The sums are not further specified, leaving this to be a general Hamiltonian for a molecule.

The first simplification to be made before the Schrödinger equation can be solved is the Born-Oppenheimer approximation, where the description of electrons and nucleons is split. We start from the full Hamiltonian (eq. 9) in the Schrödinger equation (eq. 5). By separating the solution into a electronic and nuclear wavefunction - making the total wavefunction Ψ a product of the two $\Psi_e \cdot \Psi_N$ - it is possible to find the Born-Oppenheimer approximation as the electronic Schrödinger equation [48]

$$H_e(\vec{R})\,|\Psi_e\rangle = \Big(T_e(\vec{r}) + V_{eN}(\vec{r};\vec{R}) + V_{ee}(\vec{r})\Big)\,|\Psi_e\rangle = E_e(\vec{R})\,|\Psi_e\rangle\,. \tag{10}$$

This separation is made because the electrons are lighter and move faster than the nuclei, meaning the nuclei have relatively small changes in position. The position of the nuclei $\vec{R}$ is thus a parameter in the electronic equation. The total molecular energy is found from the remaining nuclear part [47]

$$H_N\,|\Psi_N\rangle = \Big(T_N(\vec{R}) + V_{NN}(\vec{R}) + E_e\Big)\,|\Psi_N\rangle = E_{tot}\,|\Psi_N\rangle\,. \tag{11}$$

2.1.3 The relativistic Hamiltonian

In the relativistic picture, we separate the one-electron part of the electronic Hamiltonian from the two-electron operator as h and g [49];

$$H = V_{NN} + \sum_i h(i) + \frac{1}{2}\sum_{i\neq j} g(i,j). \tag{12}$$

[5] These components are also referred to as upper (Ψ^U) and lower (Ψ^L) in literature, but because the positive-energy terms tend to be bigger when working with matter, "larger" and "smaller" are the preferred terms here

The h contains the relativistic equivalent terms to T_e and V_{eN} and can be expressed [49] with the matrix

$$h_d = \begin{bmatrix} V_{eN} & c(\vec{\sigma} \cdot \vec{p}) \\ c(\vec{\sigma} \cdot \vec{p}) & V_{eN} - 2c^2 \end{bmatrix}, \tag{13}$$

where

$$V_{eN}(\vec{r}) = -\sum_A \int_{\mathbb{R}^3} \frac{\rho_A(\vec{r}')}{|\vec{r}' - \vec{r}|} \, d^3\vec{r}'; \tag{14}$$

$$\text{and} \int_{\mathbb{R}^3} \rho_A(\vec{r}) \, d^3\vec{r} = Z_A. \tag{15}$$

Here, $\vec{p}$ is the momentum, c the speed of light, and ρ_A the nuclear charge density. In the latter we see that the nuclei now no longer get treated as a point-like object, instead having a finite size. To this one-electron expression we add

$$g(1,2) = \frac{1}{r_{12}}, \tag{16}$$

which is the zeroth-order relativistic correction to the two-electron repulsion V_{ee} [46]. These expressions together in equation 12 form the Dirac-Coulomb Hamiltonian.

The full Breit operator is the first order correction term, that could be added to the 2-electron expression, giving

$$g(1,2) = \frac{1}{r_{12}} - \frac{\vec{\alpha}(1) \cdot \vec{\alpha}(2)}{r_{12}} + \frac{1}{2}\left[\frac{\vec{\alpha}(1) \cdot \vec{\alpha}(2)}{r_{12}} - \frac{(\vec{\alpha}(1) \cdot \vec{r}_{12})(\vec{\alpha}(2) \cdot \vec{r}_{12}}{r_{12}^3}\right] \tag{17}$$

with α the 4x4 Pauli matrix expression from equation 7. If this full expression is used, we get the Dirac-Coulomb-Breit (DCB) Hamiltonian, in which an approximation of quantum-electrodynamics is taken into account for the electron-electron interaction with the Breit term. The first part of the Breit term - the second term on the right-hand side of equation 17 - is called the Gaunt term and represents a magnetic interaction, and some methods are restricted to using only the Gaunt term, instead of the full Breit operator. The effect of the Breit term is small ($\mathcal{O}(c^{-2})$ correction [50]), but it's computational cost is high, so it was not included in this work.

2.1.4 X2C

While accurate, the relativistic method of using a four-component (4C) Dirac equation, DC or DCB Hamiltonian, and 4-spinor wave-function requires solving more complex integrals and performing larger calculations than the non-relativistic scheme. The level of relativistic accuracy in computational methods competes for resources with other levels of approximation and precision in the electron correlation and wavefunction descriptions. The full Breit term, for example, could not be taken into account in this thesis, though the DC 4C Hamiltonian and added Gaunt term are included in the corrections of the uncertainty estimation (see section 3.3, and results in section 4.2.3). Most of the calculations I have made, however, make use of the exact-2-component (X2C) method.

Because the upper and lower components of the wavefunction in equation 8 are not independent, but instead related through

$$\Psi^s = X\Psi^L, \text{ where } X = \frac{1}{2c}\left(1 - \frac{V - E}{2c^2}\right)^{-1} \vec{\sigma} \cdot \vec{p} \ , \tag{18}$$

some steps can be made to save computational resources [46]. Here, V is the external potential and E the energy shifted to the rest mass of the electron. Because of this relation, many approximate 2 component methods have been developed throughout the years to bridge the gap between non-relativistic and relativistic approaches. An important example is as the normalised elimination of the small component (NESC) method [51][52], which eventually lead to the X2C method, where the two-electron integration is solved in a 2C description and afterwards translated back into 4C [53][54][55]. The one-electron integrals are still solved in a 4C matrix system, and the two are combined for the full solution. The crucial element that makes X2C an "exact" method is matrix algebra; through diagonalisation and matrix transformations, the final energy calculations are equivalent to 4C calculations, but the complexity of electron interaction integrals is significantly lowered by treating them in the transformed, simpler shape. [46]

2.2 Computational methods and approximations

The Schrödinger and Dirac equations of even the smallest molecules cannot be solved analytically, thus approximations of the Hamiltonian are made. The Born-Oppenheimer approximation was already treated in section 2.1.2, and the (Dirac)-Hartree-Fock ((D)HF) method will be introduced in section 2.2.3. To lead into this, the variational principle and Slater determinant will be treated first in sections 2.2.1 and 2.2.2 respectively, as HF is a variational method where the wavefunction is expressed as a Slater determinant.

2.2.1 Variational principle

The variational principle can be used when an exact solution of the Schrödinger equation cannot be found, and methods based on it can be described as a form of informed trial-and-error. When a system is described by a Hamiltonian H with its lowest eigenvalue E_0, the Rayleigh ratio $\mathcal{E}$ [56] is defined through

$$\mathcal{E} = \frac{\int \psi^*_{trial} H \psi_{trial}\, d\tau}{\int \psi^*_{trial}\psi_{trial}\, d\tau} = \frac{\langle \psi_{trial}|\, H\, |\psi_{trial}\rangle}{\langle \psi_{trial}|\psi_{trial}\rangle}. \tag{19}$$

The trial functions can be used to solve the system, as the variation theorem states

$$\forall\ \psi_{trial}, \quad \mathcal{E} \geq E_0, \tag{20}$$

meaning; to minimise $\mathcal{E}$ is to find the ground state of the system [56]. This is done with the Rayleigh-Ritz method. The trial function is described as a sum of basis functions and coefficients [56]

$$\psi_{trial} = \sum_i c_i \psi_i, \tag{21}$$

and the variation of the Rayleigh ratio is minimised by redefining the expression

$$\mathcal{E} = \frac{\int \psi^*_{trial} H \psi_{trial}\, d\tau}{\int \psi^*_{trial}\psi_{trial}\, d\tau} = \frac{\sum c_i^* c_j H_{ij}}{\sum c_i^* c_j S_{ij}} \rightarrow \delta\mathcal{E} = \frac{\sum \delta c_i^* c_j (H_{ij} - \mathcal{E} S_{ij})}{\sum c_i^* c_j S_{ij}} = 0. \tag{22}$$

This can be solved when the numerator vanishes, which will occur for a set of equations [56] defined by

$$\det\{|H_{ij} - \mathcal{E} S_{ij}|\} = 0. \tag{23}$$

These determinants and the basis functions will be further discussed below in sections 2.2.2 and 2.4 respectively.

2.2.2 Slater determinant

A Slater determinant is an N by N determinant, with N being the amount of electrons. It contains all electrons $(1, ..., N)$ and spinorbitals $(\phi_a, ..., \phi_z)$, satisfying the Pauli principle as interchanging a pair of electrons is an antisymmetric action, and no two electrons can be in the same state as the determinant will vanish. It is often written as

$$\Psi(1, 2, ..., N) = \left(\frac{1}{N!}\right)^{1/2} \det |\phi_a(1)\phi_b(2)...\phi_z(N)|, \tag{24}$$

which shows the principal diagonal of the full determinant [56]. The $1/N!$ term is sometimes left out and just implied, as it is a normalisation factor that shows up consistently due to the indistinguishable nature of the electrons [48]. The Slater determinant is an antisymmetrised product of 1-electron orbitals, so when using it we are no longer working in a true many-particle system. In fact, it was developed in the mean-field approximation, where interactions between electrons were simplified greatly [57].

2.2.3 Hartree-Fock

In order to solve the Schrödinger equation within the Born-Oppenheimer approximation, we use another approximation where repulsion between electrons is not calculated exactly, but a field using the average positions is used to interact with each electron separately. This is why the Hartree-Fock theory is called a mean field theory [48]. It is also an ab initio theory, as it uses a potential derived without semi-empirical parameters [58]. To illustrate this, we continue with the variational method, because the Rayleigh ratio leads into the Hartree-Fock equations [56]

$$f(\vec{x}_1)\chi_i(\vec{x}_i) = \epsilon_i\chi_i(\vec{x}_1), \tag{25}$$

with the Fock operator [56];

$$f(\vec{x}_1) = h_D(\vec{x}_1) + \sum_j \mathcal{J}_j(\vec{x}_1) - \mathcal{K}_j(\vec{x}_1) \tag{26}$$

which includes the Coulomb operator [48]

$$\mathcal{J}_j(\vec{x}_1) = \int d\vec{x}_2 |\chi_j(\vec{x}_2)|^2 \frac{1}{r_{12}}, \tag{27}$$

and exchange operator [48]

$$\mathcal{K}_j(\vec{x}_1)\chi_i(\vec{x}_1) = \left[\int d\vec{x}_2 \chi_j^*(\vec{x}_2)\frac{1}{r_{12}}\chi_i(\vec{x}_2)\right]\chi_j(\vec{x}_1). \tag{28}$$

When these equations are solved numerically with basis sets we get the Roothaan equations [56]. For each spin orbital i, there is an expansion of atomic orbital basis functions $\tilde{\chi}$ [48]

$$\chi_i = \sum_\mu^K C_{\mu i}\tilde{\chi}_\mu. \tag{29}$$

When inserting this in the Hartree-Fock equations 25, multiplying by $\tilde{\chi}_\mu^*(\vec{x}_1)$ and integrating, we get a matrix equation very reminiscent of what appeared in the variational theory in section 2.2.1;

$$\sum_\nu F_{\mu\nu}C_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu}C_{\nu i} \rightarrow \mathbf{FC} = \mathbf{SC}\epsilon, \tag{30}$$

with

$$S_{\mu\nu} = \int d\vec{x}_1 c\tilde{h}i^*_\mu(\vec{x}_1)\tilde{\chi}_\nu(\vec{x}_1) \tag{31}$$

$$F_{\mu\nu} = \int d\vec{x}_1 \tilde{\chi}^*_\mu(\vec{x}_1) f(\vec{x}_1)\tilde{\chi}_\nu(\vec{x}_1) \tag{32}$$

where ϵ contains the orbital energies on the diagonal, and $\mathbf{S}$ is the overlap matrix that will vanish when a transformation to an orthogonal basis is performed. The Hartree-Fock-Roothaan model is a self-consistent field method, as $\mathbf{F}$ is solved iteratively. [48]

While this method was explained entirely in the non-relativistic regime, the Fock operator in equation 26 already shares clear similarity with the relativistic Hamiltonian in equation 12 from the Dirac description (section 2.1.3). The 1-electron component of the relativistic Hamiltonian contains terms that are comparable to the h_D and $\mathcal{J}$ terms in the Fock operator. The Coulomb operator $\mathcal{J}$ is directly the non-relativistic description of the potential V_{eN}. The exchange operator takes spin correlation into account [56], which is more explicitly tied into the Dirac equation as it has no classical analogue. Therefore, while I am not explaining the details of the relativistic Dirac-Hartree-Fock method, it can be understood from the same principles as the non-relativistic equations are based on.

2.3 Electron interaction

The approximation of an average field is necessary to make calculations of large systems possible, but it removes electron correlation from the picture. The instantaneous interactions between electrons and other quantum mechanical electron-electron effects, which can together be called electron correlation, can be recovered with a few different methods.

2.3.1 CI and CC

There are multiple ways to include electron-electron interaction on top of Hartree-Fock calculations. One of these is Configuration Interaction (CI), while another is the coupled cluster (CC) method. The CI method starts by writing the exact wavefunction Ψ as a linear combination of Slater determinants

$$\Psi = c_0\Psi_0 + \sum_{a,p} c^p_a \Psi^p_a + \sum_{a<b,p<q} c^{pq}_{ab}\Psi^{pq}_{ab} + ... \tag{33}$$

where Ψ_0 is the HF ground state wave function. Each sum is over pairs of spinorbitals, singly excited (a,p), doubly excited $(a\&b, p\&q)$, and so on. The complete inclusion of all sums accounts for the electron correlation that is not taken into account in the Hartree-Fock method, and is called full CI. However, that approach is practically difficult for many-electron systems as it is computationally expensive. Often CI gets truncated as CISD, with only single and double excitations, though that is not size extensive. This means the energy does not scale linearly with the number of particles of the system [59]. It is also not size consistent when truncated [6], meaning the energy of infinitely distanced subsystems A and B would not be the same as the energy of A and B added up where they treated as separate systems [60].

The method used in this thesis is the coupled cluster method, where the exact wavefunction is related to Ψ_0 through

$$\Psi = e^C\Psi, \tag{34}$$

[6] The QCISD method of truncation can restore size consistency for CISD [60], but any natural truncation of CI without such a correction is not size consistent

where

$$e^C = 1 + C + \frac{1}{2!}C^2 + ..., \tag{35}$$

is the series expansion defining the operator e^C (in which Ψ_0 is recovered from the 1 term), and

$$C = C_1 + C_2 + ... + C_N \tag{36}$$

further defines the Cluster operator. Here, C_n is the n-electron excitation operator, and N is the number of electrons in the system [56]. As opposed to configuration interaction, CC is size extensive [61] and size consistent [62] even when truncated [7].

In order to combine the coupled cluster method, the Schrödinger equation is transformed as

$$e^{-C}He^C \left|\Psi_0\right\rangle = E \left|\Psi_0\right\rangle, \tag{37}$$

where $e^{-C}He^C$ forms a new operator that can be can be expanded as

$$e^{-C}He^C = H + [H,C] + \frac{1}{2!}[[H,C],C] + \frac{1}{3!}[[[H,C],C],C] + \frac{1}{4!}[[[[H,C],C],C],C] \tag{38}$$

The expansion is based on the Baker-Campbell-Hausdorff formula [63], and contains commutators of operators defined by $[H,C] = HC - CH$. The series can be used because the higher order terms after the fifth commute for the coupled cluster method, meaning these terms are the only non-zero terms in the expansion. In order to solve equation 37, we will need to take a look at the C_i operators in equation 36.

