Accurate Prediction of Core Properties for Chiral Molecules using Pseudo Potentials

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von

Sebastian Marquardt, M. Sc.

aus
Darmstadt, Hessen

dem Fachbereich Chemie
der Philipps-Universität Marburg

Marburg an der Lahn
im Jahr 2018

Erstgutachter und Betreuer der Arbeit	Prof. Dr. Robert Berger
Zweitgutachter	PD. Dr. Ralf Tonner
weitere Mitglieder der Prüfungskommission	Prof. Dr. Rolf Schäfer
	Prof. Dr. Florian Kraus

Eingereicht am: 28. 05. 2018
Datum der mündlichen Prüfung: 17. 10. 2018

Als Dissertation vom Fachbereich Chemie der
Philipps-Universität Marburg (Hochschulkennziffer: 1180)
Deutscher Titel: Genaue Vorhersage der Kerneigenschaften chiraler Moleküle unter Verwendung von Pseudopotentialen

Contents

List of Tables

List of Figures

Curriculum Vitae

Abschlussarbeiten

B.Sc., *Towards an accurate description of aurophilicity within General Gradient Approximation based Density Functional Theory* bei Prof. Dr. Ursula Röthlisberger und Prof. Dr. Robert Berger, École Polytechnique Fédérale de Lausanne und Technische Universität Darmstadt, **2010**

M.Sc., *Theoretical Description and Computation of Meitner–Auger Transitions in Atomic Systems* bei Prof. Dr. Robert Berger, Technische Universität Darmstadt, **2012**

Publikationen

M. Pitzer, M. Kunitski, A. S. Johnson, T. Jahnke, H. Sann, F. Sturm, L. P. H. Schmidt, H. Schmidt-Böcking, R. Dörner, J. Stohner, J. Kiedrowski, M. Reggelin, **S. Marquardt**, A. Schießer, R. Berger, M. S. Schöffler
Direct Determination of Absolute Molecular Stereochemistry in Gas Phase by Coulomb Explosion Imaging
Science **2013**, *341*(6150), 1096–1100
DOI:10.1126/science.1240362.

M. Pitzer, G. Kastirke, M. Kunitski, P. T. Jahnke, T. Bauer, C. Goihl, F. Trinter, C. Schober, K. Henrichs, J. Becht, S. Zeller, H. Gassert, M. Waitz, A. Kuhlins, H. Sann, F. Sturm, F. Wiegandt, R. Wallauer, L. P. H. Schmidt, A. S. Johnson, M. Mazenauer, B. Spenger, S. Marquardt, **S. Marquardt**, H. Schmidt-Böcking, J. Stohner, R. Dörner, M. Schöffler, R. Berger
Absolute Configuration from Different Multifragmentation Pathways in Light-Induced Coulomb Explosion Imaging
ChemPhysChem **2016**, *17*, 2465–2472
DOI:10.1002/cphc.201501118.

S. Marquardt, S. Marquardt, M. Pitzer, K. Fehre, H. Schmidt-Böcking, R. Dörner, M. S. Schöffler, R. Berger
Theoretical study of gas phase fragmentation of multiply charged bromochlorofluoromethane (CHBrClF)
The manuscript is to be submitted for publication soon.

K. Gaul, **S. Marquardt**, T. Isaev, R. Berger
Systematic study of relativistic and chemical enhancements of $\mathcal{P}$, $\mathcal{T}$-odd effects in polar diatomic radicals
arXiv:1805.05494

S. Marquardt, R. Berger
Calculation of Core-like Properties with Pseudo Potentials
The draft is in preparation.

Abstract

Pseudo potentials (PPs) constitute perhaps the most common way to treat relativity, often in a formally non-relativistic framework, and reduce the electronic structure to the chemically relevant part. The drawback is that orbitals obtained in this picture (called pseudo orbitals (POs)) show a reduced nodal structure and altered amplitude in the vicinity of the nucleus, when compared to the corresponding molecular orbitals (MOs). Thus expectation values of operators localized in the spatial core region that are calculated with POs, deviate significantly from the same expectation values calculated with all-electron (AE) MOs. This study describes the reconstruction of AE MOs from POs, with a focus on POs generated by energy consistent pseudo Hamiltonians. The method reintroduces the nodal structure into the POs, thus providing an inexpensive and easily implementable method that allows to use nonrelativistic, efficiently calculated POs for good estimates of expectation values of core-like properties.

The discussion of the method proceeds in two parts: Firstly, the reconstruction scheme is developed for atomic cases. Secondly, the scheme is discussed in the context of MO reconstruction and successfully applied to numerous numerical examples.

Starting from the equations of the state-averaged multi-configuration self-consistent field method, used for the generation of energy consistent pseudo potentials, the electronic spectrum of the many-electron Hamiltonian is linked to the spectrum of the effective one-electron Fock operator by means of various models systems. This relation and the Topp–Hopfield–Kramers theorem, are used to show the shape-consistency of energy-consistent POs for atomic systems. *Shape-consistency* describes POs that follow distinct AOs exactly outside a core-radius r_{core}. In the cases presented here, shape-consistency holds to a high degree and it follows that in atomic systems every PO has one distinct partner in the set of AOs. The overlap integral between these two orbitals is close to one, as it is determined mainly by the spatial orbital parts outside r_{core}. Expanding, e.g., a 5s PO in occupied AOs, the 5s AOs will have the highest contribution. The POs itself contains contributions from high-energy unoccupied AOs as well (e.g. 15s), which damp the nodal structure of the POs near the nucleus. Consequently, neglecting contributions from unoccupied orbitals in a projection of the POs

reintroduces the nodal structure.

This approach is not directly suitable for the reconstruction of MOs, as they often need to be expanded in a full set of AOs at each atomic center, including all unoccupied orbitals, to properly account for the electron density distribution in the molecule. However, it is shown that the occupied MOs are well described by occupied and low-energy unoccupied AOs only and a mapping of the POs onto a basis containing only these orbitals reconstructs the nodal structure of the MO. The approach uses only standard integrals available in most quantum chemistry programs. The computational cost of these integrals scales with N^2, where N is the number of basis functions. The most time consuming step is a Gram-Schmidt orthogonalization, which scales in this implementation with MN^2, M being the number of reconstructed orbitals.

The reconstruction method is subsequently tested: Valence orbitals of atomic, closed-shell systems were reconstructed numerically exactly. The influence of numerical parameters is investigated using the molecule $BaF^{\bullet}$. It is shown that the method is basis set dependent: One has to ensure that the PO basis can be expanded exactly in the basis of AOs. Violating this rule of thumb may degrade the quality of reconstructed orbitals. Additionally, the representation of MOs by a linear combination of occupied and unoccupied AOs is investigated. For the exemplary systems, the shells included in the fitting procedure of the PP were sufficient.

Reconstruction of the alkaline earth monofluorides showed that periodic trends can be reconstructed as well. Scaling of hyperfine structure parameters with increasing atomic number is discussed. For hydrogenic atoms, the scaling should be linear, whereas small deviations from the linear behavior were observed for molecules. The scaling laws computed from reconstructed and reference orbitals were almost identical. In this context, the failure of commonly used relativistic enhancement factors beyond atomic number 100 is discussed. Applicability of the method is also tested on parity violating properties for which the main contribution is generated by the valence orbitals near the nucleus. Symmetry-independence of the method is shown by successful reconstruction of orbitals of the tetrahedral $PbCl_4$ and chiral NWHClF. The reliable reconstruction of chemical trends is shown with the help of the NWHClF derivatives NWHBrF and NWHFI.

The study of chiral compounds as, e.g., NWHClF and its group 17 derivatives, which have been proposed as paradigm for the detection of parity-violation in chiral molecules[1], remains of great importance. Especially the direct determi-

[1] D. Figgen, T. Saue, P. Schwerdtfeger, *J. Chem. Phys.* **2010**, *132*(23), 234310,

nation of absolute configuration of chiral centers is still non-trivial. The author contributed to this field with a self-written molecular dynamics (MD) program to simulate Coulomb explosions and thus to provide an insight especially into the early explosion stages directly after an instantaneous multi-ionization of the molecule CHBrClF, comparable to experiments using the Cold Target Recoil Ion Momentum Spectroscopy (COLTRIMS) technique. An algorithm for the determination of the investigated molecule's absolute configuration from time-of-flight data and detection locations of molecular fragments is included in the program. The program was used to generate experiment-equivalent data which allowed for the first time the investigation of non-racemic mixtures by the analysis routines of the experiment. The MD program includes harmonic and anharmonic bond potentials. A charge-exchange model can model partial charges in early phases of the Coulomb explosion.

Furthermore, Born–Oppenheimer MD simulations and statistical models are used to explain the relative abundance of products belonging to competing reaction channels, as obtained by photoion coincidence measurements. Additionally, qualitative statements about reaction branching ratios are made by comparing the partition functions of involved degrees of freedom. Analytic equations for partition functions of simple models are used to provide a simple formula allowing fast estimates of reaction branching ratios.

DOI:10.1063/1.3439692.

1. Introduction

1.1. Short History of Relativistic Quantum Chemistry Methods

Shortly after Heisenberg [2], Born, Heisenberg and Jordan [3] and Schrödinger [4, 5] had developed the mathematical grounds for quantum mechanics, Dirac [6] extended the wave equation to be Lorentz invariant and to incorporate the electron spin from first principles. Despite the generality of the approaches of the authors and their tremendous value for the development of (relativistic) quantum mechanics, the applicability of the equations was limited to the simplest systems. Only with the developments of Hartree [7–9] and Fock [10] and their extensions by Slater [11], the calculation of electronic structures of chemically meaningful systems became possible. However, computational limitations at the time called for simplifications, which were provided by Hellmann and Gombás in the year 1935. [12–15] They introduced pseudo potentials (PPs) to reduce the number of electrons treated explicitly in calculations of electronic structures of diatomic alkaline molecules. These PPs replaced all but one electron of an atom and were parametrized to reproduce the electronic excitation spectra of the atoms. The electronic systems of the molecules were thus reduced to effective one-electron-per-atom systems, which could be solved with the available methods. In the year 1959, Phillips and Kleinman [16] found an approach for PPs based on the atomic core orbitals (ACOs), which was developed further by Weeks and Rice [17] and has been a large success ever since. Especially in the fields of non-relativistic (NR) and scalar-relativistic (SR) quantum chemistry PPs of generalized Phillips–Kleinman (GPK) type are frequently used. However, Cohen and Heine showed in Ref. [18] that PPs of Phillips–Kleinman (PK) type are not unique, as an admixture of core orbitals to valence orbitals leaves the orbital energy unchanged. They circumvented this indeterminacy by simultaneously minimizing orbital expectation values for certain operators, e.g., for the kinetic energy. Relativistic effects were first included in effective core potential (ECP) treatments by Lee, Ermler and Pitzer [19] who fitted the PP Hamiltonian to a Dirac–Hartree–Fock (DHF) Hamiltonian. Shortly after, Kahn, Hay and

Cowan [20] used a Cowan–Griffin–Hartree–Fock (HF) Hamiltonian for the same purpose. Modern energy-consistent (EC) PPs, e.g., of Stuttgart type, are adapted to reproduce the electronic excitation spectrum of all-electron (AE) many-state Dirac–Coulomb–Breit (DCB) calculations. [21, 22]

All PP approaches have in common that the electrons in an atom are partitioned into core electrons and valence electrons. The effect of the core electrons is described by the PP and these electrons do not have to be considered in further calculations. The pseudo atomic orbitals (PAOs) have a simpler nodal structure but can be obtained with tremendously reduced effort from NR calculations. There exist many procedures to obtain PPs and their discussion as well as other alternatives have been covered in depth by various reviews. [23–26]

In contrast to NR quantum chemistry and the treatment of relativity through PPs, efficient solution methods for the many-particle Dirac–Hamiltonian were underrepresented until the late 1970s. [27] State of the art was the calculation of the four-component (4c) electronic wave function from the Dirac–Coulomb (DC) or DCB Hamiltonian. From a theoretical point of view this represents the most rigorous approach but comes at the cost of high demands for computational resources. [28, 29] In order to reduce the effort involved in the calculations of 4c wave functions, many effective two-component (2c) theories have been derived that treat the small component of the 4c wave function only implicitly. Notable are here the exact two component (X2C) solution [30–32] and the infinite order two component (IOTC) solution [33–35]. Those methods give practically identical results compared to 4c calculations, but can be computationally more efficient. On the other hand, Liu and Peng have pointed out that 4c approaches can be formulated such that they are as efficient as quasi relativistic approaches. [36] Another approach, termed direct perturbation theory (DPT), was developed by Rutkowski [37–40] and Kutzelnigg [41, 42]. They expanded a transformed Dirac equation in powers of $\frac{1}{c^2}$, with c being the speed of light. Stopkowicz and Gauß [43] obtained good agreement with DHF when applying DPT of fourth order to the lighter hydrogen halides. Schwalbach, Stopkowicz and Gauß [44] showed the significance of sixth and higher order terms of the expansion in molecules containing heavy elements, whereas third and fourth order correction terms are sufficient or molecules with lighter atoms.

The high order corrections, which are necessary to describe relativistic effects in molecules with super heavy elements, lead to the conclusion that perturbation theory with the NR Hamiltonian as limiting case might not be the best way to go. Chang, Pélissier and Durand (CPD) tried to construct a Hamiltonian which contains relativistic corrections already in first order. [45] Van Lenthe, Baerends and Snijders showed that this CPD or zeroth order regular

approximation (ZORA) Hamiltonian agrees very well with the single electron solution of the Dirac equation. [46] Later, the same authors improved their theory such that total energies and orbital energies could be described very well within the $X\alpha$ approximation. [47] Thus, chemical accuracy ($5 \times 10^{-4}\ E_\mathrm{h}$, according to the authors) came into reach for the calculation of many particle wave functions. Similar results can be obtained already with ZORA, if the electron density is partitioned into a core density and a valence density. Only the valence orbitals are computed explicitly, whereas the core density stems from another, more precise, calculation. [46] This partitioning approach is evidently central to the earlier mentioned PP approaches, from Hellmann and Gombás to the most recent formulation of EC PPs of the Stuttgart–Köln group.

With this work I aim to link the AE ansatz with the approach for widely used EC PPs and thereby show that almost 90 years of advances of the most general equations in quantum mechanics can be cut short by PPs and using simple transformations.

1.2. Core-like Molecular Properties in the PP approximation

The computational effort of many quantum chemical methods grows polynomially as N^a, where $a \geq 1$ and N is the number of electrons in the system under investigation. Very often, $a \geq 5$ and N is large even for moderately sized molecules. One of the big successes of PPs is the enormous decrease in N, accompanied with corresponding savings in computation time. The incorporation of scalar relativistic effects into the PP is independent of the system size and establishes the calculation of relativistic effects at the cost of NR calculations. Furthermore, the same NR methodology can be applied, rendering relativistic treatments commonly unnecessary and thus allowing, e.g., structure optimizations for large systems including relativistic effects. The wrong nodal structure and amplitude of the POs[1], however, hamper the calculation of some molecular properties. Schwerdtfeger notes in Ref. [24] that pseudo Hamiltonians describe only so-called *valence properties*[2] properly, defined as properties of the type

$$r^{-n}\frac{\mathrm{d}^m}{\mathrm{d}r^m}, \qquad m+n \leq 0. \tag{1.1}$$

[1] POs are the orbitals that form an eigen basis for a pseudo Hamiltonian. See Ch. 2.1 further discussion of the nomenclature.