2.3.2 CCSD and CCSD(T)

The one- and two-electron excitation operators of the coupled cluster method act equal to those of CI, namely

$$C_1\Psi_0 = \sum_{a,p} t_a^p \Psi_a^p \tag{39}$$

$$C_2\Psi_0 = \sum_{a,b,p,q} t_{ab}^{pq} \Psi_{ab}^{pq} \tag{40}$$

where t are the amplitudes [56]. As CC is defined with an exponential operator, truncation of the method at Singles (S), Singles and Doubles (SD), or any other is not the same as for CI. The expansion of e^C produces cross terms such as $C_1C_2\Psi_0$. When CI is truncated as CISD, it includes all single and double excitations, but no more. CCSD is a truncation of C as $C = C_1 + C_2$, meaning - on top of all single and double excitations - some triple and quadruple excitations are also included in the calculation. The projection equations for the CCSD method are

$$\left\langle \Psi_0 \right| H \left| (C_2 + \frac{1}{2}C_1^2)\Psi_0 \right\rangle = E_{corr}, \tag{41}$$

$$\left\langle \Psi_i^a \right| \bar{H} \left| \left(C_1 + C_2 + \frac{1}{2}C_1^2 + C_1C_2 + \frac{1}{6}C_1^3 \right) \Psi_0 \right\rangle = t_i^a E_{corr}, \tag{42}$$

[7] For a more extensive discussion on the size-extensivity of truncated CC, see [59], as there are some different models and methods of applying CC that are more or less size extensive

$$\left\langle \Psi_{ij}^{ab} \right| \bar{H} \left| \left(1 + C_1 + C_2 + \frac{1}{2}C_1^2 + C_1C_2 + \frac{1}{6}C_1^3 + \frac{1}{2}C_2^2 + \frac{1}{2}C_2C_1^2 + \frac{1}{24}C_1^4 \right) \Psi_0 \right\rangle = (t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a) E_{corr}, \quad (43)$$

with $\bar{H} = H - E_{HF}$. This is how the coupled cluster method gets combined with the Hartree-Fock method [60], solving the transformed Schrödinger equation 37. The projection method is used with the coupled cluster method instead of the variational principle, as the equations become rather unmanageable otherwise. This means the correlation energy is not related to an upper bound, as we would expect from methods based on the variation theorem.

The DIRAC programme used in this project is capable of CCSD calculations, but not of full CCSDT due to the computational cost. Instead, it used a perturbative method to include triple excitations, noted as CCSD(T). The triples correction

$$\Delta E_T(CCSD) = \left(\sum_S^S + \sum_S^D \right) \sum_t^T \sum_u^D (E_0 - E_t)^{-1} t_S V_{si} V_{tu} t_u, \quad (44)$$

gets added converged CCSD energies E_0, with V representing the perturbation operator [60]. There are a few different methods of adding triple excitations perturbatively, but CCSD(T) is currently used as the golden standard. This is because it uses not only all fourth order terms, but also one fifth order perturbative term that corrects for an imaginary frequency produced in some alternative methods [64][60].

2.4 Basis sets

The use of the variation theorem within the HF procedure leads us to use trial functions to find the lowest energy state. These trial functions are then described in equation 21 as a sum of coefficients and basis functions. These coefficients are optimised, while the basis functions remain fixed. Basis sets are used to describe the number and type of basis functions used to approximate the electronic orbitals.

2.4.1 Gaussians or Slater orbitals

Basis functions need to be centred on the atoms in the molecule, to represent the electrons moving around the heavy nucleus. The two most often used function forms to describe orbitals are Slater-type and Gaussian-type. The difference between them is quite small, as they are defined by

$$\phi_{abc}^{STO}(x, y, z) = N x^a y^b z^c e^{-\zeta r}, \quad (45)$$

$$\phi_{abc}^{GTO}(x, y, z) = N x^a y^b z^c e^{-\zeta r^2}, \quad (46)$$

with N a normalisation constant, a, b, c relating to angular momentum, and ζ the width of the orbital [56]. Slater-type orbitals represent orbitals more accurately (see Figure 3a), but are much harder to work with than Gaussian-type orbitals. The latter have analytical solutions and the useful feature that the product of two Gaussians will form a Gaussian (see Figure 3b). Therefore, combining multiple Gaussian functions to form orbitals is preferred over the use of Slater-type orbitals in many programmes, including the DIRAC programme used in this project [49].

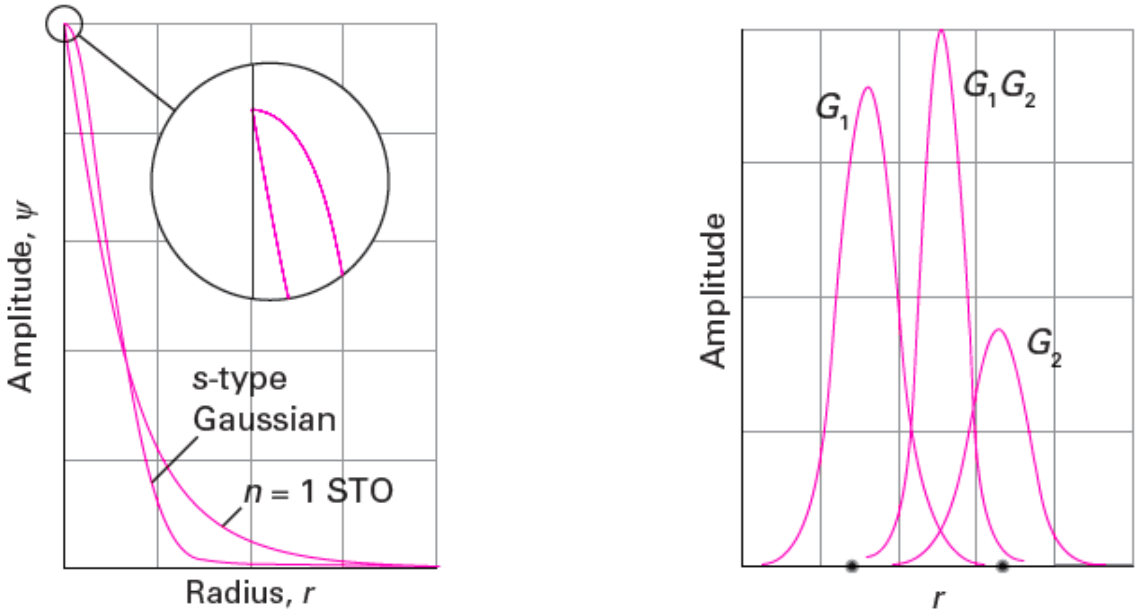


Figure 3: (a) Orbital of Slater-type (STO) has a cusp at a nucleus, while a Gaussian does not (b) Gaussian product G_1G_2 is centred between G_1 and G_2. (Figures 9.3 and 9.4 of [56])

2.4.2 Molecular orbits

Similar to how trial functions are a sum of coefficients and basis functions, molecular orbitals are approximated as a sum of coefficients and atomic orbitals represented by basis functions;

$$\chi_k = \sum_s c_{ks}\phi_s \tag{47}$$

This is known as the linear combination of atomic orbitals (LCAO) approximation [56]. The coefficients are optimised based on variational theory. This description of molecular orbitals is a necessary approximation, because their behaviour is more complex than that of atomic orbitals, especially for the bonding region. In an atom, letters like s, p, and d are used to refer to different shells or energy levels, which get replaced by σ, π, δ, et cetera in molecules. In the valence shell, where atomic orbitals in the molecule overlap most, the combined molecular orbitals are particularly mixed in character. The determination of molecular structure with these methods is done based on molecular orbital theory. [56]

2.4.3 Naming system of basis sets

In this project, the Dyall basis sets are used, which are tailored for relativistic calculations [65][66][67]. Within these basis sets, accuracy of approximation and computational cost can be balanced multiple ways.

First, the amount of functions to describe orbitals is described with the cardinality. For example, a minimal basis set contains one function to represent each occupied orbital; s-type orbitals get 1

functions, p-type get 3 for the combined p_x, p_y, and p_z orbitals, and so on. When two functions are used per orbital this is referred to as a double zeta basis set (e.g. v*Dz* or v*2z*), which adds accuracy in the representation of the orbitals. Using three or four functions per orbital, triple and quadruple zeta basis sets are the ones used most in this thesis.

Next, the difference between core and valence orbitals is of importance. For the ionisation potential, the electrons in the most outer shells tend to be more relevant than the electrons closest to the nuclei. The name of a basis set will have a *v* in it to signify split-valence, where a single set of functions is used for core electrons, and cardinality describes that of valence electrons only. It is a small compromise of accuracy, in order to save computational resources. For other molecular properties however, the core electrons have a larger effect, and split-valence basis sets provide less reliable results. For a complete basis set where all electrons are actively correlated instead of only valence orbitals, *ae* is used in the basis set name in place of the *v* (e.g. v3z and ae3z).

Furthermore, additional diffuse functions can be added to more accurately describe excited states or ions, or other situations where orbitals are extended to large distances. They are referred augmented functions, and are Gaussians with a small exponent (see equation 46). This is often described with the prefix *aug* - for augmented functions - combined with an *s* (for a single function). When the augmented function is optimised for Dyall basis sets, *a* is used instead of *aug* [67], and they can also be combined (e.g. s-aug-v4z, av4z, and s-aug-av4z).

2.4.4 CBS limit

Extending basis sets with higher cardinality and including more and more functions is too computationally expensive. However, progressively larger basis sets give more accurate results, from which it is possible to extrapolate the converged lowest energy. The complete basis set limit is found by an exponential or power law extrapolation function using the energy points at different cardinality. Further corrections can be added by comparing X2C and 4C results, or including more electrons in the valence set.

Due to the convergence rate of DHF and CC calculations being different, there are two extrapolation methods. For DHF energies, an exponential formula applies [68]

$$E(x) = E_\infty + Be^{-\alpha X}, \tag{48}$$

while for CCSD and CCSD(T) correlation energies, a power law is used [69]

$$E(x) = E_\infty + BX^{-\alpha}. \tag{49}$$

The x refers to the cardinality of the basis set, and through filling in the energies at three cardinal numbers 2, 3, and 4, the CBS energy (E_∞) and fitting parameters B and α can be found. In the exponential case (eq. 48), the joined equations are;

$$E_\infty = \frac{E(D\zeta)E(Q\zeta) - E(T\zeta)^2}{E(D\zeta) - 2E(T\zeta) + E(Q\zeta)}, \tag{50}$$

$$B = \frac{(E(D\zeta) - E(T\zeta))^4}{(E(T\zeta) - E(Q\zeta))^2 (E(D\zeta) - 2E(T\zeta) + E(Q\zeta))}, \tag{51}$$

$$\alpha = \ln\left(\frac{E(D\zeta) - E(T\zeta)}{E(T\zeta) - E(Q\zeta)}\right), \tag{52}$$

where $D\zeta$, $T\zeta$, and $Q\zeta$ refer to double, triple, and quadruple zeta [69].

For the power law in equation 49, there is no analytical solution, so instead we use an approximation where $\alpha = 3$, and find

$$E_{\infty}^{corr} = \frac{4^3 E(Q\zeta) - 3^3 E(T\zeta)}{4^3 - 3^3}, \tag{53}$$

$$B = 3^3 \left(E(T\zeta) - E_{\infty} \right), \tag{54}$$

as the CBS correlation energy [70][71][42]. For the calculations of this thesis project, the CBS-limit solutions were found using the formulas above in my own python script.

3 Methods

In this section, the theoretical knowledge of section 2 is applied to explain the strategy for the calculations reported in section 4. This is split up in the approach of bond length calculations (section 3.1) as the basis of the determination of ionisation potentials (section 3.2), and finally the procedure of uncertainty estimation (section 3.3).

For this project, the 2023 and 2024 versions of the DIRAC programme [49] were used via the Hábrók computing cluster [72]. Post-processing was done in Python [73], using libraries os, numpy [74], copy, and matplotlib.pyplot [75].

3.1 Bond length optimisation

The bond lengths of the molecules were optimised with different Dyall basis sets, starting with vXz (cardinality $X = 2, 3, 4$), and adding two layers of augmentation; one pre-optimised (avXz sets), the next obtained by a ratio (s-aug-avXz) [8]. For each basis set, the bond lengths were found with 8th degree polynomial curves fitted to at least 9 points of data from different calculations, meaning the R_e is converged to at least mÅ [9]. This was done with python, using the function "numpy.polynomial.polynomial.polyfit" [74], which returns the coefficients of a polynomial equation

$$p(x) = \sum_{i=0}^{n} c_i x^i \tag{55}$$

up to a specified degree n of the least-squares fit polynomial for the dataset. To find the equilibrium bond length, the derivative of the polynomial is taken through

$$p'(x) = \sum_{i=1}^{n} i c_i x^{i-1}. \tag{56}$$

The minimum can then be found with the function "numpy.polynomial.polynomial.polyroots" [74]. This function also gives maxima and other extrema on the complex domain. To focus on the equilibrium bond length, I first required the the potential R_e to be real, ignoring the complex results of the polyroots function through a check with the function "numpy.iscomplex". I then specified a domain with the calculated bond length - assuming that R_e would be between the smallest and

[8] TaN could not be optimised with additional augmented basis sets due to convergence issues, which are discussed in section 4.2

[9] The convergence rate of the polynomial curves depending on the degree was determined with a different molecule, as part of a tutorial in the use of the DIRAC programme developed by prof. Anastasia Borschevsky for internal use by the VSI-AIM research group

biggest bond lengths I had calculated with DIRAC - and this should leave only one minimum, which is the equilibrium bond length.

For all calculations, DHF, CCSD, and CCSD(T) methods were used. The bond length optimisation was done at the X2C level of relativity. For each molecule, an active set of spinors was chosen. In HgCl and TaN this was from -30.0 to 30.0 Hartree, while the active space in HgI was set from -50.0 to 50.0 Hartree. This choice was made after a preliminary DHF calculation determining the energy levels of the different orbitals, where -30.0 Hartree turned out to be an inappropriate cut-off for HgI as it was approximately the energy of an electronic eigenvalue. At -50.0 Hartree, there was no such issue, thus that cut-off was chosen instead. HgCl and TaN did not have this issue, and thus a comparatively smaller active energy region was kept to reduce computational cost. In HgCl, which has a ${}^2\Sigma_{1/2}$ ground state and 97 electrons total, that meant 65 electrons were included in the valence region. TaN has a ${}^1\Sigma^+$ ground state and 80 electrons, of which 52 are included in the active space. Lastly, HgI has a ${}^2\Sigma_{1/2}$ ground state and 133 electrons, of which 91 are treated as valence electrons.

The complete basis set limit extrapolation was calculated using the s-aug-avXz basis set results for HgCl and HgI, with the schemes described in section 2.4.4. For TaN, the augmented basis sets that were not optimised did not produce reliable results and because of that the avXz basis sets were used to determine the CBS bond lengths. These R_e were again found through a 8th degree polynomial fit, this time using the extrapolated energies. Specifically for the CCSD and CCSD(T) results, the correlation energies from the CBS-extrapolation were added to the DHF CBS energies before determining the R_e, ensuring all bond lengths were determined with the total energies that form the potential energy curves. All of these results are compiled in section 4.2.1, and the appendix C.1 contains some additional data.