[2] The properties are also called *valence-like* to emphasize that their expectation values depend also to a minor degree on the wave function in the core region.

Expectation values of those properties depend mainly on the electronic wave function at larger distances from the nuclei. Expectation values of *core properties* (or *core-like* properties) are primarily determined by the valence orbitals of the system near the PP. There, the POs show the wrong nodal structure and amplitude compared to AE orbitals. A direct calculation of core-like expectation values with a pseudo wave function is therefore usually not meaningful. Fig. 1.1 shows the integrands of the expectation values

$$\langle \hat{r} \rangle = \int \psi^*(r) r \psi(r) r^2 \, \mathrm{d}r \tag{1.2}$$

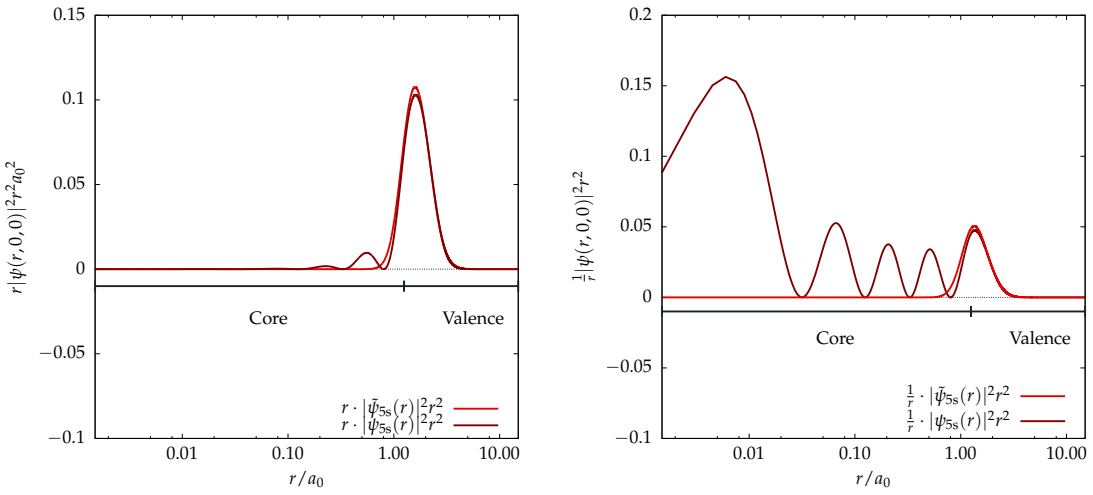
and

$$\left\langle \frac{1}{\hat{r}} \right\rangle = \int \psi^*(r) \frac{1}{r} \psi(r) r^2 \, \mathrm{d}r. \tag{1.3}$$

for PAOs and AE atomic orbitals (AOs). The expectation value of the operator $\hat{r}$ calculated with the PAO is a good approximation to the AE orbital expectation value, whereas the PAO expectation value of the $\frac{1}{\hat{r}}$ operator misses huge contributions from the core region (here, the core radius is arbitrarily marked at $1.25\, a_0$).

If core properties are to be calculated from POs then it is evidently necessary to post-process the orbitals, e.g., with an AE single-point calculation. This constitutes a computational overhead, especially as the valence orbitals of the pseudo system cannot be used as starting orbitals for further orbital variation. Thus in general, additional steps are needed to transform the POs into a form that can be used in further AE variational schemes. Overcoming this problem would render AE calculations only necessary in cases where a multicomponent calculation is explicitly required. A noteworthy attempt has been made by Daasch, McMurchie and Davidson who reconstructed the necessary ingredients in order to estimate core-like molecular properties in the configuration interaction (CI) approximation from AE atomic calculations and molecular PP CI calculations. [48] Cioslowski and Psikorz on the other hand augmented valence electron densities by atomic core electron densities in order to estimate several properties. [49] Hinds and Sandars [50] as well as Conveney and Sandars [51] constructed relativistic orbitals via phase and amplitude matching of NR orbitals at long distances.

A different approach was taken by Titov, Kozlov *et al.*. They introduced the nonvariational one-center restoration procedure (NOCR), where a shape-consistent (SC) PP is constructed together with a corresponding AE basis set. After the PP calculation the valence orbitals of the pseudo systems are replaced by their AE counterpart and the property is evaluated for the valence molec-



(a) **(b)**

Figure 1.1.: Integrand form of $\hat{r}$ and $\frac{1}{\hat{r}}$ operator integrals for the Ba 5s orbital. The function $r|\psi(r)|^2r^2$ is the integrand of the valence-like operator $\hat{r}$, whereas $\frac{1}{r}|\psi(r)|^2r^2$ represents is the integrand of the integral form of the core-like operator $\frac{1}{\hat{r}}$. The AE orbital ψ_{5s} and PP orbital $\widetilde{\psi}_{5s}$ valence-like integrands are very similar in the valence region whereas the core-like integrands have not much in common: In the core region, the PO integrand decays whereas its AE pendant has significant contributions.

ular orbitals (MOs). [52, 53] Thus a number of *parity-violating* properties was calculated with good accuracy. [28, 54–56]

If a molecular property violates parity or not, is defined by the commutator of the property's Hamiltonian and the parity operator. The parity operator inverts the signs of all spatial coordinates of a function $f(\vec{r})$. [57, p. 76]

$$\hat{\mathcal{P}}f(\vec{r}) = f(-\vec{r}) \tag{1.4}$$

If $f(\vec{r})$ is an eigenfunction of the parity operator, one has

$$\hat{\mathcal{P}}f(\vec{r}) = af(\vec{r}) \tag{1.5}$$

The most notable values of a are $+1$ and -1 and classify $f(\vec{r})$ *to be of even* $(+1)$ *or odd* (-1) *parity*. If a Hamiltonian commutes with the parity operator, i.e.,

$$\left[\hat{H}, \hat{\mathcal{P}}\right] = 0, \tag{1.6}$$

then the system *conserves* parity. Otherwise, parity is *violated* and the strength of the violation is determined by the operators involved.

I present herein a scheme for the reconstruction of SR AE MOs from PP MOs and SR atomic AE reference orbitals. The scheme is derived for EC PPs. Its applicability is shown for EC as well as SC PPs. The reconstructed MOs are used for the calculation of core-like molecular properties as, e.g., expectation values of the core-electron Coulomb operator, the (one-electron) mass-velocity Darwin (MVD) term, the hyperfine coupling constants or the electronic structure parameter of the parity-violating coupling.

1.3. Parity Violation and Chirality

Parity violation causes the total energy of enantiomers to differ. [58] However, modern experiments have not yet revealed this tiny energy difference: Avalos *et al.* [59] write that *homochirality*, i.e. the predominance of one enantiomer or one stereoisomer in nature on earth or in the universe, is thought to be the macroscopic outcome of the energy difference between enantiomers due to parity violation. The expected order of magnitude for the energy difference between enantiomers is 10^{-20}–$10^{-17} E_\mathrm{h}$. [59] However, this small energy difference can account only for a small imbalance between the enantiomeric amounts and thus would have needed a strong amplification process to result in today's homochiral biochemistry. Other hypotheses, as the *Vester–Ulbricht–hypothesis*, discuss the chirality of physical processes like the β-decay as the reason for homochirality in nature. [60] Rikken and Raupach show in Ref. [61] that photochemistry with linearly polarized light in chiral magnetic fields can also lead to energetic distinction of the enantiomers and they discuss this phenomenon as a possible cause of homochirality in earth's biochemistry.

The study of homochirality necessarily involves the study of chirality of single molecules and the molecular structure dependence of their properties. Although chirality was already observed by Pasteur in the year 1848, Ref. [62], and attributed to the arrangements of atoms in molecules by van't Hoff [63] and Le Bel [64], its direct determination remains cumbersome up to today. Bijvoet introduced in the year 1951 X-ray crystallography for the determination of absolute configuration of single crystals. [65] Inokuma *et al.* used porous complexes to absorb noncrystalline molecules and thus lifted the requirement of crystallinity. [66] However, not all stereocenters of the investigated molecule were identified unambiguously. [67] Other attempts predict NMR spectroscopy as a possible probe for chirality by coupling the induced electric dipole moment of the molecule to the magnetic moment of one nucleus and thus creating a chiral probe. [68] However, they have not been tested yet. The use of residual dipolar

couplings in NMR for enantiomer assignment still requires additional input from first principle calculations. [69] King *et al.* propose an experiment to determine chirality from the permanent antisymmetric J-coupling tensor and thus hope to provide the means to unambiguously discriminate between enantiomers. [70]

Herein, I demonstrate a method based on COLd Target Recoil Ion Momentum Spectroscopy (COLTRIMS), where high energy photons (to form core holes) or high intensity Laser beams are used to multiply ionize chiral molecules in the gas phase. Subsequently they undergo Coulomb explosion and ionic fragments are recorded by time and position sensitive detectors. These data are used to reconstruct the particles' linear momenta and thus directly determine the absolute configuration of the exploded molecule.

In Ref. [71] we presented the direct determination after Laser ionization using five-particle-coincidence events. Ionization with synchrotron radiation was used in Ref. [72], where we also used three- and four-particle coincidence events for the enantiomer discrimination. The complexity of the gas phase structures of highly-charged CHBrClF and various fragmentation channels are discussed in Ref. [73]. These works constitute the second part of this thesis.

1.4. Structure of the Thesis

Having introduced the broad topics in Ch. 1, I will briefly discuss the formalisms necessary for a complete comprehension of this work in Ch. 2. I start with the nomenclature used throughout this work in Sec. 2.1. The PP approximation is introduced in Sec. 2.2, followed by an introduction to multi-configuration self-consistent field (MCSCF) theory in Sec. 2.3. The results of the PhD work are discussed in Ch. 3, commencing with the requirements I place on a useful orbital reconstruction method in Sec. 3.1. In Sec. 3.2 I show the relation of EC POs to their AE counterparts and develop a reconstruction scheme. An elaborate protocol for its application is provided in Sec. 3.3 and complemented by a numerical example in Sec. 3.4. Sec. 3.5 summarizes all the findings of the previous sections. In Sec. 3.6 I test the method on various atomic and molecular systems. Sec. 3.7 concludes the chapter. Ch. 4 on the direct determination of absolute configuration of gas phase molecules begins with a concise introduction to the topic in Sec. 4.1. In Secs. 4.2–4.4 three publications, containing contributions from me, are presented. Ch. 5 summarizes the achievements of this thesis and provides an outlook to potential future research directions. The thesis is paraphrased in German language in Ch. 6. In the Appendix, additional data are provided that maybe helpful for a deeper understanding and reproduction of this work.

2. Theory

This chapter starts with an explanation of the nomenclature used in this work in Sec. 2.1. As the overall goal of the thesis is a reconstruction method usable for EC PP of the Stuttgart group, I will give a short introduction to PPs in general and the Stuttgart-PPs in particular in Sec. 2.2. These PPs are usually generated using state-averaged (SA) MCSCF methods, which are introduced in Sec. 2.3. These equations will be used throughout Ch. 3 to show the connection between AE orbitals and EC POs. However, a reader only interested in the outcome of this work should head directly to Ch. 3.

2.1. Nomenclature

Before turning the attention to the theoretical foundation of this work, I will discuss in short the nomenclature used in this work: A general electronic wave function of an N-electron system will be denoted by Ψ. If multiple states are discussed at the same time, they will usually be indexed by a letter in calligraphic style, e.g., $\mathcal{A}$. Slater determinants (SDs) or configuration state functions (CSFs), which may span a basis for the states, are denoted by the letter Φ. One-electron MOs are identified by ψ, whereas ϕ stands for AOs. This distinction is useful, as I will discuss PPs with respect to (w.r.t.) the electronic situation in atoms *and* molecules. For their coordinates I use the usual notation, i.e., $\vec{r}_i$ is the space coordinate of particle i and $\vec{x}_i$ the spin and space coordinate of particle i. A (more often than not Gaussian) basis for the spatial part of the AOs is marked by χ. The Gaussian basis functions (BFs) may occur contracted, denoting the primitive BFs by $\bar{\chi}$. Operators, like the Hamiltonian $\hat{H}$, are indicated by a caret "^" above the letter. Matrices are given as $\boldsymbol{H}$, whereas vectors are printed $\vec{c}$. Their elements are identified by two or one subscripts H_{ij}, c_i. The inverse of some matrix is defined by $\boldsymbol{A}^{-1}$, the transpose by $\boldsymbol{A}^T$, the complex conjugate by $\boldsymbol{A}^*$ and the adjoint by $\boldsymbol{A}^\dagger$. Throughout this work, I employ Dirac's bra-ket notation.

Concerning the discussion of PPs and related quantities, I will also use some further identifiers: A MO that was computed for a system containing a PP or,

more general, ECP Hamiltonian (see the next section for the distinction), will be denoted by $\widetilde{\psi}$, i.e., a tilde "~" is printed above the discussed quantity. The notation will be used for functions as well as for operators and other quantities. In text, the orbital $\widetilde{\psi}$ is named pseudo molecular orbital (PMO), while $\widetilde{\phi}$ is named PAO. This naming was proposed by Dolg and Cao [26] to clarify that the term *pseudo* describes the altered nature of the system's Hamiltonian. However, when the discussion does not call for a distinction between PMO and PAO, I still use the ambiguous term PO and imply the definition given by Dolg and Cao. Similarly, I discriminate between, core, valence and unoccupied atomic orbitals (ACO, AVO, AUO). A complete list of all abbreviations and symbols used in this work is placed after the appendix on pages A-37ff. and A-41ff.

2.2. Pseudo Potentials

An extensive overview over the field of PPs is found in Ch. 5 of the review of Dolg and Cao (Ref. [26]) and the reviews of Cao and Dolg [25] and Schwerdtfeger [24]. Pyykkö and Stoll discussed in an earlier review the developments of relativistic PPs in the 1990s. [23] Sec. 2.2.1 introduces the frozen core ansatz before I turn to a brief history of PPs in Sec. 2.2.2. Sec. 2.2.3 paves the way towards EC PPs and explains their parametrization.