3.2 Ionisation potentials

The energies at the determined bond lengths were used to calculate the ionisation potential. This was done by setting $x = R_e$ in equation 55 for the neutral molecule and the ion respectively, and taking the difference between the two energies. The factor used to translate from Hartree to eV was: 1 Hartree = 27.211386245981 eV. The results of this method are found in section 4.2.2 and the Appendix C.2.

For better comparison with experiment, I have also employed a method to find a vibrationally corrected adiabatic ionisation potential IP_{vc}, which can be calculated through

$$IP_{vc} = \left(E^+ + \frac{1}{2}\omega_e^+ - \frac{1}{4}\omega_e\chi_e^+\right) - \left(E^n + \frac{1}{2}\omega_e^n - \frac{1}{4}\omega_e\chi_e^n\right), \tag{57}$$

where superscripts + and n refer to the ion and neutral molecule respectively, E are the energies at the respective bond lengths, ω_e are the vibrational frequencies, and $\omega_e\chi_e$ is the anharmonicity correction [42]. Both ω_e and $\omega_e\chi_e$ can be determined from the potential energy curves used to determine R_e, which was done using the twofit programme of DIRAC [76] [49]. Because the twofit programme uses a polynomial fit as well, the degree was kept at 8 for consistency with the bond length calculations.

3.3 Uncertainty estimation

To correct for any approximations made throughout the determination of the ionisation potentials, calculations at different levels of theory and computational ability can be compared. The assumption

made in this overall uncertainty estimation is that all sources of error are independent enough to be treated separately, as they are higher-order effects [42].

First, to estimate the size of relativistic effects like the Breit term and QED corrections, the IPs of all molecules are calculated at X2C, 4C DC (Dirac-Coulomb) and 4C DC+Gaunt level with the v3z basis set using the CBS CCSD(T) equilibrium bond lengths found in previously described steps. The Gaunt term correction can only be calculated at DHF level, so the comparison is made between the SCF level ionisation potentials found with the three methods. The differences between each step of relativity form the corrections, each at the highest level of electron correlation possible. The uncertainty is based on part of the correction term from the highest level and second highest level (4C DC+Gaunt compared to 4C DC), assuming the addition of the full Breit term or QED effects would lead to a smaller contribution than this correction.

Secondly, in the domain of electron correlation, two methods are used. The IPs of the molecules are calculated with the original virtual cutoff of 30.0 Hartree (HgCl, TaN) or 50.0 (HgI), with the v3z basis set at the X2C CBS CCSD(T) R_e, which forms the base for comparison of all correction calculations done. Here, the results are compared to a calculation with the ae3z basis set and a virtual cutoff of 3300.0 Hartree to mirror the cutoff of -3300.0 Hartree, which is large enough to include all electrons. This forms the active space correction. While increasing the negative energy cut-off from here would not make a difference, as all electrons are accounted for, the virtual active space could still be increased. A percentage of the correction can therefore also be taken to represent the uncertainty of the contribution due to including an even larger virtual space. Besides that, 10% of the difference between the CCSD and CCSD(T) result in the CBS limit is taken as an estimate of the error due to higher-order excitations (full triples, quadruples, etc.). For this, no new calculation can be made with the DIRAC programme, but as CCSD(T) is known as the golden standard method, the error is expected to be relatively small. Specifically, the smaller percentage of 10% compared to the other uncertainty estimates (50% of the correction) is based on the fact that higher order excitation corrections are known to exhibit changes of sign, and will potentially cancel out [42].

Lastly, to evaluate the basis set incompleteness, two methods are used as well. To start, the difference between the CBS limit and largest basis set IP results (s-aug-av4z for HgX and av4z for TaN) is taken to base the extrapolation error on. Next, the effect of addition of diffuse functions can be compared, using results from the largest basis sets; v4z, av4z, and (where available) s-aug-av4z. The difference between the respective ionisation potentials can account for the error caused by the lack of additional diffuse functions. The difference between the best and second-best calculation is taken, but whenever a third level is available, differences between all three could show a trend. For the evaluation of basis set completeness, no new calculations are performed, as enough data from large basis set calculations is gathered prior to this step, for ionisation potential calculations described in previous section 3.2.

4 Results and discussion

4.1 Literature research

In order to have a starting point for my calculations, as well as a reference for comparison, I focussed my literature research on the equilibrium bond length, ionisation potential, and previously calculated eEDM enhancement factors. The complete tables with this data can be found in Appendix B (Tables B.1 through B.4). In section 4.1.1 and 4.1.2 below, the most relevant findings are highlighted and elaborated on.

4.1.1 HgCl and HgI

Mercury monohalides (HgF, HgCl, HgBr, and HgI) are a class of polar molecules with a ground electronic state of $^2\Sigma_{1/2}$, meaning there is one electron in an σ-type orbital that is not in a closed shell. This makes them sensitive to eEDMs in their ground state. Moreover, σ-type states are relatively simple to calculate with high accuracy. The heavier HgX molecules are easier to polarise than HgF, making them more interesting in experimental settings. [36]

While the enhancement factors of both HgCl and HgI have been calculated before [36][61][77] (see Table B.3), the ionisation potentials and bond lengths of both are only known to limited degrees of accuracy. The ionisation potential of HgCl and HgI has not been determined experimentally, although the combined ionisation and dissociation energy of HgX_2 [78][79] is often cited to approximately represent the HgX IP. Later, theoretical estimations of the IP where made [80] using normalised elimination of the small component (NESC) CCSD(T) calculations. In terms of improvement; NESC is the predecessor of the X2C method used in this project, and these previous calculations used smaller (cardinality 2 and 3) basis sets as well. Knowing this, I put my expectations on finding an ionisation potential of about 9 eV for HgCl, and about 8 eV for HgI (see Table B.1).

Many more theoretical calculations of the equilibrium bond length of both HgCl and HgI exist, with one of the most accurate to date a CCSD(T) calculation based on the approximately relativistic Douglas-Kroll-Hess Hamiltonian for HgCl expanded to the CBS limit [81]. The most recent theoretical relativistic determination of the bond length of HgI was a CCSD calculation [82], where the authors also calculated the bond length of HgCl and both ions HgI^+ and $HgCl^+$. There is some experimental data for HgCl; the bond length has been determined through electron diffraction photographs [83][84], spectrographs [85], and chemiluminescent spectra [86]. For HgI, there are not many bond length values available at the moment, and none with high accuracy. However, it is known than the R_e of mercury halides increases regularly with the increase of the atomic number of the halogen atom [87]. Thus, a bond length around 2.35 Å in HgCl can be expected, and HgI should have a larger bond length, likely around 2.8 Å (see Table B.2). The ions are both closed-shell systems with a positive charge, and consequently have a smaller bond length, differing from the neutral system by 0.1 Å or so.

4.1.2 TaN

Compared to the mercury diatomics, much less is known about tantalum nitride, though it has attracted attention as a eEDM experiment candidate recently. Tantalum is a transition metal and transition-metal-molecules often have complex spectra, though they are interesting for chemistry and astrophysical [10] research [88]. Because of this complexity, much focus has been on the first excited $^3\Delta_1$ state of TaN, for which the eEDM enhancement factor has been calculated [23][40] (see Table B.4). The ground state of TaN is $^1\Sigma^+$, a closed shell state, though the ground and first excited states lie very close together [88].

The bond length of TaN was experimentally determined through spectrometry from a hollow-cathode lamp [88] as well as laser-induced fluorescence [89]. Some theoretically calculated bond lengths exist as well, the best so far being a recent relativistic CCSD(T) calculation extended to the CBS limit [90], though that calculation does not agree with the experimental findings as well as some other theoretical estimates (see Table B.4). This means a bond length of about 1.7 Å can be expected in

[10] Specifically, diatomic transition metal nitrides like tantalum nitride are important for astrophysical research, while the focus of organic and organometallic chemistry research is on the general larger group of transition metal molecules

my calculations, but ideally we look for agreement with the 1.6830999(88) Å experimental R_e [88]. The bond length of TaN^+ is unknown, but can be expected to be a bit smaller than that of TaN, similar to the trend in HgX/ HgX^+. Different from the mercury halides, TaN has a much smaller bond length than HgCl, even though their atomic number is comparable. Tantalum nitride has a much more covalent bond than HgX, and is of triple bond order [23]. A combination of these two factors likely lead to the large difference in bond lengths between HgCl and TaN.

From my search, it seems the ionisation potential of TaN as a diatomic molecule in the gas phase remains unknown. However, as the bond length was determined with great precision, that can form an anchor in the calculation accuracy of the IP. The comparison between my results and previously determined R_e and IP will be discussed in section 4.2 for all three molecules.

4.2 Calculation results

All bond lengths are in Angstrom [Å], all ionisation potentials are given in eV. The factor used to translate from Hartree to eV was: 1 Hartree = 27.211386245981 eV [43]. The convergence from cm^{-1} to eV is exact, based on [cm^{-1}]= hc/e [eV], using the CODATA recommended values of h, c, and e [43].

4.2.1 Bond length optimisation

The bond lengths of HgCl, HgI, TaN, and their respective ions were determined for different basis sets and at the CBS limit through extrapolation, following the strategy described in section 3.1. An example of the potential energy curve is shown in Figure 4, with the equilibrium bond length of HgCl at the extrapolated CBS limit. For completeness, Figures C.1 till C.5 are included in Appendix C.1, displaying the same R_e for the other molecules and their ions. While every polynomial function is based on at least 9 calculated points, it can be seen in these figures that 11 to 13 points were taken to determine the final CBS R_e. These 12 points are bond lengths for which a (s-aug-)avXz calculation was performed at all three levels of cardinality (X=2,3,4). These added points contribute to a higher accuracy of the curve and thus R_e for the final CBS values. Figure 4 shows more calculated points above the equilibrium bond length than below, due to the fact that the R_e of smaller basis sets tends to be larger. Thus, calculations at larger bond lengths are performed, and the parabola-like shape becomes less obvious. The relationship between equilibrium bond length and basis set size is visualised in Figures 5 and 6.

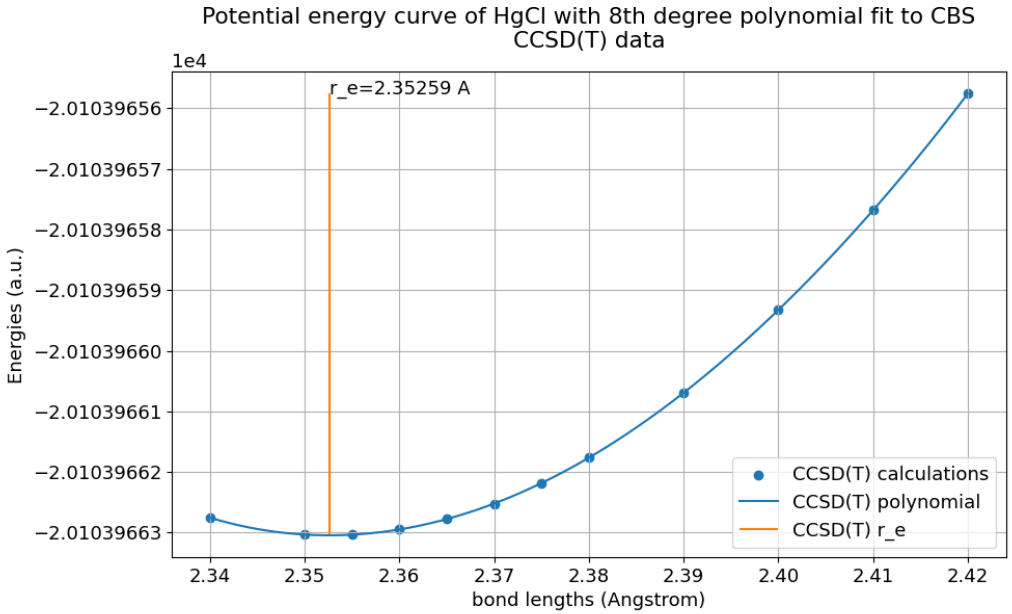


Figure 4: HgCl equilibrium bond length of the CBS X2C CCSD(T) calculation

Figures 5 and 6 are included to illustrate the trend of the equilibrium bond length decreasing with increasing basis set size. The included R_e per molecule are mentioned in Tables 1, 2, and 3. The s-aug-v3z and s-aug-v4z results were left out because s-aug-vXz calculations are at the same level of augmentation as avXz, only being slightly less accurate due to the lack of optimisation. While the differences between both single augmentation methods are small, the stepwise increase of basis set size and the trend that ensues is more clear when including only one of the methods. For completeness, the s-aug-vXz R_e are included in Tables C.1, C.2, and C.3 in the appendix (and the ionisation potentials for these basis sets in Table C.4 and C.5).

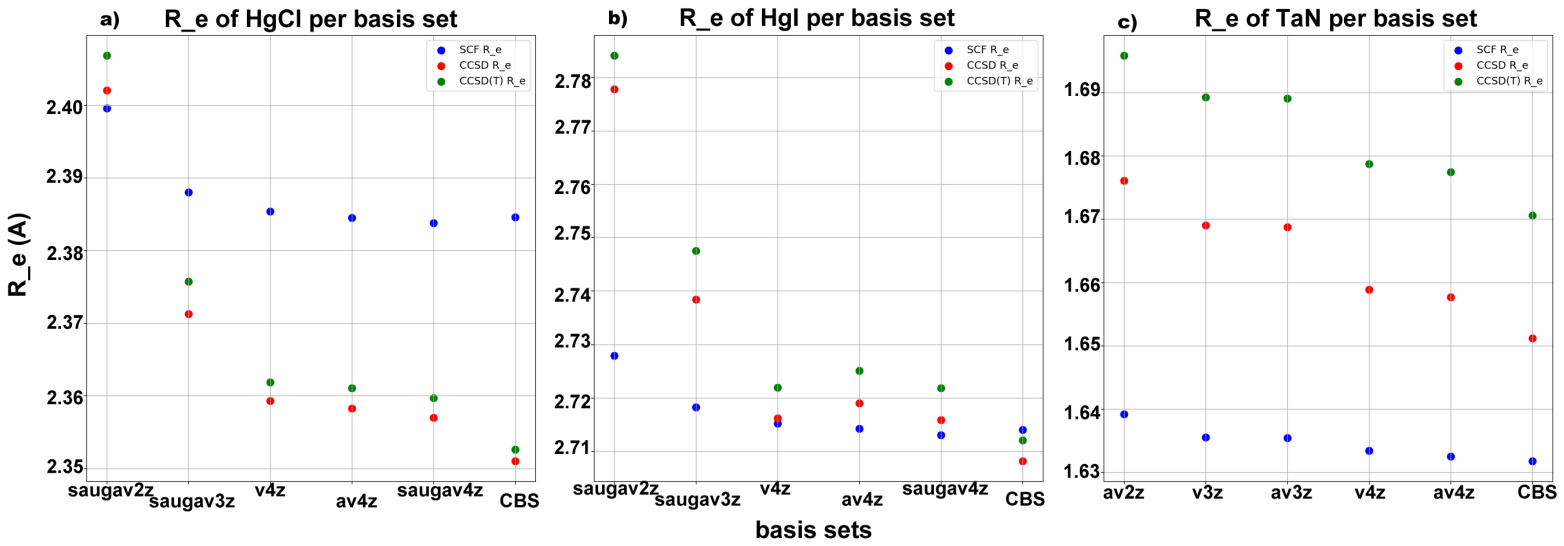


Figure 5: Equilibrium bond lengths trend with increasing size of basis set of molecules HgCl, HgI, and TaN