2.2.1. Frozen Core Approximation

The frozen core (FC) approximation underlies all quantum chemical methods that are used to determine a subset of the full electronic structure of molecular systems, while keeping the remaining subset constant. Although initially used for transferability of atomic cores to molecular calculations, McWeeny's development of the *group function* approach [74] is extensible to any other kind of frozen system. The theory of group functions starts with a definition of the wave function Ψ as an antisymmetrized product of antisymmetrized subsystem wave functions Ψ_i.

$$\begin{aligned}\Psi = M\hat{A}\,\big(&\Psi_a(\vec{x}_1,\ldots,\vec{x}_{N_a})\Psi_b(\vec{x}_{N_a+1},\ldots,\vec{x}_{N_a+N_b})\ldots\\ &\Psi_k(\vec{x}_{N_a+N_b+\ldots+1},\ldots,\vec{x}_{Na+N_b+\ldots+N_k})\big)\end{aligned} \tag{2.1}$$

$\vec{x}_i$ represents the coordinate of the ith fermion. The antisymmetrization operator $\hat{A}$ runs over all interchanges of particles between the various groups, while M ensures that the wave function remains normalized. Usually, the wave function

Ψ is not exact, as all its components have been obtained individually. However, assuming *strong orthogonality* between the group functions, i.e.,

$$\int \Psi_i^*(\vec{x}_1, \vec{x}_2, \dots \vec{x}_N)\Psi_j(\vec{x}_1, \vec{x}'_2, \dots \vec{x}'_N)\, \mathrm{d}\vec{x}_1 = 0 \qquad , \tag{2.2}$$

McWeeny established in Ref. [74] a variational theory for group functions, where an electron group function is varied in the field of all other electron groups in order to minimize its effective energy. Subsequently, each electron group function is varied in turn and the procedure is repeated until convergence. McWeeny's approach lends itself naturally to separate chemically distinct groups of electrons into, e.g., core and valence electrons. [75] A precalculated determinant that describes the atomic core electrons can then be used in calculations for the molecule and by excluding the atomic core from the variational procedure and reduces the computation time. Dyall showed in Ref. [76], how the group function ansatz for a core and a valence determinant can be used to define a valence Hamiltonian, if the core determinant is fixed.[1] The valence Hamiltonian is then (in atomic units)

$$\hat{H}_v = \sum_{i=1}^{N_v} \hat{h}_i + \sum_{\substack{i=1 \\ j=i+1}}^{N_v} \frac{1}{r_{ij}} + \hat{v}_{\mathrm{FC}} \tag{2.3}$$

$\hat{h}_i$ is the one-electron operator for particle i and $\frac{1}{r_{ij}}$ the usual non-relativistic two-electron Coulomb interaction potential. The FC potential v_{FC} contains all interactions of the core electrons with valence electrons and, in principle, the interaction of core electrons among themselves. For a closed-shell core configuration $\hat{v}_{\mathrm{FC}}$ reads as

$$\hat{v}_{\mathrm{FC}} = \sum_{i=1}^{N_v} (2J_i^c - K_i^c) - E_{\mathrm{core}} \tag{2.4}$$

with

$$J_i^c = \sum_c \int \frac{\phi_c^*(\vec{r}_j)\phi_c(\vec{r}_j)}{r_{ij}}\, \mathrm{d}\vec{r}_j \tag{2.5}$$

and

$$K_i^c\phi_k(\vec{r}_i) = \sum_c \int \frac{\phi_c^*(\vec{r}_j)\phi_k(\vec{r}_j)}{r_{ij}}\, \mathrm{d}\vec{r}_j\phi_c(\vec{r}_i). \tag{2.6}$$

[1] The core determinant is spanned by the core orbitals only, whose eigenvalues contribute significantly to the total energy. Contrariwise, valence orbitals contribute less to the total energy and are much more diffuse. Additionally, they usually form the basis for chemical bond descriptions in linear combination of atomic orbitals to molecular orbitals (LCAO-MO) calculations.

Note that the index c runs over all core electrons. E_{core} is the total energy of the determinant spanned by all core orbitals.

Although precalculated ACOs can decrease the calculation time needed to perform the iterations of the self-consistent field (SCF) algorithm, it is often much more beneficial to avoid core orbitals altogether. Then, the valence orbitals do not retain their undulations close to the nucleus due to orthogonality to the core orbitals. Many basis functions with high exponents are needed in LCAO-MO expansions to properly describe the spatial core region of the MOs. As the core region is rather unimportant for the chemical bond formation, these localized basis functions introduce a computational overhead unnecessary for the calculation of molecular structures and chemical bonds.

Computational savings are then achieved by approximating the potential v_{FC} by some other potential. Dolg and Cao [26] distinguish between ECPs, model potentials (MPs) and PPs. ECPs are any potentials v_{ECP} that are used in (2.3) as a replacement for v_{FC}. MPs on the other hand model the effect of the non-local potential v_{FC} to obtain the correct nodal structure in the valence orbitals. The PP approach uses a parametrized analytical potential to mimic the effect of v_{FC}. Usually, the parameters are determined by matching some property of the orbitals or many-electron wave function to the AE analogue. Despite all the different methods used to obtain a PP, the various PP variants come in two basic flavors: Shape-consistent PP generate a Hamiltonian, whose eigenfunctions follow the eigenfunctions of the AE Hamiltonian exactly outside a core radius r_c.

$$\widetilde{\phi}_i(r) = \begin{cases} \phi_i(r), & r \geq r_c \\ f_i(r), & r < r_c \end{cases} \tag{2.7}$$

The function $f_i(r)$ is often chosen to smooth the orbitals close to the nucleus. The other flavor is energy consistency. EC PPs are parametrized to reproduce quantum mechanical observables like the electron excitation spectrum or ionization energies of the atom. A famous example are the Stuttgart PPs developed by Dolg *et al.* in the year 1987. [77] In the following section, I will discuss some of the history of modern PPs with a special focus on EC PP.

2.2.2. Early Development of PPs

In the mid 1930s Hellmann [12, 13] derived a semi-empirical method in order to estimate molecular spectra of diatomic alkaline metals. He replaced the atomic

core by including the electrons in a local potential $V(r)$

$$V(r) = \frac{1}{r}(A\mathrm{e}^{-\kappa r} - 1). \tag{2.8}$$

The potential was derived with the Thomas–Fermi model and the parameters A and κ were chosen from experimental electronic excitation spectra. Molecular problems were then tackled by explicitly treating only one remaining valence electron per atom, thus reducing the computational work tremendously. Hellmann and Kassatotschkin [78, 79] applied the approach to the calculation of metallic bond formation and properties of alkaline and alkaline earth metals. At the same time, a very similar approach was taken by Gombás [14, 15] to predict alkaline metal bonds and their properties.

The reader may note the similarity of the potential form (2.8) to the strong nucleon interaction potential proposed by Yukawa in the year 1935. [80]

$$V(r) = -\frac{1}{r}g\mathrm{e}^{-\frac{1}{\lambda_\mathrm{C}}r} \tag{2.9}$$

g is here a coupling constant that determines the interaction strength and λ_C is the Compton wave length of the gauge boson carrying the interaction. For $\lambda_\mathrm{C} \to 0$, i.e., when the gauge boson is a Photon, and $g \to \mathrm{e}$, where e is the elementary charge, one obtains the Coulomb potential of a point charge.

Combining the strong interaction with the electromagnetic interaction, as, e.g., in the interaction of two protons, one obtains exactly Hellmann's potential. In this respect it is possible to interpret the term $\frac{A}{r}\mathrm{e}^{-\kappa r}$ in (2.8) as the Yukawa potential generated by a pseudo-lepton, to which the valence electron can couple through a pseudo boson exchange. For the PPs of Hellmann, g is usually very small (but positive) and λ_C rather large (the mass of the gauge boson is between the electron and proton mass). The potential is thus very short ranged.[2]

In the year 1959, Phillips and Kleinman [16] developed the modern theory of PPs, known today as PK PPs. By orthogonalizing a valence function ϕ_v w.r.t. a set of core orbitals ϕ_c they obtained the PO $\widetilde{\phi}$.

$$\phi_v = \widetilde{\phi} - \sum_c a_c \phi_c \tag{2.10}$$

The coefficients a_c are determined by the fact that the set of core orbitals is

[2] Hellmann gives in Ref. [12, 78] values for $\kappa = 0.6 - 1.7$ for alkaline and earth alkaline atoms in atomic units. That corresponds to a mass of $\alpha\kappa \approx 82 - 233$ times the electron mass (α being the fine structure constant).

orthonormal. Additionally, the set is orthogonal to the set of valence orbitals.

$$\langle \phi_c | \phi_v \rangle = 0 = \langle \phi_c | \widetilde{\phi} \rangle - \sum_{c'} \langle \phi_c | \phi_{c'} \rangle \, a_{c'} \tag{2.11}$$

$$\langle \phi_c | \widetilde{\phi} \rangle = \sum_{c'} \langle \phi_c | \phi_{c'} \rangle \, a_{c'} \tag{2.12}$$

$$\langle \phi_c | \widetilde{\phi} \rangle = \sum_{c'} \delta_{cc'} \tag{2.13}$$

$$a_c = \langle \phi_c | \widetilde{\phi} \rangle \tag{2.14}$$

Phillips and Kleinman started from the Hartree–Fock picture, where the Schrödinger equation reduces to the effective one-electron HF equation and formulated it for the valence electron orbital.

$$\hat{h}_{\mathrm{eff},i} \phi_v(\vec{x}_i) = \left(\hat{h}_i + \hat{v}_{\mathrm{FC},i} \right) \phi_v(\vec{x}_i) = \varepsilon_v \phi_v(\vec{x}_i) \tag{2.15}$$

Inserting (2.10) into (2.15), a pseudo-eigenvalue equation for the PO is obtained.

$$\hat{h}_{\mathrm{eff}} \left(\widetilde{\phi} - \sum_c \phi_c a_c \right) = \varepsilon_v \left(\widetilde{\phi} - \sum_c \phi_c a_c \right) \tag{2.16}$$

$$\hat{h}_{\mathrm{eff}} \widetilde{\phi} - \sum_c \hat{h}_{\mathrm{eff}} \phi_c a_c = \varepsilon_v \widetilde{\phi} - \varepsilon_v \sum_c \phi_c a_c \tag{2.17}$$

$$\hat{h}_{\mathrm{eff}} \widetilde{\phi} - \sum_c \left(\varepsilon_c - \varepsilon_v \right) \phi_c a_c = \varepsilon_v \widetilde{\phi} \tag{2.18}$$

$$\hat{h}_{\mathrm{eff}} \widetilde{\phi} + \left(\sum_c \left(\varepsilon_v - \varepsilon_c \right) |\phi_c\rangle\langle\phi_c| \right) \widetilde{\phi} = \varepsilon_v \widetilde{\phi} \tag{2.19}$$

$$\left(\hat{h}_{\mathrm{eff}} + v_{\mathrm{PK}} \right) \widetilde{\phi} = \varepsilon_v \widetilde{\phi} \tag{2.20}$$

$$v_{\mathrm{PK}} = \sum_c \left(\varepsilon_v - \varepsilon_c \right) |\phi_c\rangle\langle\phi_c| \tag{2.21}$$

The index i was dropped for brevity.

Cohen and Heine noted in Ref. [18] that the choice of the PP by Phillips and Kleinman is not unique, as any addition of the core orbitals to the valence orbitals does not change the total energy but the form of the valence orbitals. However, they point out that this problem can be lifted by applying additional constraints. Their proposal is to expand ϕ' in a set of functions f_k. The expansion coefficients are chosen to minimize the energy or to smooth out the inner oscillations of $\widetilde{\phi}$ by minimizing the radial integral $\int |\nabla \phi'| \, \mathrm{d}r$.

Weeks and Rice [17] took up the work of Phillips and Kleinman and generalized the additional potential $V_{\mathrm{PK}}(r)$. They started with the definition of core

and valence projection operators for one electron i

$$\hat{P}_c(i) = \sum_c |\phi_c(i)\rangle\langle\phi_c(i)| \tag{2.22}$$

$$\hat{P}_v(i) = \hat{1} - \hat{P}_c(i). \tag{2.23}$$

The N-electron projection operators, are obtained as products over the one-electron projection operators.

$$\hat{P}_c = \prod_{j=1}^{N_c} \hat{1}(j) \prod_{i=N_c+1}^{N_v} \hat{P}_c(i) \tag{2.24}$$

$$\hat{P}_v = \prod_{j=1}^{N_c} \hat{1}(j) \prod_{i=N_c+1}^{N_v} \hat{P}_v(i) \tag{2.25}$$

$\hat{P}_c(i)$ removes from any function the components of ϕ_c. Let the many-electron functions Φ and $\widetilde{\Phi}$ be defined as

$$\begin{aligned} \Phi &= |\phi_1(\vec{x}_1)\dots\phi_{n_c}(\vec{x}_{n_c})\phi_{n_c+1}(\vec{x}_{n_c+1})\dots\phi_n(\vec{x}_n)| \\ \widetilde{\Phi} &= |\phi_1(\vec{x}_1)\dots\phi_{n_c}(\vec{x}_{n_c})\widetilde{\phi}_{n_c+1}(\vec{x}_{n_c+1})\dots\widetilde{\phi}_n(\vec{x}_n)| \end{aligned} \tag{2.26}$$

where the vertical bars denote a Slater determinant. If the valence projection operator acts on each one-electron basis function such that

$$\phi_i = \hat{P}_v\widetilde{\phi}_i \qquad , i \in \text{valence} \tag{2.27}$$

then one gets for Φ

$$\Phi = |\phi_1(\vec{x}_1)\dots\phi_{n_c}(\vec{x}_{n_c})\phi_{n_c+1}(\vec{x}_{n_c+1})\dots\phi_n(\vec{x}_n)| \tag{2.28}$$

$$\begin{aligned} &= |\phi_1(\vec{x}_1)\dots\phi_{n_c}(\vec{x}_{n_c})(\widetilde{\phi}_{n_c+1}(\vec{x}_{n_c+1}) - \sum_c^{n_c}\phi_c(\vec{x}_{n_c+1})\,\langle\phi_c|\widetilde{\phi}_{n_c+1}\rangle)\dots \\ &\quad (\widetilde{\phi}_n(\vec{x}_n) - \sum_c^{n_c}\phi_c(\vec{x}_n)\,\langle\phi_c|\widetilde{\phi}_n\rangle)| \end{aligned} \tag{2.29}$$

Linearity of $\hat{P}_v$ allows to exploit the rules for determinants. I.e., the determinant can be expanded in sums over determinants. However, in all but one determinant a core orbital occurs at least twice. Every determinant with two identical orbitals is zero and one gets

$$\begin{aligned} \Phi &= |\phi_1(\vec{x}_1)\dots\phi_{n_c}(\vec{x}_{n_c})\phi_{n_c+1}(\vec{x}_{n_c+1})\dots\phi_n(\vec{x}_n)| \\ &= |\phi_1(\vec{x}_1)\dots\phi_{n_c}(\vec{x}_{n_c})\widetilde{\phi}_{n_c+1}(\vec{x}_{n_c+1})\dots\widetilde{\phi}_n(\vec{x}_n)| \\ &= \widetilde{\Phi} \end{aligned} \tag{2.30}$$