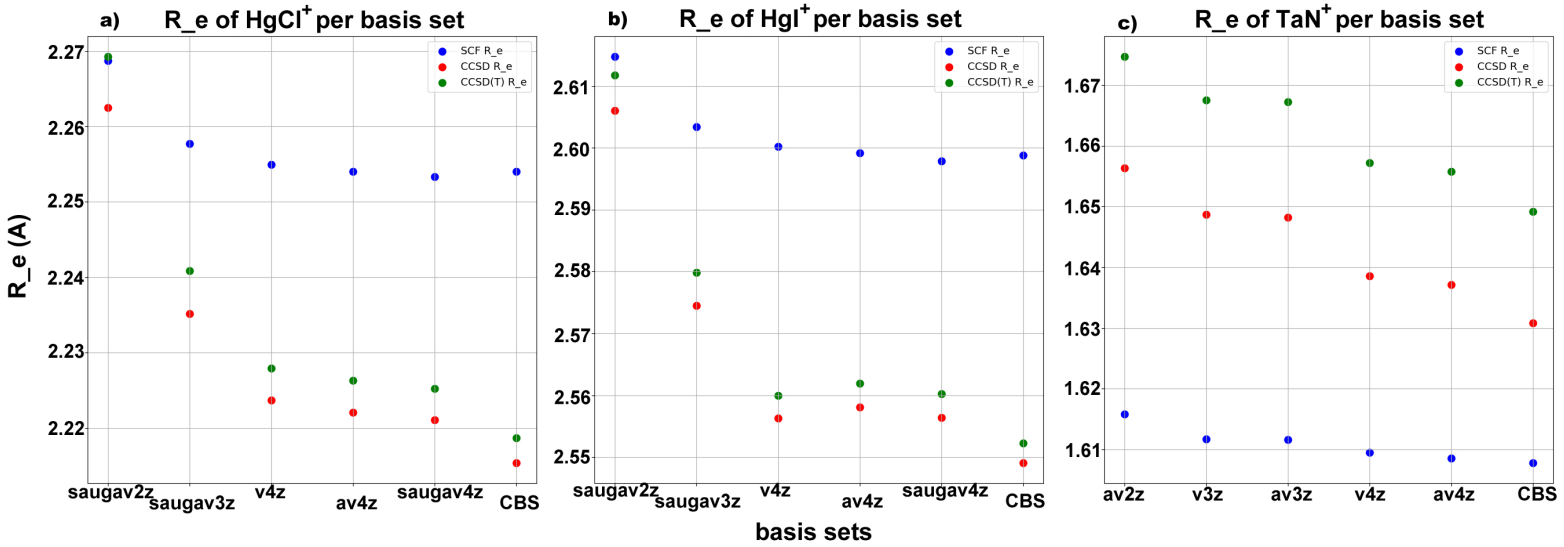


Figure 6: Equilibrium bond lengths trend with increasing size of basis set of molecules HgCl, HgI, and TaN

The coupled cluster calculations especially exhibit a large drop in bond length with increasing basis set size, while SCF bond lengths settle much more quickly. This was expected, as the convergence rate of DHF energies is in accordance with an exponential law, while the correlation energies follow a power law (see section 2.4.4 for the equations). At a cardinality of 4, the figures also show the determined bond length slightly lowers with each addition of a diffuse function. This can be explained by the added accuracy these augmented functions bring to the description of the area further away from the nuclei, and was thus also an expected trend.

In both the HgCl and HgI trends the decrease in R_e for CC calculations is more intense than for TaN. Additionally, the shift in equilibrium bond length for HgI and its ion is a bit more erratic rather than following a smooth trend. To explore this behaviour in the future, basis sets of higher cardinality might be employed. For HgI the additional exploration of non-augmented cardinality $X = 2, 3$ basis sets would show if the erratic behaviour is similar between all vXz, avXz and s-aug-avXz calculations. On the other hand, there is a bigger difference between the CCSD and CCSD(T) trends for TaN and the addition of electron interaction raises the bond length compared to the SCF results. The latter also happens for HgI. In this case, higher excitation terms would be most interesting in future research.

basis set	HgCl R_e [Å]			$HgCl^+$ R_e [Å]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
saugav2z	2.400	2.402	2.407	2.269	2.262	2.269
saugav3z	2.388	2.371	2.376	2.258	2.235	2.241
v4z	2.385	2.359	2.362	2.255	2.224	2.228
av4z	2.385	2.358	2.361	2.254	2.222	2.226
saugav4z	2.384	2.357	2.360	2.253	2.221	2.225
CBS (saugavXz)	2.385	2.351	2.353	2.254	2.215	2.219
Final (CBS CCSD(T))	2.353 ± 0.024			2.219 ± 0.024		

Table 1: HgCl and $HgCl^+$ optimised bond lengths per basis set, with the CBS limit extrapolation based on the saugavXz (X=2,3,4) results

basis set	HgI R_e [Å]			HgI^+ R_e [Å]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
saugav2z	2.728	2.778	2.784	2.615	2.606	2.612
saugav3z	2.718	2.738	2.748	2.603	2.575	2.580
v4z	2.715	2.716	2.722	2.600	2.556	2.560
av4z	2.714	2.719	2.725	2.599	2.558	2.562
saugav4z	2.713	2.716	2.722	2.598	2.556	2.560
CBS (saugavXz)	2.714	2.708	2.712	2.599	2.549	2.552
Final (CBS CCSD(T))	2.712 ± 0.024			2.552 ± 0.024		

Table 2: HgI and HgI^+ optimised bond lengths per basis set, with the CBS limit extrapolation based on the saugavXz (X=2,3,4) results

The tantalum nitride CBS extrapolation was performed using the avXz calculations, while the CBS results of HgCl and HgI both are based on s-aug-avXz energies. The difference is a layer of augmentation that could not be reached for the TaN calculations, as non-optimised augmentation calculations often did not converge at the SCF level (s-aug-vXz and s-aug-avXz). This happened especially with larger basis sets (s-aug-v4z and s-aug-av4z), although enough s-aug-v3z calculations converged to determine the R_e and the IP, both included in Table C.3 and C.5 in the appendix. The convergence issue seemed to arise from the molecule not reaching the ground state configuration, instead getting stuck in a semi-stable state somewhere between the first excited state $^3\Delta_1$ and $^1\Sigma^+$. Because the two states are known to be so close together, this is likely the cause of the behaviour. For molecular properties where augmented functions are more essential, this issue with TaN should be taken into account by analysing the electronic eigenvalues and performing a Mulliken population analysis of the valence electrons.

basis set	TaN R_e [Å]			TaN^+ R_e [Å]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
av2z	1.639	1.676	1.696	1.616	1.656	1.675
v3z	1.636	1.669	1.689	1.612	1.649	1.668
av3z	1.635	1.668	1.689	1.612	1.648	1.667
v4z	1.633	1.659	1.679	1.610	1.639	1.657
av4z	1.632	1.658	1.677	1.609	1.637	1.656
CBS (avXz)	1.632	1.651	1.671	1.608	1.631	1.649
Final (CBS CCSD(T))	1.671 ± 0.024			1.649 ± 0.024		

Table 3: TaN and TaN^+ optimised bond lengths per basis set, with the CBS limit extrapolation based on the avXz (X=2,3,4) results

For HgCl, there are many earlier determined bond lengths available to compare my results to. First, the R_e of 2.353 Å differs from the experimentally determined R_e's in Table B.2 with about 100 mÅ. However, these experimental results are mutually inconclusive, as they range from 2.23 [83] to 2.532 Å [84] with no overlap in confidence interval due to differing methods. My calculation does fall in this experimental range. The agreement with the most recent theoretical determination of 2.331 Å R_e [34] and earlier CBS-extrapolation to 2.3777 Å [81] is much better, my R_e falling in the middle and differing with less than 25 mÅ from each. These differences can be expected from different levels of relativity - the Douglas-Kroll-Hess Hamiltonian [81] being an earlier two-component approximation of 4C than X2C [49], and the other approach being quasi-relativistic [34].

The literature results to compare to for HgI are much more sparse. Experimentally, there is only the large minimum-maximum range from 2.36 to 2.81 Å [86], and 2.443 Å [84] which is unexpectedly smaller than the bond length of HgCl determined in the same paper. The latter result is thus marked as doubtful, considering the trend of increasing bond length with larger halides in HgX is well known [87]. This leaves only the range, in which my calculated R_e of 2.712 Å does fit. In 2022, the bond length of HgI was calculated at 2.7642 Å with a quasi-relativistic method and multi-reference CI [87], rather than the single reference CC X2C approach taken in this paper. This and the difference in basis set choice explains the $\sim$ 50 mÅ difference between the results. My calculation is likely more accurate due to the use of a more precise relativity correction technique, the basis set limit extrapolation, and higher order excitations, even if it is only a single reference method. CCSD calculations with the Fock space method could be employed to gain a multi-reference CCSD result in the future. This approach was not employed in this thesis as the DIRAC programme does not support the use of the perturbative triple CCSD(T) in Fock space calculations, and a multi-reference method is generally most helpful in complex open-shell states.

In contrast to the mercury halides, the experimentally determined R_e of 1.6830999(88) Å for TaN [88] agrees very well with my equilibrium bond length of 1.671 Å, differing only by 12 mÅ. From the previous theoretical estimates, only the 1.6831 Å result agrees better with experiment, likely partially due to the fact that it is based on a full 4C relativistic method [40]. However, as a smaller basis set and CI was used, there could be a favourable error cancellation leading to the better agreement. As can be seen in Table 3, my av3z R_e is actually closest to the experimental bond length. However, the bond length also increases with added electron correlation from DHF to CCSD(T), so added higher excitations calculations with the larger basis sets would together give an even more accurate R_e for TaN, if higher precision is required in the future.

Overall, the determined bond lengths are accurate to the mÅ level as equilibria for the specific applied methods due to the degree of the polynomial curves. From the agreement between the bond length of TaN and the experimental bond length, a conservative error of twice the difference is taken as the uncertainty of all R_e. The final equilibrium bond lengths determined in this project of 2.353 Å (2.219 Å) for HgCl ($HgCl^+$), 2.712 Å (2.552 Å) for HgI (HgI^+), and 1.671 Å (1.649 Å) for TaN (TaN^+) have an uncertainty of about 24 mÅ.

4.2.2 Ionisation potentials

The energies at the determined bond lengths were used to calculate the ionisation potential according to the strategy described in section 3.2. The convergence trend for the IP of HgCl with increasing basis set size can be observed in Figure 7. The trend is not quite comparable to those in the bond length illustrated in Figures 5 and 6. For the SCF calculations, the IP lowers and quickly converges towards the CBS limit, as happens for the R_e. On the other hand, the CCSD and CCSD(T) calculated ionisation potentials show a slight increase along with the basis set size. This can be explained by the energy stabilisation that occurs in the neutral molecules, which is larger than that of the ions. The addition of electron correlation effects lowers the energy of the molecules, as the system is initially slightly unstable in its neutral state. The same happens in the ion, but this system is already much more stable due to the charge interactions. The relative difference between molecular and ionic energy thus becomes larger, increasing the calculated ionisation potential. This same trend does not appear for R_e, as the shape of the potential energy curves stays similar for both the neutral molecule and the ion. What is repeated is the relatively large difference in the IP at CCSD and CCSD(T) level of TaN, while the trend of increasing IP with basis set size is less pronounced. This is reflected in the uncertainty, and further discussed in 3.3.

The IPs of HgCl and HgI are summarised in Table 4, and those of TaN in Table 5 (complete versions including s-aug-vXz results in Appendix C.2). At the highest level of precision reached here, the ionisation potentials are determined to be 9.428 eV (HgCl), 8.771 eV (HgI), and 8.461 eV (TaN).

basis set	HgCl IP [eV]			HgI IP [eV]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
saugav2z	8.669	9.339	9.347	8.034	8.622	8.633
saugav3z	8.593	9.330	9.356	7.964	8.646	8.675
v4z	8.583	9.350	9.379	7.955	8.676	8.706
av4z	8.581	9.359	9.391	7.954	8.690	8.722
saugav4z	8.580	9.359	9.391	7.952	8.691	8.723
CBS (saugavXz)	8.583	9.392	9.428	7.955	8.736	8.771

Table 4: HgCl and HgI ionisation potentials per basis set, with the CBS limit extrapolation based on the saugavXz (X=2,3,4) results

TaN IP [eV] / basis set	DHF	CCSD	CCSD(T)
av2z	7.192	8.230	8.406
v3z	7.178	8.262	8.451
av3z	7.177	8.270	8.462
v4z	7.175	8.260	8.456
av4z	7.174	8.263	8.460
CBS (avXz)	7.173	8.259	8.461

Table 5: TaN ionisation potentials per basis set, with the CBS limit extrapolation based on the avXz (X=2,3,4) results

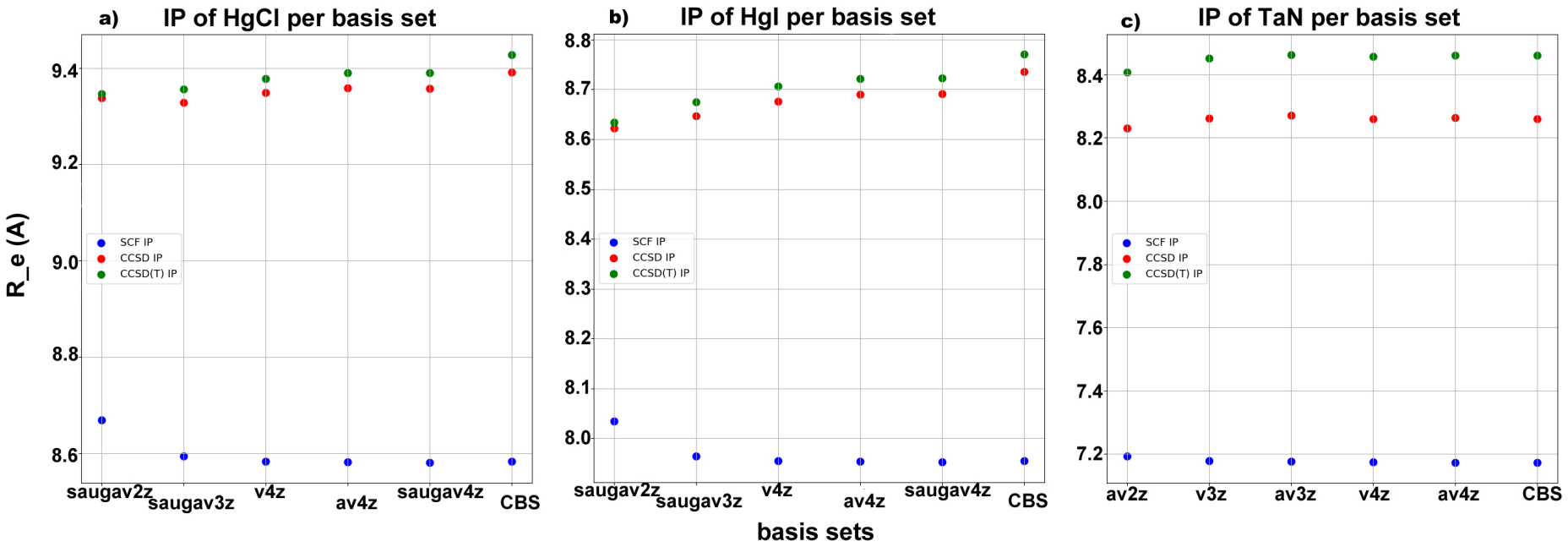


Figure 7: HgCl ionisation potential trend with increasing size of basis set

In order to determine the vibrationally corrected adiabatic ionisation potential IP_{vc}, the vibrational frequencies ω_e and anharmonicity corrections $\omega_e\chi_e$ were found from the potential energy curves of the CBS energies. These spectroscopic constants and the resulting vibrationally corrected ionisation potential IP_{vc} (calculated with equation 57 and the CCSD(T) energies and constants) are included in Table 6. As the focus of this thesis was on theoretical calculations, the addition of the IP_{vc} calculation was considered later, thus the ω_e and $\omega_e\chi_e$ calculations were only made with the CBS results. Therefore, these are not included in the basis set trend analysis.