The valence projection operator thus transforms a pseudo wave function $\widetilde{\Phi}$ into the wave function Φ with orbitals $\phi_{n_{c+1}} \dots \phi_n$ orthogonal to $\phi_1 \dots \phi_{n_c}$.

$$\begin{aligned} \Phi &= \hat{P}_v \Phi \\ &= \prod_{i=n_c+1}^{n} \hat{P}_v(i) \Phi \\ &= \prod_{i=n_c+1}^{n} \left(\hat{1}(i) - \hat{P}_c(i)\right) \widetilde{\Phi} \end{aligned} \tag{2.31}$$

A variational ansatz for the matrix element $\langle \Phi | \hat{H} | \Phi \rangle$

$$\begin{aligned} \delta \langle (\hat{1} - \hat{P}_c)\widetilde{\Phi} | \hat{H} | (\hat{1} - \hat{P}_c)\widetilde{\Phi} \rangle - E \langle (\hat{1} - \hat{P}_c\widetilde{\Phi}) | (\hat{1} - \hat{P}_c)\widetilde{\Phi} \rangle = \\ \langle (\hat{1} - \hat{P}_c)\widetilde{\Phi} | \hat{H} | (\hat{1} - \hat{P}_c)\widetilde{\Phi} \rangle - E \langle \widetilde{\Phi} | (\hat{1} - \hat{P}_c)\widetilde{\Phi} \rangle = 0 \end{aligned} \tag{2.32}$$

is used to obtain a pseudo-eigenvalue equation

$$\left(\hat{H} + \hat{V}_{\mathrm{GPK}}\right) \widetilde{\Phi} = E \widetilde{\Phi} \tag{2.33}$$

with

$$\hat{V}_{\mathrm{GPK}} = -\hat{H}\hat{P}_c - \hat{P}_c\hat{H} + \hat{P}_c\hat{H}\hat{P}_c + E\hat{P}_c. \tag{2.34}$$

The energy E in $\hat{V}_{\mathrm{GPK}}$ is fixed to the lowest eigenvalue E_{min} of (2.33).Using for $\hat{H}$ the HF Hamiltonian and spanning $\hat{P}_c$ in the ACOs, the PK equations are obtained for a single valence orbital. However, $\hat{P}_c$ is not restricted to be defined in the basis of HF core orbitals and the requirement for orthogonality of $\widetilde{\Phi}$ w.r.t. the AE core orbitals is thus lifted. Kleiner and McWeeny connected in Ref. [81] the group function approach for one core and one valence determinant to the PK PP, but adopted a more flexible PP for their purposes. Weeks and Rice state in [17] that the variation potentials of Cohen and Heine [18] can be derived as a special case of V_{GPK}.

Apart from the FC approximation for a single core-determinant, the preceding discussion was in principle exact. However, the definition of V_{GPK} still requires knowledge about the core orbitals. Practical use requires to remove these references to avoid repeated recalculation of the chemically less important core orbitals. Phillips and Kleinman already mention in Ref. [16] that the PP has to be different for different symmetries of the valence electron. Abarenkov and Heine [82], Kahn and Goddard III [83] and Schwarz [84] proposed PPs that depend on the angular momentum quantum number l.

$$V_{\mathrm{PP}}(r) = \sum_{l=0}^{N} v_l(r) \, |l\rangle\langle l| \tag{2.35}$$

where $|l\rangle\langle l|$ is the projection operator that projects out a specific angular momentum quantum number l. In Ref. [85] Kahn, Baybutt and Truhlar extended the definition of the PP.

$$V_{\mathrm{PP}}(r) = v_{L+1}(r) + \sum_{l=0}^{L} \sum_{m=-l}^{l} \left[v_l(r) - v_{L+1}(r)\right] |lm\rangle\langle lm| \tag{2.36}$$

v_{L+1} is a local potential acting on orbitals of all symmetries. For each symmetry l that is included in the core, an additional semi-local potential v_l is chosen. The potentials v_l are expanded in a series of Gaussian functions that fulfill the condition

$$r^2 \left(v_l(r) - \frac{N_c}{r} \right) = \sum_k A_{kl} r^{n_{kl}} \mathrm{e}^{-\alpha_{kl} r^2} \tag{2.37}$$

N_c is the number of core electrons included in the PP. The parameters A_{kl}, n_{kl} and α_{kl} are adjusted to a numerical form of $v_l(r)$ obtained from AE calculations.

Modern PPs all use potentials of the form

$$v_l(r) = -\frac{Z - N_c}{r} + \sum_k A_{kl} r^{-n_{kl}} \mathrm{e}^{-\alpha_{kl} r^2} \tag{2.38}$$

where the requirements on the PP define the specific choice of parameters. Z is the charge of the bare nucleus. Schwartz and Switalski [86] chose this Gaussian-screened radial potential to facilitate the calculation of PP matrix elements. Although Hellmann used an exponential instead of a Gaussian screening for his PPs, they had essentially the same form. In the spirit of Hellmann's work in the year 1935, Schwartz and Switalski [86] and Switalski and Schwartz [87] decided to optimize the PP parameters w.r.t. valence ionization energies.

Flad, Stoll and Preuß [88] and Preuß *et al.* [89] used PPs of type (2.38), fitted to experimental ionization energies of first, second and third row single-valence electron atoms, for density functional theory (DFT) calculations of molecular clusters. In the year 1982, Fuentealba *et al.* [90] included a core polarization term in the PP in order to account for major core-valence correlation effects. Dolg *et al.* [77] also adapted PPs of the Schwartz type, but optimized them to the HF valence spectra of the atoms Sc through Zn. They used single- and multi-electron fits in order to obtain their results. McMurchie and Davidson published in Ref. [91] their implementation of PP operator matrix elements. They found exact formulas for the occurring angular integrals, whereas they chose approximations to efficiently calculate the more involved radial integrals.

2.2.3. Energy-Consistent Pseudo Potentials

In Refs. [76],[92, Ch. 20], Dyall analyzed the reformulation of the HF equations in terms of valence POs. He showed that it is impossible to formulate a single Hermitian GPK PP for which all POs of the same symmetry are orthogonal. This problem is overcome by using the exact GPK PP for the lowest energy PO of a certain symmetry as an approximation to all other POs of that symmetry. The nonorthogonality of the POs is shown by expanding the PO in the AE AOs and noticing that the *core-tail overlap integral* $\langle \phi_i^{\mathrm{R}} | \phi_j^{\mathrm{R}} \rangle$ is generally nonzero.

$$\langle \widetilde{\phi}_i | \widetilde{\phi}_j \rangle = \langle \phi_i | \phi_j \rangle + \sum_{c,c'} \langle \phi_c | \widetilde{\phi}_i \rangle \, \langle \widetilde{\phi}_j | \phi'_c \rangle = \delta_{ij} + \langle \phi_i^{\mathrm{R}} | \phi_j^{\mathrm{R}} \rangle \,. \tag{2.39}$$

Expanding the Fock matrix in the these nonorthogonal POs $\widetilde{\phi}_i$ introduces a different metric than unity which effectively leads to a compression of the (valence) eigenvalue spectrum compared to the HF spectrum. Furthermore, the POs are not normalized anymore, in contrast to the true valence eigenfunctions of the HF Hamiltonian, ϕ_i. If the POs would be chosen normalized, then the consequence, according to Dyall, is the alteration of the valence two-electron interaction and the PP. The effect is reduced Coulomb repulsion and a more attractive PP in the spatial valence region. SC orbitals try to circumvent this shortcoming by enforcing the norm of a PO outside the core to match that of the corresponding AE AO exactly. However, this ansatz changes the valence energy and corresponding PPs are state-specific.

The PPs of the Stuttgart type, i.e., those of Preuß, Stoll and Dolg, are energy-consistent. I.e., the PP is fitted such that the valence electrons' excitation spectrum of an atom coincides with a reference spectrum. This procedure generates energetically averaged PP, which are suitable for a proper description of valence properties in many systems (see last paragraph of Ref. [76] or [92, p. 414]).

The functional S, which is minimized to obtain a EC PP, reads

$$S = \sum_{\mathcal{I}}^{\mathcal{K}} v_{\mathcal{I}} \left(\widetilde{E}_{\mathcal{I}} - E_{\mathrm{ref},\mathcal{I}} + \Delta E \right)^2 . \tag{2.40}$$

$\widetilde{E}_{\mathcal{I}}$ is the energy of state $\widetilde{\Psi}_{\mathcal{I}}$ calculated with a PP, $E_{\mathrm{ref},\mathcal{I}}$ the energy of the corresponding AE reference state, ΔE a parameter that determines the offset between the absolute energies of PP and reference state energies and $v_{\mathcal{I}}$ a weight factor. The sketch in Fig. 2.1 shows how the states in the reference system relate to the states in the PP system. The parameters A_{kl} and α_{kl} of (2.38) are optimized by using a least-squares algorithm. A typical choice for the reference energies are

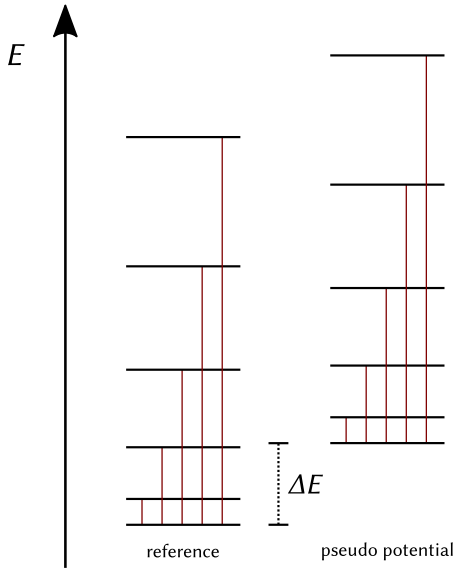


Figure 2.1.: Sketch of an energy diagram corresponding to an AE and a PP system. The PP is fit such that the spectra coincide. An additional parameter ΔE allows flexibility for absolute positioning of the levels.

the energies of multi-configuration Dirac–Hartree–Fock (MCDHF) calculations including the Breit correction. The inactive space of such a calculation incorporates all core electrons. The PP system is then calculated by a similar method, e.g. multi-configuration Hartree–Fock (MCHF), employing the same configuration space as for the reference. Modern fits are performed in a two-component formalism and subsequently spin-orbit averaged to obtain SR PPs for the use in NR calculations. Metz [93] derived analytic first and second order derivatives of S w.r.t. the PP parameters for an efficient optimization using a Quasi-Newton Levenberg–Marquardt algorithm.

Cao and Dolg state in Ref. [25] that the shape of pseudo valence orbitals (PVOs) agrees very well with that of DHF AE atomic valence orbitals (AVOs) in the spatial valence region. In his dissertation, Metz demonstrated the same findings for various atoms. [93] The similarity for the Ba 5s and 6s orbital is showcased in Fig. 2.2. The POs were obtained using the NR unrestricted Hartree–Fock (UHF) method and a `def2-TZVP` basis set. A SR ZORA UHF scheme was used for the AE calculation, employing the basis set `even-temp_v7` (see App. A).

2.3. MCSCF Ansatz

HF theory provides the means to obtain reasonable electronic structures for many closed-shell ground state molecules, as long as no atoms with complicated spin situations are involved. If dynamic correlation effects are important, perturbation theory (PT), CI and coupled cluster (CC) methods represent a swiss-army knife to tackle these effects on the basis of the HF many-electron wave function. Full configuration interaction (Full-CI) even gives the exact total energies in a given basis, but is usually prohibitively expensive to calculate. However, when the HF ground state is a bad reference function, the calculation of dynamic correlation effects cannot improve the results. Including multiple reference wave functions in the CI, incorporates static correlation and significantly improves electronic structures of these systems. This class of methods is usually called *multi reference (MR)* methods.

In the application of *Multi-configuration (MC)* methods usually a CI calculation is performed, but the MO coefficients of the HF reference wave function are optimized simultaneously to the CI coefficients. An overview over the equations involved can be found in Ref. [94, Ch. 12]. The MC methods come in different fla-

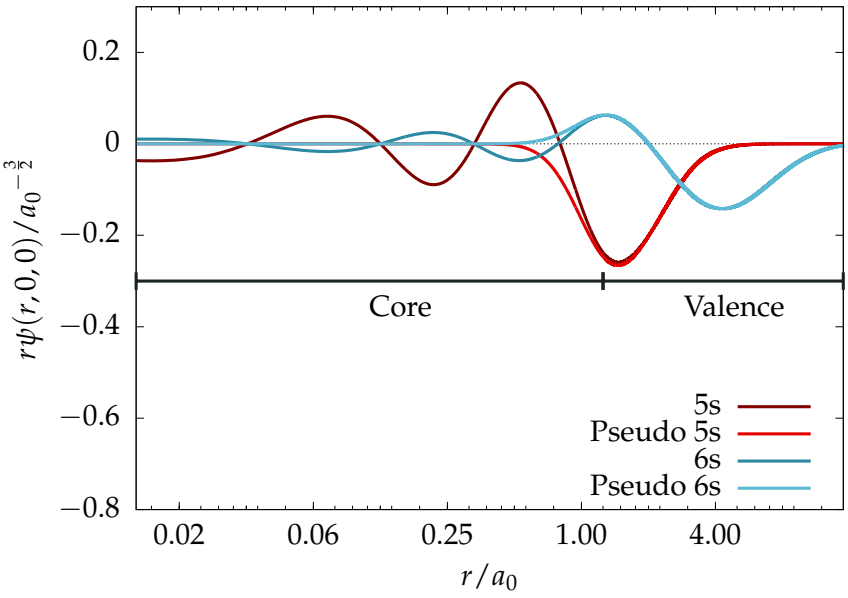


Figure 2.2.: Comparison of Ba 5s and 6s AE AOs and POs. In dark colors the SR AE orbitals are shown, whereas light colors indicate there PP counterparts. The orbitals are in good agreement outside the core region, whereas they differ within the core region due to the reduced nodal structure of POs.

vors, namely as single-state, state-specific (SS) and state-averaged (SA) methods. Single-state MCSCF wave functions are optimized on *one* previously selected electronic state. All other states are usually not well described. In contrast, SS-MCSCF wave functions are constructed from *several* non-orthogonal reference states where each is optimized to a single electronic state. The prize to pay is the calculation of Hamiltonian matrix elements over non-orthogonal MOs since each state is now expanded in a different set of MOs. SA-MCSCF circumvents this by defining an average energy E which equals the sum of all $\mathcal{N}$ weighted state energies $E_{\mathcal{I}}$.

$$E = \sum_{\mathcal{I}}^{\mathcal{N}} w_{\mathcal{I}} E_{\mathcal{I}} = \sum_{\mathcal{I}} w_{\mathcal{I}} \langle \Psi_{\mathcal{I}} | \hat{H} | \Psi_{\mathcal{I}} \rangle \tag{2.41}$$