From the comparison between the terms for the molecule and the ion, it is clear that these corrections almost fully cancel out for TaN, leading to an IP correction of approximately 0.4 meV. For HgI, the difference between ω_e^n and ω_e^+ is about 40 cm^{-1}, and while $\omega_e\chi_e^n$ and $\omega_e\chi_e^+$ are nearly the same size, they are opposite in sign. Still, this only leads to a vibrational correction of about 2 meV for HgI. In HgCl, the $\omega_e\chi_e$ terms approximately cancel out, while the difference between ω_e^n and ω_e^+ is enough to cause IP_{vc} to be about 50 meV larger than IP.

Spectroscopic constants	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
	HgCl [cm^{-1}]			HgCl^+ [cm^{-1}]		
ω_e	309.9	302.9	297.6	396.5	394.5	384.5
$\omega_e\chi_e$	−0.988	8.23	8.14	−2.70	8.14	8.02
	HgI [cm^{-1}]			HgI^+ [cm^{-1}]		
ω_e	158.9	140.7	135.4	185.7	182.9	179.4
$\omega_e\chi_e$	−4.01	−2.22	−3.34	2.46	4.63	4.62
	TaN [cm^{-1}]			TaN^+ [cm^{-1}]		
ω_e	1237.5	1170.8	1107.7	1244.8	1175.9	1114.3
$\omega_e\chi_e$	2.54	3.22	3.98	2.60	3.59	4.08
IP_{vc} (CCSD(T)) [eV]			HgCl	HgI	TaN	
			9.434	8.773	8.461	

Table 6: Spectroscopic constants ω_e and $\omega_e\chi_e$ determined from CBS potential energy curves, and calculated IP_{vc}

The spectroscopic constants ω_e and $\omega_e\chi_e$ have been determined experimentally for some of the included molecules. For HgCl, ω_e is reported as 292.61 cm^{-1}, about 5 cm^{-1} smaller than my estimate. In HgI the difference is a bit larger, with the experimental value being 125.0 cm^{-1}, 10 cm^{-1} below my estimate. It appears the addition of higher excitations is important for the ω_e values to be more comparable with the experimental data, as Table 6 shows a gradual decrease of ω_e from DHF to CCSD(T) for both HgX molecules. For TaN, ω_e has not been determined purely experimentally, though a

non-relativistic DFT calculation reporting 1102.7 cm^{-1} exists and only differs from my calculations by 5 cm^{-1}. Considering these differences, a conservative uncertainty estimate of 15 cm^{-1} can be taken for my ω_e calculations, based taking 1.5 times the difference between the HgI results and experiment.

Next, $\omega_e\chi_e$ has only been measured with high accuracy for HgCl at 1.6025 cm^{-1}, which is very different from my estimate. The value has been estimated for HgI at 1.0 cm^{-1} as well, which shows the stark difference between $\omega_e\chi_e$ for HgCl and HgI in my calculations is likely fully due to inaccuracy rather than a representation of real differences between both molecules. With this, the error is conservatively set to 10 cm^{-1}.

4.2.3 Uncertainty estimation

According to the strategy explained in section 3.3, three correction terms were calculated for the ionisation potentials using energies at the CBS equilibrium bond lengths. The first correction, accounting for a higher active space, comes from a comparison of X2C CCSD(T) v3z IP to the IP calculated from with the ae3z basis set. The active space was increased from the range of -30 to 30 Hartree, to a range of -3300 to 3300 Hartree, while all other settings were kept constant. This was done for both HgCl and TaN, but for HgI a smaller basis set was used - v2z with -50 to 50 Hartree active and ae2z with active -3300 to 3300 Hartree - as the larger size of the molecule made the ae3z calculation computationally much more expensive. However, all HgI calculations were made with a larger active space than HgCl and TaN. This means the HgI calculations include many more virtual orbitals already, leading to a much smaller correction term and error from additional active virtual orbitals. As can be seen in Table 8 virtual space is one of the biggest sources of uncertainty for HgCl and TaN, which could be lowered significantly in future calculations if an active space similar to that of HgI is taken instead.

For the second correction term, the X2C CCSD(T) v3z IP was taken again, this time compared to a calculation using the 4C Dirac-Coulomb Hamiltonian (and CCSD(T) v3z). X2C calculations were made throughout the thesis, as a 4 component method requires more computational resources. The correction is small, < 8 meV, showcasing the accuracy of X2C calculations.

Third, an additional relativity correction was acquired with the difference between the 4C DC DHF ionisation potential and the DC+Gaunt DHF IP, both calculated with the v3z basis set. This correction nearly cancels out the previous, especially for HgCl. However, higher order relativistic effects are known to sometimes cancel out, so this is not unexpected. These three corrections have been added to the ionisation potentials and the vibrationally corrected IP_{vc} per molecule, as shown in Table 7.

Method	HgCl	HgI	TaN
X2C-CCSD(T) CBS IP	9.428323	8.770635	8.460532
IP_{vc}	9.433711	8.773117	8.460934
+ Active space	+0.012917	-0.001120	+0.025803
+ 4C DC	+0.007328	+0.005784	+0.007212
+ Gaunt	-0.007723	-0.007646	-0.008720
Total corrected IP	9.440844	8.767653	8.484827
Total corrected IP_{vc}	9.446232	8.770135	8.485229
IP + corrections & uncertainty	9.411 ± 0.033	8.768 ± 0.032	8.485 ± 0.039
IP_{vc} + corrections & uncertainty	9.446 ± 0.035	8.770 ± 0.035	8.485 ± 0.042

Table 7: Corrected ionisation potentials [eV]

From the relativistic correction from the 4C DC and DC+Gaunt DHF calculations, the uncertainty due to the unaccounted for higher order relativity terms was determined. This is about 4 meV for each molecule, or about 10% of the total uncertainty. Considering that it is a conservative estimate (as higher order effects can also cancel out) and 4 component calculations are computationally heavy, the addition of these higher order effects should not be the main focus for future research when considering the ionisation potential. When turning to the eEDM enhancement factor, which is a relativistic effect, calculations at higher orders of relativity likely become more relevant again.

There are three remaining sources of uncertainty in Table 8, the first of which are related to the choices of basis set. The largest error factor in both mercury diatomics is the CBS extrapolation. Comparatively, the TaN ionisation potentials have converged much better in this regard, which could also be seen in the trend in Figure 7 (section 4.2.2). Likely the cardinality is important for HgCl and HgI, and more accurate results could be gained in the future with the use of basis sets with cardinality 5. For TaN quadruple order basis sets seem perfectly adequate, and increasing the cardinality should not be a main focus.

The difference between v4z and av4z ionisation potentials is bigger than that between av4z and s-aug-av4z, as can be expected for a larger amount of augmented functions. This leads to an uncertainty of 2 meV for TaN where the extra level of augmentation could not be reached, and an error under 0.2 and 0.5 meV for HgCl and HgI respectively. These combined results point to the high accuracy reached using optimised augmented diffuse functions, and smaller benefit of additional augmentation on top of that.

Lastly from this table, in TaN half of the total uncertainty (20 out of 40 meV) is caused by the available level of excitations, as was expected from the trend in Figure 7 (section 4.2.2). For this molecule a calculation of CCSDT energies could be made to compare

to the CCSD(T) IP in the future to gain a much more accurate ionisation potential. Such a calculation would also be beneficial for the mercury halides, but less significant as that error is only about 3.5 meV and on par with the relativity uncertainties.

The total error is reported in Table 8 in meV, and this finally gives us the corrected ionisation potentials with uncertainty estimates for all three molecules as reported in Table 7. As hypothesised, these values fall in the range that allows the molecules to be trapped at cryogenic temperatures.

Category	Error source	HgCl	HgI	TaN
Basis set	Cardinality in CBS extrapolation	18.729	23.870	0.095
	Augmentation	0.177	0.452	1.943
Correlation	Virtual space	6.458	0.560	12.902
	Triple and higher excitations*	3.588	3.511	20.162
Relativity	Breit and QED	3.862	3.823	4.360
Spectroscopic	ω_e	1.860	1.860	1.860
constants	$\omega_e\chi_e$	1.240	1.240	1.240
Total uncertainty IP [meV]		32.814	32.216	39.462
Total uncertainty IP_{vc} [meV]**		35.294	34.696	41.942

*Higher excitations error is taken as 10% of difference, compared to 50% for all other cases

**$\Delta_{IP_{vc}} = \Delta_{IP} + \Delta_{\omega_e} + (1/2)\Delta_{\omega_e\chi_e}$

Table 8: Sources of uncertainty in calculated IP of molecules [meV]

5 Conclusion

In this research project the ionisation potentials of the molecules HgCl, HgI, and TaN were determined with relativistic coupled cluster calculations and a complete basis set-limit extrapolation. The three diatomic molecules are potential eEDM experiment candidates, but many of their properties are not well known. As the ionisation potential can provide a deeper understanding of the electronic structure, the goal of this thesis was to determine this with benchmark accuracy. Therefore, on top of the X2C CCSD(T) CBS IP, a vibrationally corrected IP_{vc} (for better comparison with experiment) as well as added correction terms for higher active space and 4 component relativistic calculations were determined. Lastly, the uncertainty of the ionisation potentials were estimated based on identifying differences between increasingly better levels of calculation and comparing to available spectroscopic data.

Through the uncertainty analysis, it became clear future calculations of the ionisation potential in these molecules could benefit in three main ways. The HgCl and HgI CBS extrapolations were less well converged than the extrapolation for TaN. In the mercury diatomics, calculations at higher cardinality ($X > 4$) could lower the uncertainty by over 50%. Secondly, the addition of higher excitations (full triples and beyond) would be the biggest factor in improving the accuracy of the IP for TaN, potentially decreasing the error by 40%. As a third point, HgCl and TaN calculations were made with an active space of -30 to 30 Hartree, while HgI calculations had an active space of -50 to 50 Hartree. The virtual space corrections of HgI were found with lower cardinality calculations than the other two molecules, but due to the larger initial active space these corrections were much smaller for HgI compared to HgCl and TaN. Future calculations, especially for the IP of TaN, could thus benefit from the use of an active space around -50 to 50 Hartree. The relativistic corrections and error due to lack of Breit and QED corrections are not the most significant for the ionisation potential of these three molecules, though they might be more relevant in the eEDM enhancement factor calculations as that is a relativistic property.

The final values determined are $IP = 9.411 \pm 0.033$ eV and $IP_{vc} = 9.446 \pm 0.0035$ eV (HgCl); $IP = 8.768 \pm 0.032$ eV and $IP_{vc} = 8.770 \pm 0.035$ eV (HgI); $IP = 8.485 \pm 0.039$ eV and $IP_{vc} = 8.485 \pm 0.042$ eV (TaN). These are high enough to allow the molecules to be trapped at cryogenic temperatures and used in eEDM and other Standard Model-testing research. The molecules HgCl, HgI, and TaN thus join the list of prospective eEDM experiment candidates able to be trapped with state-of-the-art techniques, at least on account of their high ionisation potentials. We can look forward to future experiments with lower statistical uncertainty considering this new knowledge, and perhaps the measurement of the eEDM is on the horizon.

References

[1] J. D. Wells, "Prof. von Jolly's 1878 prediction of the end of Theoretical Physics as Reported by Max Planck", (2016).

[2] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", Physics Letters B **716**, 1–29 (2012).

[3] M. Pospelov and A. Ritz, "Electric dipole moments as probes of new physics", Annals of Physics **318**, 119–169 (2005).

[4] J. Morrison, *Modern Physics with Modern Computational Methods* (Academic Press, 2020).

[5] J. F. J. van den Brand, "Introductory Course: Discrete symmetries", Vrije Universiteit Amsterdam.

[6] R. Lehnert, "CPT symmetry and its violation", Symmetry **8**, 114 (2016).

[7] M. S. Safronova, D. Budker, D. DeMille, D. F. J. Kimball, A. Derevianko, and C. W. Clark, "Search for new physics with atoms and molecules", Reviews of Modern Physics **90**, 025008 (2018).

[8] R. Aaij, C. Abellan Beteta, B. Adeva, M. Adinolfi, C. Adrover, A. Affolder, Z. Ajaltouni, J. Albrecht, F. Alessio, M. Alexander, et al., "First observation of CP violation in the decays of B_s^0 mesons", Physical Review Letters **110**, 221601 (2013).

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2 π decay of the K_2^0 meson", Physical Review Letters **13**, 138 (1964).

[10] M. Denis, P. A. Haase, R. G. Timmermans, E. Eliav, N. R. Hutzler, and A. Borschevsky, "Enhancement factor for the electric dipole moment of the electron in the BaOH and YbOH molecules", Physical Review A **99**, 042512 (2019).

[11] T. H. Wright, "Stray fields and the electron's electric dipole moment", PhD thesis (University of Colorado at Boulder, 2024).

[12] T. S. Roussy, L. Caldwell, T. Wright, W. B. Cairncross, Y. Shagam, K. B. Ng, N. Schlossberger, S. Y. Park, A. Wang, J. Ye, et al., "An improved bound on the electron's electric dipole moment", Science **381**, 46–50 (2023).

[13] ACME Collaboration et al., "Improved limit on the electric dipole moment of the electron", Nature **562**, 355–360 (2018).

[14] K. I. Baklanov, A. N. Petrov, A. V. Titov, and M. G. Kozlov, "Progress toward the electron electric-dipole-moment search: Theoretical study of the PbF molecule", Physical Review A **82**, 060501 (2010).

[15] M. Jung, "A robust limit for the electric dipole moment of the electron", Journal of High Energy Physics **2013**, 1–15 (2013).

[16] V. S. Prasannaa, "The search for the electric dipole moment of the electron (eEDM) in mercury halides using the relativistic coupled cluster theory", PhD thesis (Indian Institute of Astrophysics, 2017).

[17] NL-eEDM collaboration, P. Aggarwal, H. L. Bethlem, A. Borschevsky, M. Denis, K. Esajas, P. A. Haase, Y. Hao, S. Hoekstra, K. Jungmann, et al., "Measuring the electric dipole moment of the electron in BaF", The European Physical Journal D **72**, 1–10 (2018).

[18]D. Mukherjee, B. K. Sahoo, H. S. Nataraj, and B. P. Das, "Relativistic coupled cluster (RCC) computation of the electric dipole moment enhancement factor of francium due to the violation of time reversal symmetry", The Journal of Physical Chemistry A **113**, 12549–12557 (2009).

[19]D. DeMille, "Diatomic molecules, a window onto fundamental physics", Physics Today **68**, 34–40 (2015).

[20]B. L. Augenbraun, Z. D. Lasner, A. Frenett, H. Sawaoka, C. Miller, T. C. Steimle, and J. M. Doyle, "Laser-cooled polyatomic molecules for improved electron electric dipole moment searches", New Journal of Physics **22**, 022003 (2020).