$w_{\mathcal{I}}$ are weighting coefficients, which are in general not identical to the ones used in (2.40). The electronic states $\Psi_{\mathcal{I}}$ are generated from a common set of orbitals and SDs. $\hat{H}$ is the Hamiltonian, which can be expanded in a second quantization formulation in a set of orbitals $\{\psi_i\}$. I assume here that the Hamiltonian contains only one-electron contributions

$$h_{pq} = \int \psi_p^*(\vec{x}_1) \hat{h}(\vec{x}_1) \psi_q(\vec{x}_1) \, \mathrm{d}\vec{x}_1 \tag{2.42}$$

and two-electron contributions

$$g_{pqrs} = \int \psi_p^*(\vec{x}_1) \psi_q(\vec{x}_1) \hat{g}(\vec{x}_1, \vec{x}_2) \psi_r^*(\vec{x}_2) \psi_s(\vec{x}_2) \, \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 \quad . \tag{2.43}$$

The zero-electron contributions (e.g., repulsion of nuclei) have been regarded as a constant shift to the total energy (Born–Oppenheimer approximation) and thus do not appear in the electronic Hamiltonian anymore. The full operator is then

$$\hat{H} = \sum_{pq} h_{pq} \hat{E}_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} \hat{e}_{pqrs} \tag{2.44}$$

$\hat{E}_{pq}$ is the one-electron singlet excitation operator that excites one electron from spin-orbital ψ_p to spin-orbital ψ_q.

$$\hat{E}_{pq} = \hat{a}_{\alpha,p}^\dagger \hat{a}_{\alpha,q} + \hat{a}_{\beta,p}^\dagger \hat{a}_{\beta,q} = \sum_{\sigma=\alpha,\beta} \hat{a}_{\sigma,p}^\dagger \hat{a}_{\sigma,q} \tag{2.45}$$

$\hat{a}_i^\dagger$ creates an electron in orbital i, if the operator acts on a ket-state.

$$\hat{a}_i^\dagger \, |\mathrm{vac}\rangle = |\psi_i\rangle \tag{2.46}$$

$\hat{a}_i$ is the corresponding annihilation operator. $\hat{e}_{pqrs}$ is the two-electron excitation operator.[3]

$$\hat{e}_{pqrs} = \sum_{\sigma,\tau} \hat{a}^\dagger_{\sigma,p} \hat{a}^\dagger_{\tau,r} \hat{a}_{\tau,s} \hat{a}_{\sigma,q} \tag{2.47}$$

Following the arguments of Ref. [94, Ch. 12.7], one can parametrize the SA energy using the orbital rotation operator $\hat{\kappa}$ and the state-transfer operator $\hat{R}$ in their real valued form.

$$\hat{\kappa} = \sum_{p>q} \kappa_{pq} E^-_{pq} \tag{2.48}$$

$$\hat{E}^-_{pq} = \hat{E}_{pq} - \hat{E}_{qp} \tag{2.49}$$

$$\hat{R} = \sum_{\substack{\mathcal{K}>\mathcal{J} \\ \mathcal{J}\leq\mathcal{N}}} R_{\mathcal{K}\mathcal{J}} \left(|\Psi_\mathcal{K}\rangle \langle\Psi_\mathcal{J}| - |\Psi_\mathcal{J}\rangle \langle\Psi_\mathcal{K}| \right) \tag{2.50}$$

Index $\mathcal{J}$ runs over all states included in the averaging, whereas index $\mathcal{K}$ runs over the orthogonal complement as well. The parametrized energy $\bar{E}$ is

$$\bar{E}(\boldsymbol{\kappa}, \boldsymbol{R}) = \sum_{\mathcal{I}}^{\mathcal{N}} w_\mathcal{I} \langle\Psi_\mathcal{I}| \mathrm{e}^{-\hat{R}} \mathrm{e}^{-\hat{\kappa}} \hat{H} \mathrm{e}^{\hat{\kappa}} \mathrm{e}^{\hat{R}} |\Psi_\mathcal{I}\rangle \tag{2.51}$$

The elements κ_{pq} and $R_{\mathcal{K}\mathcal{J}}$ of the matrices $\boldsymbol{\kappa}$ and $\boldsymbol{R}$ can then be iteratively adjusted in order to minimize the total energy $\bar{E}$. Optimization of the SA energy is performed using a second order method, e.g., the Newton–Raphson method or SCF methods. Helgaker, Jørgensen and Olsen favor the versatile second order methods due to their simple application to a range of different methods. [94, Ch. 10.9.6]

Methods based on the Newton–Raphson scheme usually expand the total energy $\bar{E}$ in a Taylor series around $\vec{\lambda} = \vec{0}$ up to second order, where $\vec{\lambda}$ is the column vector containing all parameters.

$$\bar{E} \approx E^{(0)} + \vec{\lambda}^T \vec{E}^{(1)} + \frac{1}{2} \vec{\lambda}^T \boldsymbol{E}^{(2)} \vec{\lambda} + \mathcal{O}(\vec{\lambda}^3) \tag{2.52}$$

The superscript in parentheses denote the derivative w.r.t. the parameters in $\vec{\lambda}$, i.e., $\vec{E}^{(1)}$ denotes the gradient and $\boldsymbol{E}^{(2)}$ the Hessian. The zeroth order term is

[3] Refer to Ch.1–3 of Ref. [94] for a conclusive introduction in the topic of second quantization.

given by

$$E^{(0)} = \sum_{\mathcal{I}} w_{\mathcal{I}} \langle \Psi_{\mathcal{I}} | \hat{H} | \Psi_{\mathcal{I}} \rangle \tag{2.53}$$

$$= \sum_{\mathcal{I}} w_{\mathcal{I}} \left(\sum_{pq} h_{pq} \langle \Psi_{\mathcal{I}} | \hat{E}_{pq} | \Psi_{\mathcal{I}} \rangle + \frac{1}{2} \sum_{pqrs} g_{pqrs} \langle \Psi_{\mathcal{I}} | \hat{e}_{pqrs} | \Psi_{\mathcal{I}} \rangle \right) \tag{2.54}$$

$$= \sum_{\mathcal{I}} w_{\mathcal{I}} \left(\sum_{pq} h_{pq} D_{pq}^{\mathcal{I}\mathcal{I}} + \frac{1}{2} \sum_{pqrs} g_{pqrs} P_{pqrs}^{\mathcal{I}\mathcal{I}} \right) \quad , \tag{2.55}$$

where $D_{pq}^{\mathcal{I}\mathcal{I}}$ and $P_{pqrs}^{\mathcal{I}\mathcal{I}}$ are the one- and two-electron density matrices for the state $\Psi_{\mathcal{I}}$.

For the SA-MCSCF wave function, the gradient w.r.t. the orbital rotation parameters is

$$\vec{E}^{(1)}(\kappa_{mn}) = \sum_{\mathcal{I}}^{\mathcal{N}} w_{\mathcal{I}} \langle \Psi_{\mathcal{I}} | \left[\hat{E}_{mn}^{-}, \hat{H} \right] | \Psi_{\mathcal{I}} \rangle \tag{2.56}$$

and w.r.t. the state transfer parameters

$$\vec{E}^{(1)}(R_{\mathcal{K}\mathcal{J}}) = -2 \left(w_{\mathcal{J}} - \sum_{\mathcal{I}}^{\mathcal{N}} \delta_{\mathcal{K}\mathcal{I}} w_{\mathcal{I}} \right) \langle \Psi_{\mathcal{J}} | \hat{H} | \Psi_{\mathcal{K}} \rangle \, . \tag{2.57}$$

The matrix elements of the Hessian have been reported elsewhere [94, p. 639].

It is possible to choose a linear parametrization of the electronic states instead. Each state is then expanded in a set of SDs or configurations.

$$|\Psi_{\mathcal{I}}\rangle = \sum_{I} C_{I\mathcal{I}} |\Phi_I\rangle \tag{2.58}$$

This changes the gradient w.r.t. to the expansion coefficients.

$$E^{(1)}(C_{I\mathcal{I}}^{*}) = w_{\mathcal{I}} \sum_{J} \left(\langle \Phi_I | \hat{H} | \Phi_J \rangle - E_{\mathcal{I}} \langle \Phi_I | \Phi_J \rangle \right) C_{IJ} \tag{2.59}$$

The price to pay is that the coefficients $C_{I\mathcal{I}}$ need to fullfil the condition for unitarity as well, which is automatically fullfilled for the coefficients κ_{pq} and $R_{\mathcal{K}\mathcal{J}}$. The advantage of any second order (Newton–Raphson-like) optimization scheme is that it simplifies the change from one parametrization to another as only the definition of the gradient and Hessian have to be changed.

3. Development of a Reconstruction Method

In this chapter I discuss the reconstruction of AE molecular orbitals from PMOs. I start with the fundamental conditions for a useful reconstruction scheme in Sec. 3.1. The theoretical foundation of the scheme is discussed in Sec. 3.2, where also the implemented equations are found through discussion of various model systems. Links to a general procedure are indicated. I continue by providing an implementation protocol in Sec. 3.3, which is complemented by a simple numerical example in Sec. 3.4. In Sec. 3.6 numerical tests are discussed and the general applicability of the scheme is shown. Sec. 3.7 summarizes the results presented in this chapter.

3.1. Requirements on the Method

Since the developments of Hellmann and Gombás, PP approximations to the ECP are obtained in EC ways, i.e., the POs are ensured to reproduce orbital energies, electron excitation spectra or other energetic quantities of AE calculations. I aim in this work for a transformation of Stuttgart POs to their AE analogues, as the Stuttgart PPs introduced in Sec. 2.2 are widely used nowadays and, e.g., the standard PPs employed when using the program TURBOMOLE [95]. The following criteria should guide the development:

Applicability The scheme should be applicable to molecules as well as atoms.

Black-box The average user should not need any knowledge about the procedure.

Efficiency The reconstruction procedure should be more efficient than a corresponding (scalar-) relativistic calculation. Furthermore, it should serve as a good starting point for subsequent AE wave function reoptimizations.

Accuracy The reconstructed orbitals should be in good agreement with corresponding relativistic calculations.

Generality The scheme should be in principle exact. Any deviation w.r.t. AE calculations should be rooted in the semi-empirical nature of used PPs. All extensions should be derivable from the principle set of equations.

Practicality The scheme should be applicable to as many types of PP as possible.

Backwards compatibility The scheme should be compatible with *old wave functions*. I.e., those wave functions can be used in the scheme without further modification or even a recalculation with a modified SCF algorithm.

Already at this point it is clear that some of those requirements are mutually exclusive. E.g., *accuracy* usually drops with *efficiency* and *generality* does not comply with *practicality*, since a scheme which can be used for every PP cannot be derivable from a special type of PP. In the following, I will, however, prioritize *efficiency* and *practicality*.

3.2. Properties of Energy-Consistent Pseudo Potentials

In this section, I discuss the connection between energy and shape of pseudo orbitals. I start with a discussion of several model systems: The first consists of a HF ground state and two ionic Koopmans' states, i.e., states created through addition or removal of an electron from the HF determinant. I show that for this model system EC PP generate orbital energies that coincide with the orbital energies of the AE HF ground state. The next model systems are also of the *Koopmans type*, including four and five states. Additionally, a configuration interaction singles (CIS) model is discussed. For these systems, the PO energies are shown to coincide approximately with AE orbital energies. Subsequently, the form of EC PP is discussed and compared to PP of the PK type. The general expression for a PP in the MCSCF case is presented within the natural orbital approximation and links to the discussed models are illustrated. The POs are then shown to be SC within the limits of the model systems. Finally, the SC of EC POs is discussed using the example of the Ba 5s orbital.

3.2.1. Koopmans Three-States Approximation

The first model system consists of a HF ground state Ψ_0, a Koopmans cationic states $\Psi_{\mathcal{C}}$ and a Koopmans anionic state $\Psi_{\mathcal{A}}$. The states have the following electron configuration:

$$\begin{array}{cccc} \phi_b & — & — & \uparrow \\ \phi_a & \uparrow\downarrow & \uparrow & \uparrow\downarrow \\ \phi_i & \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow \\ & \Psi_0 & \Psi_{\mathcal{C}} & \Psi_{\mathcal{A}} \end{array}$$

ϕ_i denotes an inactive, i.e., a core, orbital, which is only present in the AE system and not in the pseudo system. The Hamiltonian for this three state system is diagonal, as states with different particle numbers do not couple.

$$\boldsymbol{H} = \begin{pmatrix} E_0 & 0 & 0 \\ 0 & E_{\mathcal{C}} & 0 \\ 0 & 0 & E_{\mathcal{A}} \end{pmatrix} \tag{3.1}$$

The SA energy is

$$E = w_0 \langle \Psi_0 | \hat{H} | \Psi_0 \rangle + w_{\mathcal{C}} \langle \Psi_{\mathcal{C}} | \hat{H} | \Psi_{\mathcal{C}} \rangle + w_{\mathcal{A}} \langle \Psi_{\mathcal{A}} | \hat{H} | \Psi_{\mathcal{A}} \rangle \tag{3.2}$$

with the weights

$$\begin{aligned} w_0 &= 1 \\ w_{\mathcal{C}} = w_{\mathcal{A}} &= 0 \end{aligned} \tag{3.3}$$

I.e., the minimum SA energy is identical to the energy of the HF ground state. Other weights are possible, but then the optimized orbitals are not necessarily HF orbitals.

Some arbitrary PP shall now be parametrized to match the spectrum of the all-electron valence Hamiltonian. This can be ensured by minimization of the functional S in (2.40). For an ideal PP, S equals 0 and one gets for all terms individually

$$\begin{aligned} v_{\mathcal{I}} \left(\widetilde{E}_{\mathcal{I}} - E_{\mathcal{I}} + \Delta E \right) &= 0 \\ \widetilde{E}_{\mathcal{I}} - E_{\mathcal{I}} + \Delta E &= 0 \qquad , \end{aligned} \tag{3.4}$$

assuming that $v_{\mathcal{I}} \neq 0$. (3.4) holds for all states included in the optimization functional S. Thus, one can equate (3.4) for two distinct states $\mathcal{I}$ and $\mathcal{J}$.

$$\begin{aligned} \widetilde{E}_{\mathcal{I}} - E_{\mathcal{I}} + \Delta E &= \widetilde{E}_{\mathcal{J}} - E_{\mathcal{J}} + \Delta E \\ \widetilde{E}_{\mathcal{I}} - \widetilde{E}_{\mathcal{J}} &= E_{\mathcal{I}} - E_{\mathcal{J}} \end{aligned} \tag{3.5}$$

I.e., the PP generates the same energy spectrum as the AE system.