[21]K. Gaul, N. R. Hutzler, P. Yu, A. M. Jayich, M. Iliaš, and A. Borschevsky, "CP-violation sensitivity of closed-shell radium-containing polyatomic molecular ions", Physical Review A **109**, 042819 (2024).

[22]L. Anderegg, N. B. Vilas, C. Hallas, P. Robichaud, A. Jadbabaie, J. M. Doyle, and N. R. Hutzler, "Quantum control of trapped polyatomic molecules for eEDM searches", Science **382**, 665–668 (2023).

[23]L. V. Skripnikov, A. N. Petrov, N. S. Mosyagin, A. V. Titov, and V. V. Flambaum, "TaN molecule as a candidate for the search for a T, P-violating nuclear magnetic quadrupole moment", Physical Review A **92**, 012521 (2015).

[24]I. Kozyryev and N. R. Hutzler, "Precision measurement of time-reversal symmetry violation with laser-cooled polyatomic molecules", Physical Review Letters **119**, 133002 (2017).

[25]A. Singh, L. Maisenbacher, Z. Lin, J. J. Axelrod, C. D. Panda, and H. Müller, "Dynamics of a buffer-gas-loaded, deep optical trap for molecules", Physical Review Research **5**, 033008 (2023).

[26]S. K. Lamoreaux, J. P. Jacobs, B. R. Heckel, F. J. Raab, and N. Fortson, "New constraints on time-reversal asymmetry from a search for a permanent electric dipole moment of ^{199}Hg", Physical Review Letters **59**, 2275 (1987).

[27]A.-M. Mårtensson-Pendrill and P. Öster, "Calculations of atomic electric dipole moments", Physica Scripta **36**, 444 (1987).

[28]W. C. Griffith, M. D. Swallows, T. H. Loftus, M. V. Romalis, B. R. Heckel, and E. N. Fortson, "Improved limit on the permanent electric dipole moment of ^{199}Hg", Physical Review Letters **102**, 101601 (2009).

[29]B. Graner, Y. Chen, E. G. Lindahl, and B. R. Heckel, "Reduced limit on the permanent electric dipole moment of ^{199}Hg", Physical Review Letters **116**, 161601 (2016).

[30]M. G. Kozlov and L. N. Labzowsky, "Parity violation effects in diatomics", Journal of Physics B: Atomic, Molecular and Optical Physics **28**, 1933 (1995).

[31]Y. Y. Dmitriev, Y. G. Khait, M. G. Kozlov, L. N. Labzovsky, A. O. Mitrushenkov, A. V. Shtoff, and A. V. Titov, "Calculation of the spin-rotational hamiltonian including P-and P, T-odd weak interaction terms for HgF and PbF molecules", Physics Letters A **167**, 280–286 (1992).

[32]E. R. Meyer and J. L. Bohn, "Prospects for an electron electric-dipole moment search in metastable ThO and Th F^+", Physical Review A **78**, 010502 (2008).

[33]Z. Yang, J. Li, Q. Lin, L. Xu, H. Wang, T. Yang, and J. Yin, "Laser-cooled HgF as a promising candidate to measure the electric dipole moment of the electron", Physical Review A **99**, 032502 (2019).

[34]S. Elmoussaoui and W. Chmaisani, "Theoretical electronic structure and rovibrational calculations with spin-orbit effect of the HgCl low-lying electronic states", Chemical Physics Letters **744**, 137209 (2020).

[35]P. Pyykko, "Relativistic effects in structural chemistry", Chemical Reviews **88**, 563–594 (1988).

[36]V. S. Prasannaa, A. C. Vutha, M. Abe, and B. P. Das, "Mercury monohalides: Suitability for electron electric dipole moment searches", Physical Review Letters **114**, 183001 (2015).

[37]V. S. Prasannaa, M. Abe, V. M. Bannur, and B. P. Das, "Theoretical analysis of effective electric fields in mercury monohalides", Physical Review A **95**, 042513 (2017).

[38]T. Fleig, "TaO+ as a candidate molecular ion for searches of physics beyond the standard model", Physical Review A **95**, 022504 (2017).

[39]Design and F. of an Apparatus for TaO High-Resolution Laser Spectroscopy, "Chung, timothy", MA thesis (University of Nevada, Las Vegas, 2022).

[40]T. Fleig, M. K. Nayak, and M. G. Kozlov, "TaN, a molecular system for probing P, T-violating hadron physics", Physical Review A **93**, 012505 (2016).

[41]V. V. Flambaum, D. DeMille, and M. G. Kozlov, "Time-reversal symmetry violation in molecules induced by nuclear magnetic quadrupole moments", Physical Review Letters **113**, 103003 (2014).

[42]A. A. Kyuberis, L. F. Pašteka, E. Eliav, H. A. Perrett, A. Sunaga, S. M. Udrescu, S. G. Wilkins, R. F. Garcia Ruiz, and A. Borschevsky, "Theoretical determination of the ionization potentials of CaF, SrF, and BaF", Physical Review A **109**, 022813 (2024).

[43]J. E. Sansonetti and W. C. Martin, *NIST Chemistry Webbook, NIST Standard Reference Database Number 108. Handbook of basic atomic spectroscopic data* (National Institute of Standards and Technology, 2013).

[44]P. J. Mohr and B. N. Taylor, "CODATA recommended values of the fundamental physical constants: 1998", Reviews of modern physics **72**, 351 (2000).

[45]H. Nakatsuji, "Discovery of a general method of solving the schrodinger and dirac equations that opens a way to accurately predictive quantum chemistry", Accounts of Chemical Research **45**, 1480–1490 (2012).

[46]J. Autschbach, "Perspective: relativistic effects", The Journal of chemical physics **136** (2012).

[47]D. Doeglas, "Relativistic coupled cluster calculations of molecular properties of BaF and PbF required in high-precision experiments", PhD thesis (University of Groningen, 2019).

[48]C. D. Sherrill, "An introduction to Hartree-Fock molecular orbital theory", School of Chemistry and Biochemistry Georgia Institute of Technology (2000).

[49]T. Saue, R. Bast, A. S. P. Gomes, H. J. A. Jensen, L. Visscher, I. A. Aucar, R. Di Remigio, K. G. Dyall, E. Eliav, E. Fasshauer, et al., "The DIRAC code for relativistic molecular calculations", The Journal of chemical physics **152** (2020).

[50]K. G. Dyall and K. Fægri, *Introduction to relativistic quantum chemistry* (Oxford University Press, 2007).
[51]K. G. Dyall, "Interfacing relativistic and nonrelativistic methods. I. normalized elimination of the small component in the modified Dirac equation", The Journal of chemical physics **106**, 9618–9626 (1997).
[52]M. Filatov and K. G. Dyall, "On convergence of the normalized elimination of the small component (NESC) method", Theoretical Chemistry Accounts **117**, 333–338 (2007).
[53]K. G. Dyall, "An exact separation of the spin-free and spin-dependent terms of the Dirac-Coulomb-Breit Hamiltonian", The Journal of chemical physics **100**, 2118–2127 (1994).
[54]M. Iliaš and T. Saue, "An infinite-order two-component relativistic Hamiltonian by a simple one-step transformation", The Journal of chemical physics **126** (2007).
[55]W. Liu and D. Peng, "Exact two-component Hamiltonians revisited", The Journal of chemical physics **131** (2009).
[56]P. W. Atkins and R. S. Friedman, *Molecular quantum mechanics* (Oxford university press, 2011).
[57]J. C. Slater, "The theory of complex spectra", Physical Review **34**, 1293 (1929).
[58]T. Tsuneda, "Hartree-Fock method", in *Density functional theory in quantum chemistry* (Springer, 2014), pp. 35–63.
[59]H. Sekino and R. J. Bartlett, "On the extensivity problem in coupled-cluster property evaluation", in *Advances in quantum chemistry*, Vol. 35 (Elsevier, 1999), pp. 149–173.
[60]K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A fifth-order perturbation comparison of electron correlation theories", Chemical Physics Letters **157**, 479–483 (1989).
[61]M. Abe, G. Gopakumar, M. Hada, B. P. Das, H. Tatewaki, and D. Mukherjee, "Application of relativistic coupled-cluster theory to the effective electric field in YbF", Physical Review A **90**, 022501 (2014).
[62]W. Kutzelnigg, "The relativistic many body problem in molecular theory", Physica Scripta **36**, 416 (1987).
[63]D. Boer, "Lecture notes on Lie Groups in Physics", Faculty of Science and Engineering of the University of Groningen (2023-2024).
[64]J. Rezac, L. Simova, and P. Hobza, "CCSD[T] describes noncovalent interactions better than the CCSD(T), CCSD(TQ), and CCSDT methods", Journal of chemical theory and computation **9**, 364–369 (2013).
[65]K. G. Dyall, "Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the 5 d elements Hf–Hg", Theoretical Chemistry Accounts **112**, 403–409 (2004).
[66]K. G. Dyall, "Relativistic quadruple-zeta and revised triple-zeta and double-zeta basis sets for the 4p, 5p, and 6p elements", Theoretical Chemistry Accounts **115**, 441–447 (2006).
[67]K. G. Dyall, "Relativistic double-zeta, triple-zeta, and quadruple-zeta basis sets for the light elements H–Ar", Theoretical Chemistry Accounts **135**, 128 (2016).

[68]D. Feller, "The use of systematic sequences of wave functions for estimating the complete basis set, full configuration interaction limit in water", The Journal of chemical physics **98**, 7059–7071 (1993).
[69]V. Vasilyev, "Online complete basis set limit extrapolation calculator", Computational and Theoretical Chemistry **1115**, 1–3 (2017).
[70]T. Helgaker, W. Klopper, H. Koch, and J. Noga, "Basis-set convergence of correlated calculations on water", The Journal of chemical physics **106**, 9639–9646 (1997).
[71]A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, "Basis-set convergence in correlated calculations on Ne, N_2, and H_2O", Chemical Physics Letters **286**, 243–252 (1998).
[72]Center for Information Technology at the University of Groningen provided access to the Hábrók high performance computing cluster.
[73]Python Software Foundation. Python Language Reference, version 12.7. Available at http://www.python.org.
[74]NumPy library for scientific computing in Python. Documentation available at https://numpy.org/doc/stable/index.html.
[75]Matplotlib library for visualisation in Python. Documentationavailable at https://matplotlib.org/stable/.
[76]DIRAC software reference and manual, available at https://www.diracprogram.org/doc/release-19/index.html.
[77]V. S. Prasannaa, B. K. Sahoo, M. Abe, and B. P. Das, "Significance of non-linear terms in the relativistic coupled-cluster theory in the determination of molecular properties", Symmetry **12**, 811 (2020).
[78]R. W. Kiser, J. G. Dillard, and D. L. Dugger, "Mass spectrometry of inorganic halides", in *Mass spectrometry in inorganic chemistry* (ACS Publications, 1966).
[79]H. M. Rosenstock, K. Draxl, B. W. Steiner, and J. T. Herron, *Energetics of gaseous ions*, Vol. 6, Supplement No. 1 (Journal of Physical and Chemical Reference Data, 1977).
[80]D. Cremer, E. Kraka, and M. Filatov, "Bonding in mercury molecules described by the normalized elimination of the small component and coupled cluster theory", ChemPhysChem **9**, 2510–2521 (2008).
[81]N. B. Balabanov and K. A. Peterson, "Mercury and reactive halogens: the thermochemistry of Hg^+Cl_2, Br_2, BrCl, ClO, and BrO", The Journal of Physical Chemistry A **107**, 7465–7470 (2003).
[82]D. Majumdar, S. Roszak, and J. Leszczynski, "Probing the structures and thermodynamic characteristics of the environment polluting mercuric halides, cyanides and thiocyanates", Chemical Physics Letters **501**, 308–314 (2011).
[83]L. R. Maxwell and V. M. Mosley, "Internuclear distances in Se_2, Te_2, and HgCl by electron diffraction", Physical Review **57**, 21 (1940).
[84]N.-H. Cheung and T. A. Cool, "Franck-Condon factors and r-centroids for the $B^2\Sigma$-$X^2\Sigma$ systems of HgCl, HgBr, and HgI", Journal of Quantitative Spectroscopy and Radiative Transfer **21**, 397–432 (1979).
[85]A. K. Rai, S. B. Rai, and D. K. Rai, "Spectral study of the DX system of the diatomic mercury chloride molecule", Journal of Physics B: Atomic and Molecular Physics **15**, 3239 (1982).

[86]F. Zhang, D. Oba, and D. W. Setser, "A flowing-afterglow study of the quenching reactions of mercury ($^3P_2$) and mercury ($^3P_0$) atoms by halogens, interhalogens, and polyatomic halide molecules", Journal of Physical Chemistry **91**, 1099–1114 (1987).

[87]S. Zhao, X. Mo, R. Li, H. Guo, and B. Yan, "Extensive spin–orbit multi-reference study on low-lying electronic states of HgI", Journal of Quantitative Spectroscopy and Radiative Transfer **277**, 107993 (2022).

[88]R. S. Ram, J. Liévin, and P. F. Bernath, "Emission spectroscopy and ab initio calculations for TaN", Journal of Molecular Spectroscopy **215**, 275–284 (2002).

[89]T. Steimle, D. L. Kokkin, Y. Kim, R. J. Mawhorter, and C. Linton, "Characterization of the [18.42] 0^+–$X^1\Sigma^+$(0,0) band of tantalum nitride, TaN", Chemical Physics Letters **664**, 138–142 (2016).

[90]X. Chen, S. Yan, R. Zhang, and C. Ning, "A high-resolution photoelectron spectroscopic and computational study of TaX^-(X= C, N, O)", The Journal of Chemical Physics **161** (2024).

[91]W. R. Wadt, "The electronic structure of HgCl and HgBr", Applied Physics Letters **34**, 658–660 (1979).

[92]D. A. Kleier and W. R. Wadt, "Molecular structure of mercurous halides: mercurous fluoride and mercurous chloride", Journal of the American Chemical Society **102**, 6909–6913 (1980).

[93]J. Barbe, "Convenient relations for the estimation of bond ionicity in AB type compounds", Journal of Chemical Education **60**, 640 (1983).

[94]P. Schwerdtfeger, P. D. W. Boyd, S. Brienne, J. S. McFeaters, M. Dolg, M.-S. Liao, and W. H. E. Schwarz, "The mercury-mercury bond in inorganic and organometallic compounds. A theoretical study", Inorganica chimica acta **213**, 233–246 (1993).

[95]M. Kaupp and H. G. von Schnering, "Origin of the unique stability of condensed-phase Hg_2^{2+}. An ab initio investigation of M^I and M^{II} species (M= Zn, Cd, Hg)", Inorganic Chemistry **33**, 4179–4185 (1994).

[96]M.-S. Liao, Q.-E. Zhang, and W. H. E. Schwarz, "Properties and stabilities of MX, MX_2, and M_2X_2 compounds (M= Zn, Cd, Hg; X= F, Cl, Br, I)", Inorganic Chemistry **34**, 5597–5605 (1995).

[97]S. Gnanakaran and R. M. Hochstrasser, "Vibrational relaxation of HgI in ethanol: Equilibrium molecular dynamics simulations", The Journal of chemical physics **105**, 3486–3496 (1996).

[98]A. F. Khalizov, B. Viswanathan, P. Larregaray, and P. A. Ariya, "A theoretical study on the reactions of Hg with halogens: atmospheric implications", The Journal of Physical Chemistry A **107**, 6360–6365 (2003).

[99]J. A. Tossell, "Calculation of the energetics for oxidation of gas-phase elemental Hg by Br and BrO", The Journal of physical chemistry A **107**, 7804–7808 (2003).