Now, what is the effect of an ideal PP on the POs? Koopmans' theorem approximates the ionization energy of a HF system by the orbital energy of the ionized orbital. The total energies of all investigated AE states are

$$\begin{aligned} E_0 &= 2h_{ii} + J_{ii} + 2h_{aa} + J_{aa} + 4J_{ia} - 2K_{ia} && (3.6)\\ &= 2\epsilon_i + 2\epsilon_a - J_{aa} - J_{ii} - 4J_{ia} + 2K_{ia} \\ E_{\mathcal{C}} &= 2h_{ii} + J_{ii} + h_{aa} + 2J_{ia} - K_{ia} && (3.7)\\ &= 2\epsilon_i + \epsilon_a - J_{aa} - J_{ii} - 4J_{ia} + K_{ia} \\ E_{\mathcal{A}} &= 2h_{ii} + J_{ii} + 2h_{aa} + J_{aa} + 4J_{ia} - 2K_{ia} + h_{bb} + 2J_{ib} - K_{ib} + 2J_{ab} - K_{ab} && (3.8)\\ &= 2\epsilon_i + 2\epsilon_a - J_{aa} - J_{ii} - 4J_{ia} + 2K_{ia} + \epsilon_b \end{aligned}$$

and those of the pseudo states are

$$\begin{aligned} \widetilde{E}_0 &= 2\widetilde{h}_{aa} + \widetilde{J}_{aa} = 2\widetilde{\epsilon}_a - \widetilde{J}_{aa} && (3.9)\\ \widetilde{E}_{\mathcal{C}} &= \widetilde{h}_{aa} = \widetilde{\epsilon}_a - \widetilde{J}_{aa} && (3.10)\\ \widetilde{E}_{\mathcal{A}} &= 2\widetilde{h}_{aa} + \widetilde{J}_{aa} + \widetilde{h}_{bb} + 2\widetilde{J}_{ab} - \widetilde{K}_{ab} = 2\widetilde{\epsilon}_a - \widetilde{J}_{aa} + \widetilde{\epsilon}_b && (3.11) \end{aligned}$$

The energy differences are

$$\begin{aligned} E_0 - E_{\mathcal{C}} &= \epsilon_a && (3.12)\\ E_{\mathcal{A}} - E_0 &= \epsilon_b && (3.13) \end{aligned}$$

and the corresponding differences for the pseudo system are

$$\begin{aligned} \widetilde{E}_0 - \widetilde{E}_{\mathcal{C}} &= \widetilde{\epsilon}_a && (3.14)\\ \widetilde{E}_{\mathcal{A}} - \widetilde{E}_0 &= \widetilde{\epsilon}_b \qquad . && (3.15) \end{aligned}$$

Thus, the ideal PP ensures orbital energy consistency (compare to (3.5)).

$$\begin{aligned} \widetilde{\epsilon}_a &= \epsilon_a \\ \widetilde{\epsilon}_b &= \epsilon_b \end{aligned} \qquad (3.16)$$

Inserting the corresponding energies in (3.4), one finds the definition of ΔE.

$$\Delta E = 2\epsilon_i - J_{ii} - 4J_{ia} + K_{ia} \qquad (3.17)$$

The AE Fock matrix is defined by the matrix elements in the MO basis (see [94, p. 622])

$$F_{mn} = \sum_q h_{nq} D_{mq} + \sum_{qrs} g_{nqrs} P_{mqrs} \qquad , \qquad (3.18)$$

with the one- and two-electron integrals h_{pq} and g_{pqrs} and the one- and two-electron density matrices D_{pq} and P_{pqrs} defined as in Ch. 2.3. Its valence orbital part coincides only with the PP Fock matrix for state Ψ_0, the HF ground state.

$$\boldsymbol{F}_0 = \begin{pmatrix} \epsilon_i & 0 & 0 & 0 \\ 0 & \epsilon_i & 0 & 0 \\ 0 & 0 & \epsilon_a & 0 \\ 0 & 0 & 0 & \epsilon_a \end{pmatrix} \quad , \widetilde{\boldsymbol{F}}_0 = \begin{pmatrix} \widetilde{\epsilon}_a & 0 \\ 0 & \widetilde{\epsilon}_a \end{pmatrix} \tag{3.19}$$

The AE valence and pseudo Fock matrices of the ionic states differ. E.g., the Fock matrix of the cationic system reads

$$\boldsymbol{F}_{\mathcal{C}} = \begin{pmatrix} \epsilon_i - J_{ia} & 0 & h_{ia} - g_{aiaa} \\ 0 & \epsilon_i - J_{ia} + K_{ia} & 0 \\ h_{ia} - g_{iaaa} & 0 & \epsilon_a - J_{aa} \end{pmatrix} \quad , \widetilde{\boldsymbol{F}}_{\mathcal{C}} = \begin{pmatrix} \widetilde{\epsilon}_a - \widetilde{J}_{aa} \end{pmatrix} \tag{3.20}$$

Summarizing, one could also fit the PP using the valence Fock matrices of the neutral system. But using the ionic Fock matrices instead could give a different PP.

3.2.2. Koopmans Four- and Five-States Approximation

Complexity of the previously discussed problem can be increased by adding a second anionic state, where orbital ϕ_c is occupied by a β electron.

ϕ_c	—	—	—	↓
ϕ_b	—	—	↓	—
ϕ_a	↑↓	↓	↑↓	↑↓
ϕ_i	↑↓	↑↓	↑↓	↑↓
	Ψ_0	$\Psi_{\mathcal{C}}$	$\Psi_{\mathcal{A}1}$	$\Psi_{\mathcal{A}2}$

By analogy with the previous example one can calculate the four energies of the states. The first three total energies are identical to E_0, $E_{\mathcal{C}}$ and $E_{\mathcal{A}}$ from the previous example. The remaining energy is

$$E_{\mathcal{A}2} = 2h_{ii} + J_{ii} + 2h_{aa} + J_{aa} + 4J_{ia} - 2K_{ia} + h_{cc} + 2J_{ic} - K_{ic} + 2J_{ac} - K_{ac} \tag{3.21}$$

and

$$\widetilde{E}_{\mathcal{A}2} = 2\widetilde{h}_{aa} + \widetilde{J}_{aa} + \widetilde{h}_{cc} + 2\widetilde{J}_{ac} - \widetilde{K}_{ac} \quad , \tag{3.22}$$

respectively. Then, the difference $E_{\mathcal{A}2} - E_0$ is

$$E_{\mathcal{A}2} - E_0 = \epsilon_c \tag{3.23}$$

and for the pseudo systems

$$\widetilde{E}_{\mathcal{A}2} - \widetilde{E}_0 = \widetilde{\epsilon}_c \tag{3.24}$$

I.e., in this example again, the AE orbital energies are reproduced exactly for the pseudo system using an ideal PP.

Adding now an additional state with another electron in orbital ϕ_d creates additional couplings due to two-electron interaction of orbitals ϕ_b and ϕ_d.

ϕ_d	—	—	—	—	↿
ϕ_c	—	—	—	↿	—
ϕ_b	—	—	↿	—	↿
ϕ_a	⇂	↿	⇂	⇂	⇂
ϕ_i	⇂	⇂	⇂	⇂	⇂
	Ψ_0	Ψ_C	$\Psi_{\mathcal{A}1}$	$\Psi_{\mathcal{A}2}$	$\Psi_{\mathcal{A}3}$

The first four energy differences are exactly as before. However, the fourth total energy difference,

$$E_{\mathcal{A}3} - E_{\mathcal{A}1} = \epsilon_d + J_{bd} - K_{bd} \tag{3.25}$$

contains also two-electron interaction energies. An ideal PP that reproduces the spectrum of this Hamiltonian, will thus reproduce exactly the HF orbital energies ϵ_a through ϵ_c, whereas the ϵ_d is only matched by $\widetilde{\epsilon}_d$ up to a constant Δ_{bd}.

$$\begin{aligned} \epsilon_d &= \widetilde{\epsilon}_d + \Delta_{bd} \\ \Delta_{bd} &= (\widetilde{J}_{bd} - J_{bd}) - (\widetilde{K}_{bd} - K_{bd}) \end{aligned} \tag{3.26}$$

Note that this notation does *not* imply differences in the two-electron interaction energies alone! The one-electron contributions to the orbital energies ϵ_d and $\widetilde{\epsilon}_d$ may differ as well.

3.2.3. CIS Approximation

In the next example, I will discuss the effect of a CIS approximation. Again, I start with the definition of three states.

ϕ_c	—	—	↿
ϕ_b	—	↿	—
ϕ_a	⇂	↿	↿
ϕ_i	⇂	⇂	⇂
	Ψ_0	$\Psi'_{\mathcal{A}}$	$\Psi'_{\mathcal{B}}$

The excited states are primed, as they are not eigenstates of the Hamiltonian. In the basis of these states, the Hamiltonian reads

$$\boldsymbol{H}' = \begin{pmatrix} E_0 & 0 & 0 \\ 0 & E_{\mathcal{A}} & E' \\ 0 & E' & E_{\mathcal{B}} \end{pmatrix} \tag{3.27}$$

Brioullin's theorem formally decouples the singly excited block from the HF ground state. Assuming the off-diagonal elements E' to be small, one can expand the diagonalized Hamiltonian $\boldsymbol{H}$ in powers of E'.

$$\boldsymbol{H} \approx \begin{pmatrix} E_0 & 0 & 0 \\ 0 & E_{\mathcal{A}} & 0 \\ 0 & 0 & E_{\mathcal{B}} \end{pmatrix} + \begin{pmatrix} E_0 & 0 & 0 \\ 0 & -\frac{E'^2}{E_{\mathcal{B}}-E_{\mathcal{A}}} & 0 \\ 0 & 0 & \frac{E'^2}{E_{\mathcal{B}}-E_{\mathcal{A}}} \end{pmatrix} \tag{3.28}$$

The first term alone is correct up to first order, whereas the sum is correct up to third order. Considering the first term only, an EC PP can be fitted to the energy differences $E_{\mathcal{B}} - E_{\mathcal{A}}$ and $E_{\mathcal{A}} - E_0$. The energies in the first order approximation are

$$E_{\mathcal{A}} = 2\epsilon_i + \epsilon_b - J_{ii} - 2J_{ia} + K_{ia} - J_{ab} + K_{ab} + h_{aa} \tag{3.29}$$
$$E_{\mathcal{B}} = 2\epsilon_i + \epsilon_c - J_{ii} - 2J_{ia} + K_{ia} - J_{ac} + K_{ac} + h_{aa} \tag{3.30}$$

and

$$\widetilde{E}_{\mathcal{A}} = \widetilde{\epsilon}_b - \widetilde{J}_{ab} + \widetilde{K}_{ab} + \widetilde{h}_{aa} \tag{3.31}$$
$$\widetilde{E}_{\mathcal{B}} = \widetilde{\epsilon}_c - \widetilde{J}_{ac} + \widetilde{K}_{ac} + \widetilde{h}_{aa} \quad . \tag{3.32}$$

The state energy differences are then

$$E_{\mathcal{A}} - E_0 = \epsilon_b - \epsilon_a - J_{ab} + K_{ab} \tag{3.33}$$
$$E_{\mathcal{B}} - E_0 = \epsilon_c - \epsilon_a - J_{ac} + K_{ac} \tag{3.34}$$

and the equivalent expression for the pseudo system. Again, the ideal PP exactly reproduces these energy differences. Hence one gets for the orbital energy differences

$$\epsilon_b - \epsilon_a = \widetilde{\epsilon}_b - \widetilde{\epsilon}_a - \Delta_{ab} \tag{3.35}$$
$$\epsilon_c - \epsilon_a = \widetilde{\epsilon}_c - \widetilde{\epsilon}_a - \Delta_{ac} \tag{3.36}$$

where Δ_{ai} is defined as in (3.26). If the Δ_{ai} vanish, then the HF AE orbital energy *spectrum* is correctly reproduced in the pseudo system. The absolute orbital energies ϵ_i and $\widetilde{\epsilon}_i$ can, however, differ by a constant shift.

3.2.4. PP parametrization

The Roothaan–Hall SCF equations may be formulated to provide all occupied *and* virtual MOs. [94, Ch. 10.6] For the Koopmans four state problem in Ch. 3.2.2, the Fock matrix[1] is

$$\boldsymbol{F} = \begin{pmatrix} 2\epsilon_a & 0 & 0 \\ 0 & 2\epsilon_b & 0 \\ 0 & 0 & 2\epsilon_c \end{pmatrix} \tag{3.37}$$

As shown in the previous sections, the orbital energies $\widetilde{\epsilon}_i$ coincide with ϵ_i for an ideal PP.

$$\boldsymbol{F} = \widetilde{\boldsymbol{F}} = \begin{pmatrix} 2\widetilde{\epsilon}_a & 0 & 0 \\ 0 & 2\widetilde{\epsilon}_b & 0 \\ 0 & 0 & 2\widetilde{\epsilon}_c \end{pmatrix} \tag{3.38}$$

Note that both matrices $\boldsymbol{F}$ and $\widetilde{\boldsymbol{F}}$ are spanned in *different* sets of MOs. $\boldsymbol{F}$ contains the one- and two-electron terms, whereas $\widetilde{\boldsymbol{F}}$ contains additionally the PP term.

$$\widetilde{\boldsymbol{F}} = \widetilde{\boldsymbol{h}} + \widetilde{\boldsymbol{g}} + \widetilde{\boldsymbol{V}} \tag{3.39}$$

Thus, the parametrization of the PP in a given PO basis is known. It contains six parameters, that can be determined by bringing the one- and two-electron matrices of (3.39) to the left hand side (LHS).

$$\widetilde{\boldsymbol{V}} = \begin{pmatrix} \widetilde{V}_{11} & \widetilde{V}_{12} & \widetilde{V}_{13} \\ \widetilde{V}_{12}^* & \widetilde{V}_{22} & \widetilde{V}_{23} \\ \widetilde{V}_{13}^* & \widetilde{V}_{23}^* & \widetilde{V}_{33} \end{pmatrix} \tag{3.40}$$

$$\widetilde{\boldsymbol{V}} = \begin{pmatrix} 2\epsilon_a - \widetilde{h}_{aa} - \widetilde{g}_{aa} & -\widetilde{h}_{ab} - \widetilde{g}_{ab} & -\widetilde{h}_{ac} - \widetilde{g}_{ac} \\ & 2\epsilon_b - \widetilde{h}_{bb} - \widetilde{g}_{bb} & -\widetilde{h}_{bc} - \widetilde{g}_{bc} \\ & & 2\epsilon_c - \widetilde{h}_{cc} - \widetilde{g}_{cc} \end{pmatrix} \tag{3.41}$$

$$= \begin{pmatrix} 2(E_0 - E_{\mathcal{C}}) - \widetilde{h}_{aa} - \widetilde{g}_{aa} & -\widetilde{h}_{ab} - \widetilde{g}_{ab} & -\widetilde{h}_{ac} - \widetilde{g}_{ac} \\ & 2(E_0 - E_{\mathcal{A}1}) - \widetilde{h}_{bb} - \widetilde{g}_{bb} & -\widetilde{h}_{bc} - \widetilde{g}_{bc} \\ & & 2(E_0 - E_{\mathcal{A}2}) - \widetilde{h}_{cc} - \widetilde{g}_{cc} \end{pmatrix} \tag{3.42}$$

$\widetilde{\boldsymbol{V}}$ depends explicitly on the eigenvalues of the many-electron Hamiltonian but also through the one-electron and two-electron matrix integrals implicitly

[1] Here, the spatial orbital representation is used to avoid unnecessarily large matrices.

on the POs. However, any practical implementation of the PP relies on the parametrization of the operator $\hat{V}$ and not its matrix elements. An example was given in (2.38). These parameters have to be identified self-consistently. An implementation of such a PP fitting by Gaul [96] was available to me.

3.2.5. Comparison to PK Potentials

The discussion of the Koopmans' models and the CIS model showed that in all cases the fitting of the PP to the energy expectation values of the investigated states generates a PP for which the HF orbital energy spectrum is approximately energy-consistent. In contrast, the PK PP reproduces only one orbital energy exactly, whereas the whole spectrum is compressed. [76] Two properties of the PK PP may be at the root of this problem. The first is the single parameter which enters the PK PP regardless of system size. In contrast, an EC PP contains many more parameters. In the previously discussed case, six parameters would be needed for a proper fit. Secondly, the AE orbital-energies and the core orbitals are calculated once and for all and subsequently used to construct the PK PP. Generation of an EC PP requires an iterative procedure instead, as the PP and the POs depend on each other.

3.2.6. Fock-like Equations in MCSCF

Up to now I always assumed a HF ground state wave function, for which the Fock equation for orbital ϕ_i,

$$\hat{F}_{ii}\phi_i = \sum_j \epsilon_{ji}\phi_j \tag{3.43}$$

can be unitarily transformed into a pseudo eigenvalue equation

$$F_{ii}\phi'_i = \epsilon'_{ii}\phi'_i \qquad . \tag{3.44}$$

For open-shell HF or MCSCF wave functions, this is not generally possible anymore (see, e.g., Refs. [97] and [98, Ch. 8]). In MCSCF theory, one has a Fock operator for each pair of electrons, such that

$$\sum_j \hat{F}_{ij}\phi_j = \sum_j \epsilon_{ji}\phi_j \tag{3.45}$$

If it is possible to choose the ϕ_j to diagonalize the one-electron density matrix (natural orbital expansion), then the Fock-operators decouple to (3.43). For each

orbital ϕ_i there exists then an individual Fock operator. However, simultaneous diagonalization of $\boldsymbol{\epsilon}$ is not generally possible. [98]

Let me assume for the moment that such a natural orbital transformation indeed can be found. The elements of a generalized Fock matrix are given by (3.18) and its diagonal elements are

$$F_{mm} = \sum_q h_{mq} D_{mq} + \sum_{qrs} g_{mqrs} P_{mqrs} \quad . \tag{3.46}$$

The expression for the total energy

$$E = \sum_{pq} h_{pq} D_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} P_{pqrs} \tag{3.47}$$

can therefore be recast to incorporate the diagonal elements of the Fock matrix.

$$E = \sum_p F_{pp} - \frac{1}{2} \sum_{pqrs} g_{pqrs} P_{pqrs} = \sum_{\mathcal{I}} w_{\mathcal{I}} \left(\sum_p F^{\mathcal{I}}_{pp} - \frac{1}{2} \sum_{pqrs} g_{pqrs} P^{\mathcal{I}}_{pqrs} \right) \tag{3.48}$$

The weighted difference between two state energies $v_{\mathcal{I}} E_{\mathcal{I}} - v_{\mathcal{J}} E_{\mathcal{J}}$ is

$$v_{\mathcal{I}} E_{\mathcal{I}} - v_{\mathcal{J}} E_{\mathcal{J}} = \sum_p \left(v_{\mathcal{I}} F^{\mathcal{I}}_{pp} - v_{\mathcal{J}} F^{\mathcal{J}}_{pp} \right) - \frac{1}{2} \sum_{pqrs} g_{pqrs} \left(v_{\mathcal{I}} P^{\mathcal{I}}_{pqrs} - v_{\mathcal{J}} P^{\mathcal{J}}_{pqrs} \right) \tag{3.49}$$

Without approximation, the full expression cannot be reduced any further. Using a HF ground state as state $\mathcal{I}$, an anionic Koopmans state as state $\mathcal{J}$ and unit weights,

$$v_{\mathcal{I}} = v_{\mathcal{J}} = 1 \quad , \tag{3.50}$$

one reobtains the Koopmans example discussed in Sec. 3.2.1. The orbital-energy consistency, which was ensured by selection of an EC PP in the previous sections, is thus a boundary case of the more general MCSCF EC PP. In order to elaborate the consequences of EC PPs further, I will assume in the following, that EC implies orbital-energy consistency.

3.2.7. PO Shape

In this section, I discuss the shape of EC POs. As the core-tail is more or less determined by the PP that smooths the oscillations of the wave function close to the nucleus, I am interested in the behavior of the orbital at large distances from the nucleus.

In the appendix to Ref. [99], Topp and Hopfield show with reference to Kramers [100] that matching the ground and excited states for any energy-independent potential to the excitation energies of a real potential (e.g., an experimental spectrum) generates the correct logarithmic derivatives of one valence orbital w.r.t. the radial coordinate at and outside the core radius r_c.

$$-\frac{1}{2}r^2\phi_E^2\frac{\mathrm{d}}{\mathrm{d}E}\left(\frac{\partial\phi_E}{\partial r}\frac{1}{\phi_E}\right)\bigg|_{r=r_c} = \int\int\int_0^{r_c}\phi_E^*(\vec{x})\phi_E(\vec{x})\,\mathrm{d}\vec{x} \tag{3.51}$$

Eq. (3.51) defines the charge contained in the core up to a constant. The valence tail of the potential, i.e., far from the nucleus, is assumed to be of pure Coulomb type and the charge outside the core is defined up to a constant factor as well. The factor is determined by the total charge. The authors conclude that any potential that generates the correct ground and excited state energies will lead to the correct orbital shape in the spatial valence region.

PPs of the Stuttgart type are of Coulomb form: Dolg states in Ref. [26, Sec. 5.3.2] that the only considered values for n_{kl} in

$$v_l(r) = -\frac{Z-N_c}{r} + \sum_k A_{kl}r^{-n_{kl}}\mathrm{e}^{-\alpha_{kl}r^2} \tag{2.38}$$

are from the set $\{0,1,2\}$, but modern Stuttgart PPs use only the value $n_{kl} = 2$. Fig. 3.1 shows, that PPs of type (2.38) constitute effectively screened r^{-n} potentials and decay much faster. Within $2\,a_0$, the screened potentials have decayed, while the unscreened potentials are still non-negligible. A test charge further away from the nucleus would feel mainly the Coulomb potential due to the remaining charge $\frac{Z-N_c}{r}$. As this constitutes the situation for which the Topp–Hopfield–Kramers theorem was derived, I follow the conclusion of Dolg [26] and Schwerdtfeger [24] that an EC PO will follow the shape of a corresponding AE orbital, although no distinct core radius is determined.

3.2.8. AO Contributions to POs

Let me now turn to the contributions of AE AOs to POs, which I need in order to find a proper reconstruction method. The assumption is that if one knows which AE orbitals contribute to a PO, then one can reconstruct the AE orbitals by projecting out *unwanted* parts of the PO.

Dyall explains in Ref. [76], that shape-consistent POs cannot reproduce the same orbital energy spectrum as orbitals obtained in AE calculations. The PO

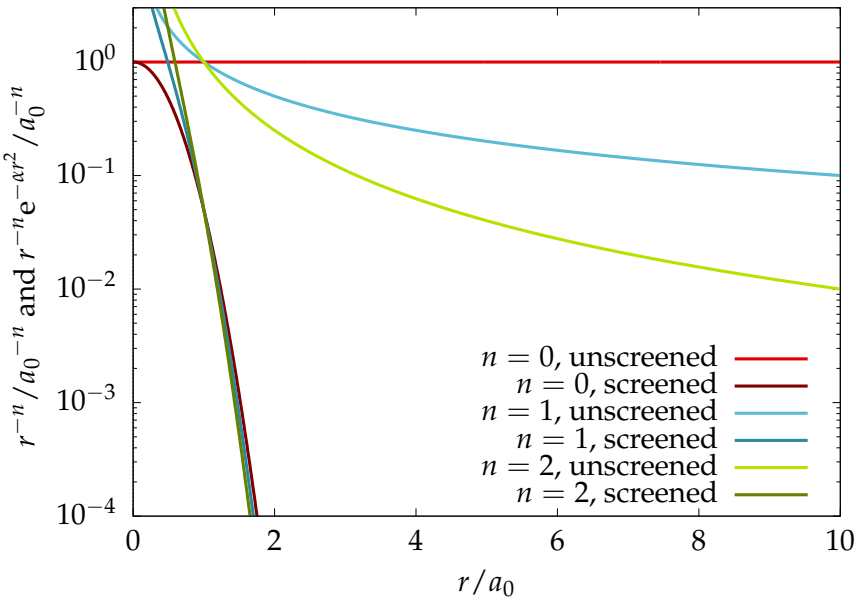


Figure 3.1.: Screening of r^{-n} potentials. The screening of potentials of the type r^{-n} due to an unnormalized Gaussian function $e^{-\alpha r^2}$ with exponent $\alpha = 3a_0^{-2}$ is shown. For $n = 1$, one obtains a (screened) Coulomb potential.

spectrum is compressed and the valence energy (energy of the valence electron determinant) is changed due to admixture of virtual orbitals in the POs. By projecting out the unoccupied orbitals, one can effectively achieve (valence) energy-consistency while sacrificing shape-consistency. Dyall speculates that energy-consistently fitted PPs allow a balance between a good representation of the valence energy, orbital spectrum and orbital shape. However, Schwerdtfeger mentions in Ref. [24] that EC PO are often practically SC, supporting the conclusion of Dyall.

I analyze in Tab. 3.1 the lowest UHF Ba PO (obtained with a Stuttgart PP), the 5s orbital. Shown are the contributions of HF AE AOs to the PO. The major contribution comes ($\approx$ 99.3%) from the Ba 4s and 5s orbitals. Small amounts of high energy virtual orbitals are needed to remove oscillations close to the

Table 3.1.: AE orbital contributions to Ba 5s PO. Only contributions with absolute value larger than 0.02 (core) and 0.03 (unoccupied orbitals) are shown.

	ϕ_{3s}	ϕ_{4s}	ϕ_{5s}	ϕ_{15s}	ϕ_{16s}
$\widetilde{\phi}_{5s}$	−0.022	−0.328	−0.941	0.052	−0.037

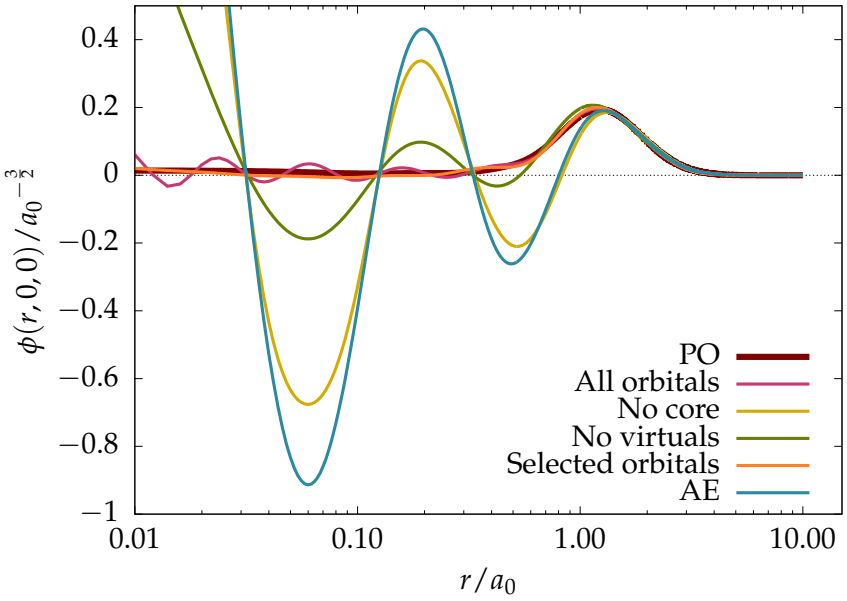


Figure 3.2.: Reexpansion of Ba 5s PO in AE AOs. The 5s PO is reexpanded using different AE orbital sets. The results are compared to the AE and PP solutions. A specific selection of orbitals removes the oscillations in the core completely, whereas a reconstruction using all orbitals introduces high frequency oscillations in the core region while maximizing the valence overlap. Ignoring the core-orbitals in the reexpansion generates almost a AE 5s orbital, indicating the presence of core orbitals in POs.

nucleus. In Fig. 3.2 I show the same 5s PO, reexpanded in different subsets of the AE AOs. The linear combination that approximates the shape of the PO (visually) the best is found using the contributions given in Tab. 3.1. Using significantly more than these two virtuals, introduces high frequency oscillations of the orbital close to the nucleus. The reason is that the contributions are determined by the overlap integral of AE AOs and the PO. In Fig. 3.3 the radial increase of the overlap integral

$$\int_0^r \phi_a(\vec{r}\,')\phi_b(\vec{r}\,')\,\mathrm{d}\vec{r}\,' \tag{3.52}$$

with $a, b \in \{\text{AE AO, PO}\}$ is shown. For all combinations of a and b the major contribution to the integral comes from the region $r > 1\,a_0$.

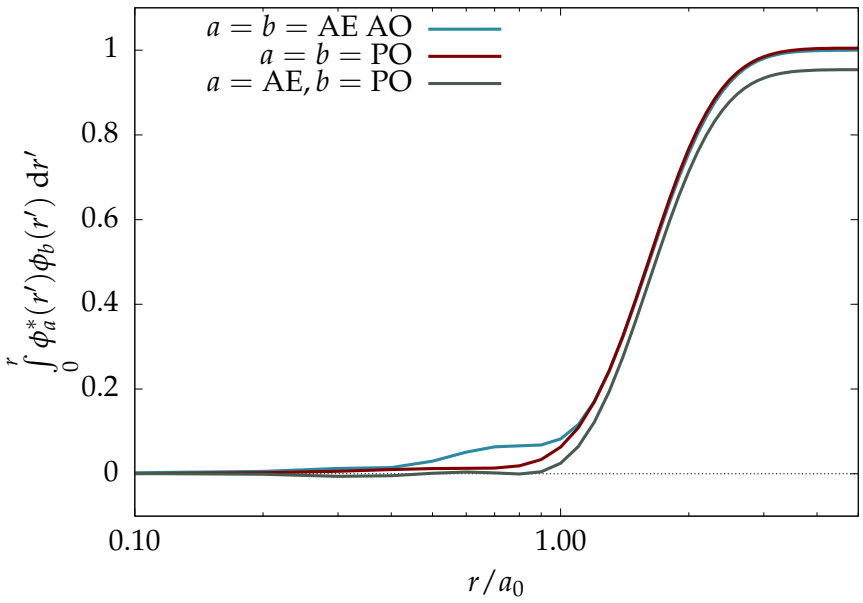


Figure 3.3.: Partial Ba 5s orbital overlap integrals. The overlap integral is calculated between orbitals ϕ_a and ϕ_b from 0 to r, where a and b were selected from the set $\{\mathrm{AE}, \mathrm{PO}\}$. The integral shows nearly identical behavior for the AE AO and PO. Only close to the core one sees an additional bump due to the nodal structure of the AE 5s orbital. Despite the difference in the absolute value, the overlap between AE AO and PO also follows the same shape and thus allows orbital reconstruction by using the overlap integrals over the full space, rather then overlap integrals over the valence space only.