[100]B. C. Shepler, N. B. Balabanov, and K. A. Peterson, "Ab initio thermochemistry involving heavy atoms: An investigation of the reactions Hg^+ IX (X= I, Br, Cl, O)", The Journal of Physical Chemistry A **109**, 10363–10372 (2005).

[101]V. M. Kovba and E. A. Pazyuk, "Calculations of the electronic states of the HgI and HgTe molecules by the SCF-Xα-SW method", Russian Journal of Physical Chemistry **80**, 1285–1288 (2006).

[102]E. Kraka, M. Filatov, and D. Cremer, "Comparison of gold bonding with mercury bonding", Croatica Chemica Acta **82**, 233–243 (2009).

[103]J. Kim, H. Ihee, and Y. S. Lee, "Spin-orbit density functional and ab initio study of HgX_n (X= F, Cl, Br, and I; n= 1, 2, and 4)", The Journal of chemical physics **133** (2010).

[104]W. Zou, M. Filatov, and D. Cremer, "Exploring bonding in heavy atom chemistry with Dirac-exact methods", Current Inorganic Chemistry (Discontinued) **3**, 220–234 (2013).

[105]M. Abe, V. S. Prasannaa, and B. P. Das, "Application of the finite-field coupled-cluster method to calculate molecular properties relevant to electron electric-dipole-moment searches", Physical Review A **97**, 032515 (2018).

[106]M. Zhou and L. Andrews, "Reactions of laser-ablated niobium, tantalum, and rhenium atoms with nitrogen atoms and molecules. Infrared spectra and density functional calculations of the metal nitride and dinitride molecules", The Journal of Physical Chemistry A **102**, 9061–9071 (1998).

Appendix

A Abbreviations glossary

This overview includes abbreviations used throughout the report, as well as abbreviations used in Tables B.1 till B.4 in Appendix B.

- 2C = two-component, referring to the type of Hamiltonian
  - X2C = exact two-component method, described in section 2.1.4
  - 4C = four-component, mostly referring to the relativistic Dirac-Coulomb Hamiltonian, described in section 2.1.3
- AVTZ-PP = aug-cc-p-VTZ-PP basis set, relativistic effective core potentials (REPs)
- B3LYP = a hybrid functional added to DFT, less reliable than CCSD(T)
- CASSCF = complete active-space self-consistent field (method)
- CBS = complete basis set, generally extrapolated from multiple calculations
- CC = coupled cluster, an electron correlation method described in section 2.3.1
  - CCSD = coupled cluster, with single and double excitations, including some cross terms. Described in section 2.3.2
  - CCSD(T) = coupled cluster, CCSD with added perturbative triple excitations. Described in section 2.3.2
  - CCSDT = coupled cluster with single, double, and triple excitations (non-perturbative triples)
- CI = Configuration Interaction
  - CISD = configuration interaction
  - MRCI = multi-reference configuration interaction
  - MRCI+Q = MRCI with the Davidson correction
  - QCISD = CISD corrected for size consistency
- CRENBL = a large, valence, shape-consistent method
- CV = core valence, used in basis sets (added core valence functions)
  - QCISD = quadratic CI with single and double substitutions
- DC = Dirac-Coulomb, referring to the 4 component Hamiltonian, explained in 2.1.3
- DFT = density functional theory
- (D)HF = (Dirac-)Hartree-Fock, explained in section 2.2.3
- ECP = effective core potential
  - RECCP = relativistic effective core core potential
  - NRECP = non-relativistic effective core potential
- EDM = electric dipole moment, or eEDM for electron EDM, explained in section 1.2
- FFCC = finite field perturbation theory adapted to relativistic CCSD method
- GTO = Gaussian type orbital, explained in section 2.4.1
- H = Hamiltonian
- HgX = mercury halides HgCl and HgI. When referring to additional molecules not in this paper this also includes HgF and HgBr

- HOMO = highest occupied molecular orbital
- IVO = improved virtual orbital
- LDA (or LDF) = local density functional
- LUMO = lowest unoccupied molecular orbital
- M06 = hybrid-meta, a functional more appropriate for estimating thermochemical energies. M06-L is known as the meta-generalized gradient approximation (GGA)
- MP2 = Moller-Plesset many-body perturbation theory at second order, a method of adding electron correlation through the Rayleigh-Schrödinger perturbation theory
- NESC = normalised elimination of the small component
- PBE0 = hybrid, a functional generally reasonable for molecular geometries and vibrational frequencies
- PEC(s) = potential energy curve(s)
- P/ POL = polarization, used in basis sets (added polarization functions), or in general techniques (like POL-CI, being polarization CI)
- PP = pseudopotential, RPP = relativistic pseudopotential
- RKR = Rydberg-Klein-Rees method, generally used for constructions potential curves
- RCCSD = relativistic CCSD
  - LERCCSD = linear approximation of RCCSD
  - nLERCCSD = including non-linear terms in RCCSD
- SCF = self consistent field, in this work also used to refer to Dirac-Hartree-Fock calculations
- SM = the standard model, mentioned in section 1.1
- SO/ SOC = spin-orbit coupling (effects)
- SR = scalar relativity
- STO = Slater type orbital, explained in section 2.4.1
- XZ = (where X is D, T, Q, or a number) X-zeta, in reference to basis set with cardinality X

B Literature results HgX and TaN

HgCl IP **[eV]**	**HgI** IP **[eV]**	**Method**	**Extra detail**	**Year**	**Source**
12.06 ± 0.26	11.3 ± 0.4	Experimental*	mass spectrometry from $HgX_2 \rightarrow HgX^+ + X$	1966	[78]
12.06 ± 0.26	11.3 ± 0.4	Experimental*	photoionisation data from $HgX_2 \rightarrow HgX^+ + X$, cites [78]	1977	[79]
9.518	8.859	theory; NESC	DFT B3LYP; basis: [15s13p8d5f] DZ(core) TZ (val) (Hg), Dunning aug-cc-pVTZ (Cl), Dyall [12s11p7d2f] (I)	2008	[80]
9.258	8.581	theory; NESC	CCSD(T); basis: [15s13p8d5f] DZ(core) TZ (val) (Hg), Dunning aug-cc-pVTZ (Cl), Dyall [12s11p7d2f] (I)	2008	[80]

val. is short for valence region. Abbreviations can be found in Appendix A

*Values are expected to differ significantly from the actual IPs of HgX molecules, but mentioned here as the source is often referenced in literature

Table B.1: HgCl and HgI ionisation potential (IP)

HgCl R_e [Å]	**HgI R_e [Å]**	**Method**	**Extra details**	**Source**	
2.23 ± 0.03	-	Experimental	(HgCl) electron diffraction photographs	1940	[83]
2.27 ± 0.03	-	Experimental	(Hg-Cl from $HgCl_2$) electron diffraction photographs	1940	[83]
2.532 ± 0.04	2.443 ± 0.04	Experimental	Simulated emission intensity distribution fit with electron diffraction data	1979	[84]
2.3948	-	Experimental	Photographing 3rd order concave grating spectrograph	1982	[85]
2.419 - 2.434[@]	2.36 - 2.81[@]	Experimental	RKR potentials fitted to chemiluminescent spectra	1987	[86]
2.42 ± 0.04	-	theory; RECCP (Hg), NRECP (Cl)	SCF with IVO; basis: DZ+P	1979	[91]
2.39	-	theory; with GVBTWO prog.	SCF; basis: DZ+P (Hg), DZ+P s-aug (Cl)	1980	[92]
2.41	-	theory; with GVBTWO prog.	SCF POL-CI; basis: RECP (Hg), NRCP (Cl)	1980	[92]
2.54	-	theory	Bond length and ionicity relation equation	1983	[93]
2.504	-	theory; pseudovalence H	HF; basis: 9s7p6d1f (Hg), [Ne]-core SZ core+DZ val.+3d(1.8Z) d-pol. (Cl)	1993	[94]
2.450	-	theory; pseudovalence H	MP2; basis: 9s7p6d1f (Hg), [Ne]-core SZ core+DZ val.+3d(1.8Z) d-pol. (Cl)	1993	[94]
2.50	-	theory; pseudovalence H	LDA; basis: 9s7p6d1f (Hg), [Ne]-core SZ core+DZ val.+3d(1.8Z) d-pol. (Cl)	1993	[94]
2.441	-	theory; quasi-rel. Hg	HF; basis: segm. (8s7p6d)/[6s5p3d] val (Hg PP), segm. (5s5p1d)/[3s3p1d] val (Cl)	1994	[95]
2.408	-	theory; quasi-rel. Hg	MP2; basis: segm. (8s7p6d)/[6s5p3d] val (Hg PP), segm. (5s5p1d)/[3s3p1d] val (Cl)	1994	[95]
2.46	2.80	theory	LDF rel. DFT; basis: STO V3Z aug 2p pol (Hg), v2z aug-d-pol (Cl&I)	1995	[96]
2.408	-	theory	PP rel. DFT; basis: STO V3Z aug 2p pol (Hg), v2z aug-d-pol (Cl&I)	1995	[96]
-	2.8	theory; with GAUSSIAN92 prog.	RECP; basis: v2z-p-pol (Hg), v2z-dp-pol (I)	1996	[97]
2.612	-	theory; with GAUSSIAN98 A.7 prog.	B3LYP; basis: LanL2Dz (Hg&Cl); ECP+VDZ	2003	[98]
2.455	-	theory; with GAUSSIAN98 A.7 prog.	B3LYP; basis: EC60MWB(9s9p6d4f) (Hg), aug-cc-pVQZ (Cl)	2003	[98]
2.433	-	theory; with GAUSSIAN98 A.7 prog.	QCISD; basis: LanL2Dz (Hg), 6-311G(2df) (Cl)	2003	[98]
2.398	-	theory; with GAUSSIAN98 A.7 prog.	QCISD; basis: EC60MWB(9s9p6d4f) (Hg), 6-311G(2df) (Cl)	2003	[98]
2.394	-	theory; using RECPs	QCISD; basis: RECPs, valence electron only, +polarisation functions	2003	[99]
2.376	-	theory; using RECPs	CCD; basis: RECPs, valence electron only, +polarisation functions	2003	[99]

prog. is short for programme, H for Hamiltonian, rel. for relativistic. Abbreviations can be found in Appendix A

[@]Minimum-maximum range

Table B.2: HgCl and HgI equilibrium bond distances

HgCl R_e [Å]	**HgI R_e [Å]**	**Method**	**Extra details**	**Source**	
2.3870	-	theory; Douglas-Kroll-Hess H	CCSD(T); basis: AVQZ	2003	[81]
2.3824	-	theory; Douglas-Kroll-Hess H	CCSD(T); basis: AV5Z	2003	[81]
2.3798	-	theory; Douglas-Kroll-Hess H	CCSD(T); basis: CBS1, 3-point formula	2003	[81]
2.3777	-	theory; Douglas-Kroll-Hess H	CCSD(T); basis: CBS2, 2-point formula	2003	[81]
2.3541	2.7075	theory; RPP	CCSD(T), SO+SR, with MOLPRO; basis: CBS+CV	2005	[100]
-	2.75	theory; quasi-rel.	SCF-X_α-SW; 30% overlapping atomic spheres	2006	[101]
2.460	2.820	theory; NESC	DFT B3LYP; basis: DZc+TZv (Hg), aug-cc-pVTZ (Cl), (22s16p12d2f)[12s11p7d2f] (I)	2008	[80]
2.274 ($HgCl^+$)	2.598 (HgI^+)	theory; NESC	DFT B3LYP; basis: DZc+TZv (Hg), aug-cc-pVTZ (Cl), (22s16p12d2f)[12s11p7d2f] (I)	2008	[80]
2.461	2.820	theory; rel. NESC/B3LYP DFT	basis: Dyall VDZ core VTZ val. (Hg), Dunning aug-cc-pVTZ (Cl), Dyall (22s16p12d2f)[12s11p7d2f] (I)	2009	[102]
2.274 ($HgCl^+$)	2.598 (HgI^+)	theory; rel. NESC/B3LYP DFT	basis: Dyall VDZ core VTZ val. (Hg), Dunning aug-cc-pVTZ (Cl), Dyall (22s16p12d2f)[12s11p7d2f] (I)	2009	[102]
2.364	2.793	theory; 2C SO	DFT CCSD(T)*; basis: RECP CRENBL	2010	[103]
2.392	2.744	theory; 2C SO	DFT CCSD(T)*; basis: RECP AVTZ-PP	2010	[103]
2.394	2.739	theory; rel.	CCSD; basis: rel. SBKJ VDZ ECP, + aug-f& g pol. (Hg), or + aug-d& f pol. func. (Cl, I)	2011	[82]
2.258 ($HgCl^+$)	2.566 (HgI^+)	theory; rel.	CCSD; basis: rel. SBKJ VDZ ECP, + aug-f& g pol. (Hg), or + aug-d& f pol. func. (Cl, I)	2011	[82]
2.389	2.736	theory; NESC	PBE0 functional; basis: def2-QZVPP for Cl, SARC for Hg & I**	2013	[104]
2.331	-	theory; quasi-rel. ECP	MRCI+Q, CASSCF; basis: ECP78MWB (Hg), ECP10MWB (Cl)	2020	[34]
-	2.7642	theory; quasi-rel., with MOLPRO[!]	MRCI+Q + SOC with SD; basis: cc-pwCV5z-PP, with ECP60MDF (Hg), or with ECP46MDF (I)	2022	[87]

*Multiple DFT functionals were used (PBE0 and M06, and some corrections for these), but CCSD(T) was the reference value

**Recontracted basis sets

[!]The MOLPRO2012 programme that uses molecular valence model Hamiltonian, quasi-rel. ECP with mass-velocity and Darwin effects in 1C term and spin orbit effects in 2C term

Table B.2: (continued) HgCl and HgI equilibrium bond distances

HgCl E_{eff} [GV/cm]	**HgI E_{eff} [GV/cm]**	**Hamiltonian**	**Extra details**	**Year**	**Source**
103.57	96.85	eff. eEDM H	DF; basis: CCPVDZ for Cl, Dyall's c3v for Hg & I	2015	[36]
113.56	109.30	4C H	CCSD, accuracy of 5%; basis: 22s19p12d9f1g (Hg), 12s8p1d (Cl), 21s15p11d (I)	2015	[36]
114.31	109.56	4C H	FFCC CCSD; basiss: Dyall cDZV (Hg), cc-pVDZ (Cl, I)	2018	[105]
104.33	99.27	4C H	DF; basis: Dyall QZ (Hg, I), Dunning cc-VQZ (Cl); cut-off virtuals at 1000 a.u.	2020	[77]
112.51	110.00	4C H	LERCCSD; basis: Dyall QZ (Hg, I), Dunning cc-VQZ (Cl); cut-off virtuals at 1000 a.u.	2020	[77]
110.94	107.38	4C H	nLERCCSD; basis: Dyall QZ (Hg, I), Dunning cc-VQZ (Cl); cut-off virtuals at 1000 a.u.	2020	[77]

Abbreviations can be found in Appendix A

Table B.3: Theoretically determined E_{eff} [GV/cm] for HgCl and HgI

TaN	Type	Method	Details	Year	Source
1.6830999(88)	R_e (Å)	Experimental	Spectrometry of emission from tantalum hollow-cathode lamp	2002	[88]
1.685	R_e (Å)	Experimental	high-res. laser-induced fluorescence spectra of the $[18.42]0^+ - X^1\Sigma^+$ band of TaN	2016	[89]
1.690	R_e (Å)	Theory; NR	DFT BP86 functional; basis: Los Alamos ECP+Dz (Ta), D95* (N)	1998	[106]
1.706	R_e (Å)	Theory; quasi-relativistic	CASSCF/CMRCI; basis: Wood-Boring PPs, v2z +1 Gaussian pol. orbital per atom	2002	[88]
1.69*	R_e (Å)	Theory; GRECP, 2C	CCSD(T); basis: MBas	2015	[23]
1.6831	R_e (Å)	Theory; DC H	HF+MR_{12}-CISD(18)+T$^{@}$; basis: Dyall v3z+val.&core-correlating exponents (Ta), Dunning cc-pVTZ-DK (N)	2016	[40]
1.710	R_e (Å)	Theory; RECP	using Gaussian09 prog.; CCSD(T); basis: CBS limit extrapolation	2024	[90]
34.9*	E_{eff} (GV/cm)	Theory; GRECP, 2C	CCSD(T)**; basis: 15s10p10d5f2g uncontracted Gaussian (Ta), aug-ccpVQZ (N)	2015	[23]
36.0*	E_{eff} (GV/cm)	Theory; DC H	HF+MR_{12}^{+T}-CISD(18)$^{@}$ + external triples + Δ spinors + 4s,4p core-valence corrections; basis: see above	2016	[40]

pol. short for polarisation, prog. for programme. Abbreviations can be found in Appendix A

*For $^3\Delta_1$ state of TaN

**Result is a combination of CCSD, CCSD(T), and correlation, basis set & core corrections

$^{@}$The 12 in MR_{12} refers to the number of Kramers pairs, the 18 in CISD(18) to the accumulated number of electrons

Table B.4: TaN literature overview of R_e, E_{eff}

C Complete calculation results HgX and TaN

C.1 Equilibrium bond lengths

basis set	HgCl R_e [Å]			$HgCl^+$ R_e [Å]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
saugav2z	2.400	2.402	2.407	2.269	2.262	2.269
saugv3z	2.388	2.373	2.378	2.258	2.235	2.241
saugav3z	2.388	2.371	2.376	2.258	2.235	2.241
v4z	2.385	2.359	2.362	2.255	2.224	2.228
saugv4z	2.384	2.357	2.360	2.253	2.221	2.226
av4z	2.385	2.358	2.361	2.254	2.222	2.226
saugav4z	2.384	2.357	2.360	2.253	2.221	2.225
CBS (saugavXz)	2.385	2.351	2.353	2.254	2.2157	2.219

Table C.1: HgCl and $HgCl^+$ optimised bond lengths per basis set, with the CBS limit extrapolation based on the saugav2z, saugav3z, and saugav4z results

basis set	HgI R_e [Å]			HgI^+ R_e [Å]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
saugav2z	2.728	2.778	2.784	2.615	2.606	2.612
saugv3z	2.718	2.732	2.740	2.604	2.572	2.577
saugav3z	2.718	2.738	2.748	2.603	2.575	2.580
v4z	2.715	2.716	2.722	2.600	2.556	2.560
saugv4z	2.714	2.715	2.721	2.599	2.555	2.559
av4z	2.714	2.719	2.725	2.599	2.558	2.562
saugav4z	2.713	2.716	2.722	2.598	2.556	2.560
CBS (saugavXz)	2.714	2.708	2.712	2.599	2.549	2.552

Table C.2: HgI and HgI^+ optimised bond lengths per basis set, with the CBS limit extrapolation based on the saugav2z, saugav3z, and saugav4z results

basis set	TaN R_e [Å]			TaN^+ R_e [Å]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
av2z	1.639	1.676	1.696	1.616	1.656	1.675
v3z	1.636	1.669	1.689	1.612	1.649	1.668
saugv3z	1.635	1.669	1.689	1.612	1.648	1.667
av3z	1.635	1.668	1.689	1.612	1.648	1.667
v4z	1.633	1.659	1.679	1.610	1.639	1.657
av4z	1.632	1.658	1.677	1.609	1.637	1.656
CBS (avXz)	1.632	1.651	1.671	1.608	1.631	1.649

Table C.3: TaN and TaN^+ optimised bond lengths per basis set, with the CBS limit extrapolation based on the av2z, av3z, and av4z results

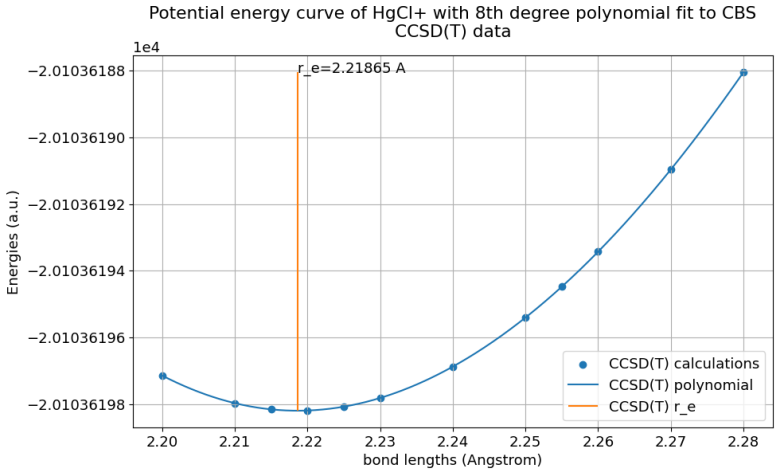


Figure C.1: $HgCl^+$ equilibrium bond length of CBS CCSD(T)

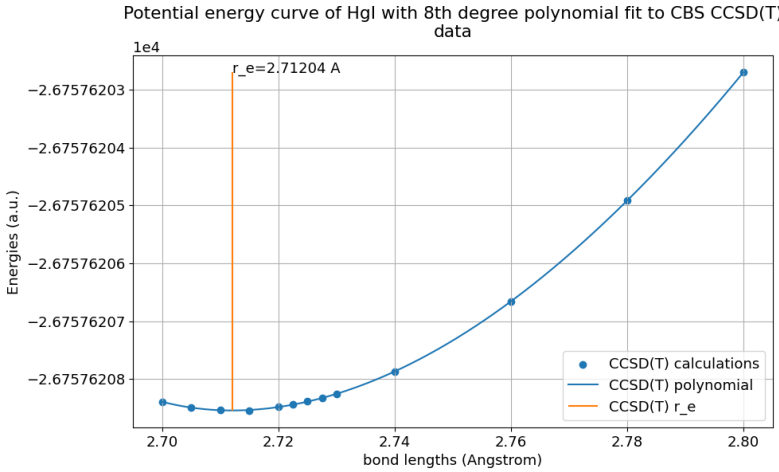


Figure C.2: HgI equilibrium bond length of CBS CCSD(T)

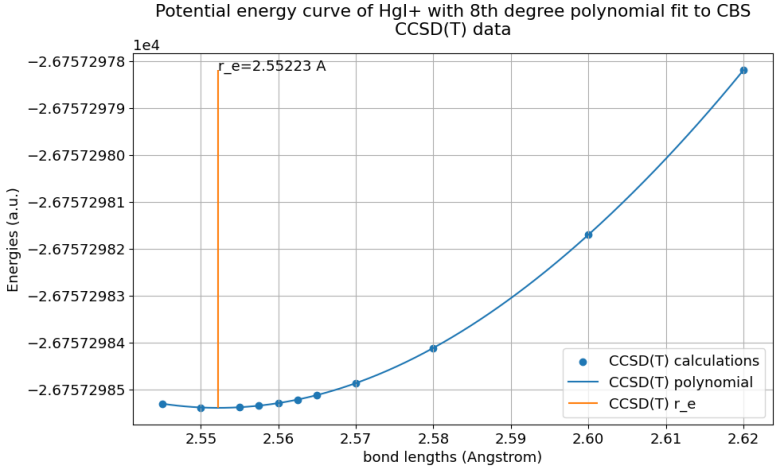


Figure C.3: HgI^+ equilibrium bond length of CBS CCSD(T)

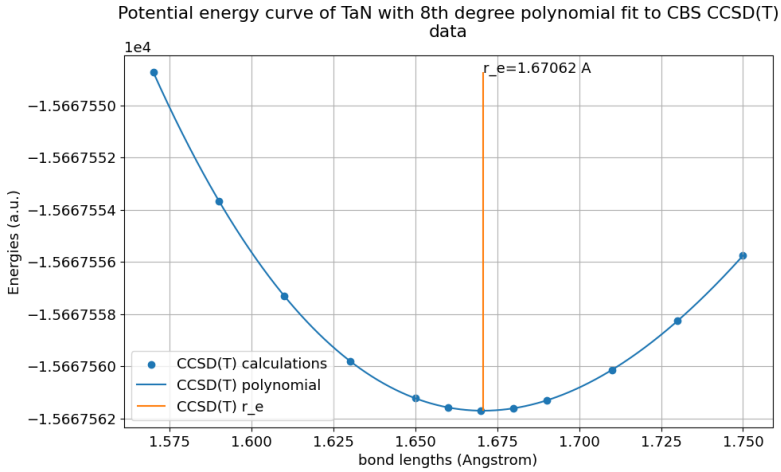


Figure C.4: TaN equilibrium bond length of CBS CCSD(T)

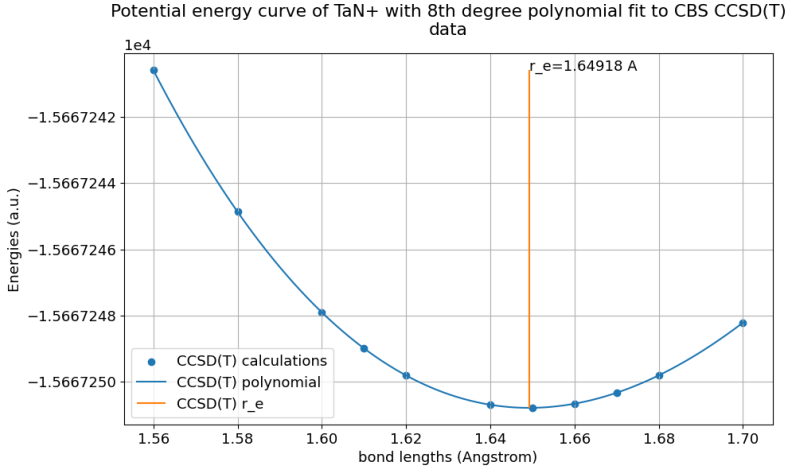


Figure C.5: TaN^+ equilibrium bond length of CBS CCSD(T)

C.2 Ionisation potentials

basis set	HgCl IP [eV]			HgI IP [eV]		
	DHF	CCSD	CCSD(T)	DHF	CCSD	CCSD(T)
saugav2z	8.669	9.339	9.347	8.034	8.622	8.633
saugv3z	8.593	9.328	9.355	7.968	8.640	8.668
saugav3z	8.593	9.330	9.356	7.964	8.646	8.675
v4z	8.583	9.350	9.379	7.955	8.676	8.706
saugv4z	8.580	9.358	9.390	7.954	8.686	8.722
av4z	8.581	9.359	9.391	7.954	8.690	8.722
saugav4z	8.580	9.359	9.391	7.952	8.691	8.723
CBS (saugavXz)	8.583	9.392	9.428	7.955	8.736	8.771

Table C.4: HgCl and HgI ionisation potentials per basis set, with the CBS limit extrapolation based on the saugav2z, saugav3z, and saugav4z results

TaN IP [eV] / basis set	DHF	CCSD	CCSD(T)
av2z	7.192	8.230	8.406
v3z	7.178	8.262	8.451
saugv3z	7.153	8.247	8.439
av3z	7.177	8.270	8.462
v4z	7.175	8.260	8.456
av4z	7.174	8.263	8.460
CBS (avXz)	7.173	8.259	8.461

Table C.5: TaN ionisation potentials per basis set, with the CBS limit extrapolation based on the av2z, av3z, and av4z results

C.3 Uncertainty calculations

setting 1	IP [eV]	setting 2	IP [eV]	difference [eV]	difference [%]
Basis set quality (base setting: X2C CCSD(T))					
saugav4z	9.3909	CBS	9.4283	0.0375	0.3989
v4z	9.3789	av4z	9.3905	0.0117	0.1248
av4z	9.3905	saugav4z	9.3909	0.0004	0.0038
Correlation - active space (base setting: X2C CCSD(T))					
v3z (-30 to 30 a.u.)	9.3158	ae3z (-3300 to 3300 a.u.)	9.3287	0.0129	0.1387
Correlation - excitations (base setting: X2C saugav4z)					
CCSD	9.3924	CCSD(T)	9.4283	0.0359	0.3820
Level of relativity (base setting: v3z)					
X2C CCSD(T)	9.3158	4C DC CCSD(T)	9.3231	0.0073	0.0787
4C DC DHF	8.6044	4C DC+Gaunt DHF	8.5967	-0.0078	-0.0898
X2C DHF	8.5982	4C DC+Gaunt DHF	8.5967	-0.0016	-0.0181

Table C.6: Differences in ionisation potentials [eV] of HgCl for different settings, at bond lengths $R_e^n = 2.35259$ Å and $R_e^+ = 2.21865$ Å

setting 1	IP [eV]	setting 2	IP [eV]	difference [eV]	difference [%]
Basis set quality (base setting: X2C CCSD(T))					
saugav4z	8.7229	CBS	8.7706	0.0477	0.5473
v4z	8.7060	av4z	8.7220	0.0160	0.1834
av4z	8.7220	saugav4z	8.7229	0.0009	0.0104
Correlation - active space (base setting: X2C CCSD(T))					
v2z (-50 to 50 a.u.)	8.5224	ae2z (-3300 to 3300 a.u.)	8.5212	-0.0011	-0.0131
Correlation - excitations (base setting: X2C saugav4z)					
CCSD	8.7355	CCSD(T)	8.7706	0.0351	0.4019
Level of relativity (base setting: v3z)					
X2C CCSD(T)	8.6343	4C DC CCSD(T)	8.6401	0.0058	0.0670
4C DC DHF	7.9834	4C DC+Gaunt DHF	7.9757	-0.0076	-0.0958
X2C DHF	7.9797	4C DC+Gaunt DHF	7.9757	-0.0040	-0.0497

Table C.7: Differences in ionisation potentials [eV] of HgI for different settings, at bond lengths $R_e^n = 2.71204$ Å and $R_e^+ = 2.55223$ Å

setting 1	IP [eV]	setting 2	IP [eV]	difference [eV]	difference [%]
Basis set quality (base setting: X2C CCSD(T))					
av4z	8.4603	CBS	8.4605	0.0002	0.0022
v4z	8.4565	av4z	8.4603	0.0039	0.0460
Correlation - active space (base setting: X2C CCSD(T))					
v3z (-30 to 30 a.u.)	8.4512	ae3z (-3300 to 3300 a.u.)	8.4770	0.0258	0.3053
Correlation - excitations (base setting: X2C av4z)					
CCSD	8.2589	CCSD(T)	8.4605	0.2016	2.4413
Level of relativity (base setting: v3z)					
X2C CCSD(T)	8.4512	4C DC CCSD(T)	8.4584	0.0072	0.0853
4C DC DHF	7.1916	4C DC+Gaunt DHF	7.1829	-0.0087	-0.1213
X2C DC DHF	7.1842	4C DC+Gaunt DHF	7.1829	-0.0013	-0.0177

Table C.8: Differences in ionisation potentials [eV] of TaN for different settings, at bond lengths $R_e^n = 1.67062$ Å and $R_e^+ = 1.64918$ Å