3.2.9. Summary

I showed in this section that various models allow to formulate EC PPs which are also orbital-energy consistent. These approximations constitute boundary cases of more involved MCSCF adjusted EC PP. Hence I assumed that EC PP can be considered orbital-energy consistent. Topp and Hopfield have shown before that wave functions in a Coulomb potential, which have the same spectrum, are also shape-consistent in the spatial valence region. Indeed, my investigation of a PO showed that shape-consistency holds in a weak sense for EC PPs: Although one cannot formulate a specific core radius after which the PO follows the same shape as its AE counterpart, there is only little difference between the PO and the AE AO shape at large distance to the nucleus. From these results I concluded that the transformation of the atomic PO to AE AO can be based solely on the overlap metric (see Fig. 3.3). I.e., overlap integrals are the only quantities that determine the contribution of each AE AO to the final orbital. The nodal structure is reintroduced by reexpanding, e.g., a 5s PO in a finite set of AE AOs. As the

5s PO has the largest overlap with its 5s AE counterpart, this AO contributes the most in the reexpansion. If the finite set of AE AOs is not able to cancel the undulations of the 5s AO, then the reexpanded "PO" will have a shape very similar to that of the 5s AE AO. Including instead all AE AO in the reconstruction can lead to a full projection (i.e. (nearly) exact representation) of the PO into the new basis. Projection of the PO is avoided by carefully selecting the orbital space for the reconstruction, e.g., using only occupied and low-lying virtual orbitals. One should note that the incomplete reexpansion, which reintroduces the nodal structure of the orbital, necessarily generates an orbital with a norm smaller than unity. To provide a proper orbital, a renormalization is necessary, which, however, redistributes electrons in the system. This electron redistribution has much more impact on molecules than on atoms and is further discussed later.

3.3. Mathematical Grounds and Reconstruction Protocol

In this section I make the transition from atomic to molecular systems. I first provide the mathematical working equations, from which a stepwise protocol is created to guide through the process of orbital reconstruction. Secondly, I show in greater detail, which components are necessary for a successful reconstruction.

3.3.1. Mathematical Formulation of Reconstruction Procedure

The reconstruction of MOs is performed by a mapping of the PMOs $\widetilde{\psi}_i$, usually expanded in a basis $\{\widetilde{\chi}_\mu\}$

$$|\widetilde{\psi}_i\rangle = \sum_\mu |\widetilde{\chi}_\mu\rangle \, \widetilde{c}_{\mu i}, \tag{3.53}$$

onto a new basis. The new basis $\{\phi\}$ contains a selection of valence and virtual AE AOs of the pseudo atom. Additionally, all AE AOs of atoms without PP are included, but orthogonalized w.r.t. core, valence and virtual orbitals of the pseudo atom to obtain a basis that is least-distorted from the atomic orbital basis of the AE atom. Furthermore, linear dependencies can be avoided by neglecting basis functions of atoms without PP that are linear dependent to the core AOs.[2]

[2]For all calculations performed in this thesis, no basis functions were neglected. However, very diffuse basis functions in very small molecules could introduce linear dependencies that would be removed via this route.

A transformation into this orthogonal basis is described by the mapping operator $\hat{P}$.

$$\hat{P} = \sum_k |\phi_k\rangle\langle\phi_k| \tag{3.54}$$

Expanding ϕ_k in a finite basis set of Gaussians $\{\chi_\mu\}$,

$$|\phi_k\rangle = \sum_\mu |\chi_\mu\rangle\, m_{\mu k}, \tag{3.55}$$

where $m_{\mu k}$ are the expansion coefficients, one obtains for $\hat{P}$

$$\hat{P} = \sum_{k\mu\nu} |\chi_\mu\rangle\, m_{\mu k} (m_{k\nu})^* \,\langle\chi_\nu|\,. \tag{3.56}$$

The PMO is now mapped onto the new basis. As the new basis excludes the necessary orbitals to remove the nodal structure of the AO at the nucleus, one effectively reintroduces the nodal structure.

$$\begin{aligned}
|\psi_i\rangle &= \hat{P}\,|\widetilde{\psi}_i\rangle && (3.57)\\
&= \sum_{k\mu\nu\mu'} |\chi_\mu\rangle\, m_{\mu k} (m_{k\nu})^* \,\langle\chi_\nu|\widetilde{\chi}'_\mu\rangle\, \widetilde{c}_{\mu' i} && (3.58)\\
&= \sum_{k\mu\nu\mu'} |\chi_\mu\rangle\, m_{\mu k} (m_{k\nu})^* \,(\boldsymbol{S}_{BA})_{\nu\mu'}\, \widetilde{c}_{\mu' i} && (3.59)\\
&= \sum_{\mu} |\chi_\mu\rangle\, c_{\mu i} && (3.60)
\end{aligned}$$

$\boldsymbol{S}_{BA}$ is the overlap matrix between the Gaussian basis sets $\{\chi_\mu\}_B$ and $\{\widetilde{\chi}_\mu\}_A$. All reconstructed MOs can be obtained in a single step, using the matrix formulation of (3.59).

$$\boldsymbol{C}_B = \boldsymbol{M}\boldsymbol{M}^\dagger \boldsymbol{S}_{BA} \widetilde{\boldsymbol{C}} \tag{3.61}$$

As the POs were obtained within the FC approximation anyway and all components of $\hat{P}$ are chosen orthogonal to the atomic core orbitals, the ACOs simply complement the reconstructed MOs without further processing to provide a full set of MOs.

3.3.2. Stepwise Protocol

The first ingredients are the PMOs and an AE calculation for the atom where the PP was employed. The latter one can be precomputed and saved for the use in other reconstructions. A core space is defined by the core electrons contained in

the PP. Additionally, a valence space needs to be selected which will in general be larger than the occupied orbital space of the atom to correctly span the valence part of the PMOs. The orbitals which are to be reconstructed are then defined. Lastly, (3.61) is calculated to get the MO coefficients in terms of AE orbitals.

Before I state the protocol, let me introduce the vocabulary to help addressing the proper systems and bases:

Heavy atom: Atom, that carries a PP in the PP calculation

Light atom: Atom, that does not carry a PP in any calculation

$\{\widetilde{\chi}^{(\mathrm{P})}\}$: BFs at the heavy atom in a PP calculation.

$\{\chi^{(\mathrm{P})}\}$: BFs at the heavy atom in an AE calculation.

$\{\chi^{(\mathrm{R})}\}$: BFs at all light (*remaining*) atoms.

$\{\widetilde{\chi}\}$: The full basis for the system with PP. $\{\widetilde{\chi}\} = \{\widetilde{\chi}^{(\mathrm{P})}\} \cup \{\chi^{(\mathrm{R})}\}$

$\{\chi\}$: The full basis for the reconstructed AE system. $\{\chi\} = \{\chi^{(\mathrm{P})}\} \cup \{\chi^{(\mathrm{R})}\}$

$\{\widetilde{\psi}\}$: Subset of the PMOs of the molecule. $\widetilde{\psi}_i = \sum_\mu \widetilde{\chi}_\mu \widetilde{c}_{\mu i}$, i is the orbital index.

$\{\phi^{(\mathrm{P})}\}$: AE AOs of the heavy atom. $\phi_i^{(\mathrm{P})} = \sum_\mu \chi_\mu^{(\mathrm{P})} d_{\mu i}$, i indicates the orbital index.

$\{\phi^{(\mathrm{P})}\}_{\mathrm{core}}$: AE ACOs of the heavy atom. $\phi_i^{(\mathrm{P})} = \sum_\mu \chi_\mu^{(\mathrm{P})} d_{\mu i}, \qquad i = \{1, \ldots, n_{\mathrm{core}}\}$.

$\{\phi^{(\mathrm{P})}\}_{\mathrm{select}}$: Subset of AE valence and virtual AOs of the heavy atom. $\phi_i^{(\mathrm{P})} = \sum_\mu \chi_\mu^{(\mathrm{P})} d_{\mu i}, \qquad i = \{n_{\mathrm{core}}+1, \ldots, n_{\mathrm{core}}+n_{\mathrm{val}}, \ldots, n_{\mathrm{core}}+n_{\mathrm{val}}+n_{\mathrm{select}}\}$.

$\{\phi^{(\mathrm{R})}\}$: The orthogonalized basis at all light atoms.
$\{\phi^{(\mathrm{R})}\} = \{\chi^{(\mathrm{R})}\} - \left(\{\phi^{(\mathrm{P})}\}_{\mathrm{core}} \cup \{\phi^{(\mathrm{P})}\}_{\mathrm{select}}\right)$

$\{\phi\}$: $\{\phi\} = \{\phi^{(\mathrm{P})}\} \cup \{\phi^{(\mathrm{R})}\}$

The protocol goes as follows:

1. Calculate PMOs $\{\widetilde{\psi}\}$ for the molecule in the Gaussian basis $\{\widetilde{\chi}\}$.
2. Calculate AE AOs $\{\phi^{(P)}\}$ for the heavy atom.
3. From the Gaussian BFs $\{\chi^{(\mathrm{P})}\}$ used in step 2 and the BFs of all light atoms $\{\chi^{(\mathrm{R})}\}$ used in step 1, construct the combined basis $\{\chi\}$.

4. Identify the core orbitals $\{\phi^{(P)}\}_{\text{core}}$ from the orbitals of an AE heavy atom calculation $\{\phi^{(P)}\}$ (i.e., the orbitals contained in the PP) and select the valence (and virtual) orbitals $\{\phi^{(P)}\}_{\text{select}}$ that can describe the valence part of the PMO. An analysis as performed in Sec. 3.2.8 is helpful.
5. Orthonormalize the basis functions $\{\chi^{(R)}\}$ w.r.t. the atomic orbitals $\{\phi^{(P)}\}_{\text{core}} \cup \{\phi^{(P)}\}_{\text{select}}$ to obtain the basis $\{\phi^{(R)}\}$.
6. Build the new valence basis from the union $\{\phi^{(P)}\}_{\text{select}} \cup \{\phi^{(R)}\}$. Its matrix representation is named $\boldsymbol{M}_{\perp,\text{val}}$. The corresponding matrix representation of the union $\{\phi^{(P)}\}_{\text{core}} \cup \{\phi^{(P)}\}_{\text{select}} \cup \{\phi^{(R)}\}$ is named $\boldsymbol{M}_{\perp}$.
7. Calculate the overlap matrix $\boldsymbol{S}_{\widetilde{\chi}\chi}$ of bases $\{\widetilde{\chi}\}$ and $\{\chi\}$. Note that $\boldsymbol{S}_{\widetilde{\chi}\chi} = \boldsymbol{S}_{\chi\widetilde{\chi}}^{T}$.
8. Select the PMO coefficients $\widetilde{\boldsymbol{C}}_{\text{select}}$ that describe the orbitals which shall be reconstructed.
9. Transform these coefficients into the basis $\{\chi\}_B$.

$$\boldsymbol{C}_{B,\text{val}} = \boldsymbol{M}_{\perp\text{val}} \boldsymbol{M}_{\perp\text{val}}^{\dagger} \boldsymbol{S}_{BA} \widetilde{\boldsymbol{C}}_{\text{select}} \tag{3.62}$$

10. Renormalize $\boldsymbol{C}_{B,\text{val}}$ to give $\boldsymbol{C}_{B,\text{val,norm}}$.
11. Augment $\boldsymbol{C}_{B,\text{val,norm}}$ by the core orbitals in $\boldsymbol{M}_{\perp}$, which gives $\boldsymbol{C}_{B,\text{norm}}$.
12. Use $\boldsymbol{C}_{B,\text{norm}}$ for the calculation of core-like properties.

Sec. 3.6 investigates the applicability and peculiarities of this protocol.

3.4. Numerical Example

Let me discuss an illustrative example: The molecule under consideration is LiH, where the inner 1*s* shell of Li is condensed in a PP. Minimal bases (Tab. 3.2) are used to construct the AOs of Li and H. In the following, I will need two basis definitions: The letter A describes the set $\{\chi_{\text{Li}-\text{PP}}, \chi_{\text{H}}\}$ and B identifies the set $\{\chi_{\text{Li}}, \chi_{\text{H}}\}$. Both sets contain the corresponding BFs given in Tab. 3.2.

A PP was created for the Li-PP basis in Tab. 3.2 that reproduces the Li 2s HF orbital energy of $\varepsilon_{2\text{s}} = -0.180124$, using the program PPFIT implemented by Konstantin Gaul during an internship. [96] The PP parameters are given in Tab. 3.3. These calculations are simple enough to constitute an easily recalculable, despite (or because of) their methodical crudeness.

Table 3.2.: Minimal basis for LiH. s-type functions of STO-3G basis for Li and H of Hehre, Stewart and Pople [101] were taken from the Basis Set Exchange Library [102].

Atom	Type	Exponents	Contraction coefficients
Li	s	16.119 575	0.154 329
		2.936 201	0.535 328
		0.794 650	0.444 635
	s	0.636 290	−0.099 967
		0.147 860	0.399 513
		0.048 089	0.700 115
Li-PP	s	0.636 290	−0.099 967
		0.147 860	0.399 513
		0.048 089	0.700 115
H	s	0.168 856	0.444 635
		0.623 913	0.535 328
		3.425 250	0.154 329

Table 3.3.: Minimal basis PP for Li. The PP consists of one s-type function that was fitted to the 2s orbital energy of Li.

Angular momentum	Exponential coefficient	Coefficients	n
s	3.000000	19.773 661	2
p	1.000000	0.000 000	2

Calculation of PMOs

The LiH PMOs are spanned in basis A with the atoms separated by $d \approx 2.885\, a_0$. (The exact value is printed in App. D.)

$$\widetilde{\psi}_1(\vec{x}) = c_{11}\chi_{\mathrm{Li}-PP} + c_{21}\chi_{\mathrm{H}} \tag{3.63}$$

$$\widetilde{\psi}_2(\vec{x}) = c_{12}\chi_{\mathrm{Li}-PP} + c_{22}\chi_{\mathrm{H}} \tag{3.64}$$

The coefficient matrices[3] for the α and β electron are[4]

$$\widetilde{\boldsymbol{C}}^T(\alpha) = \begin{pmatrix} 0.524891 & 0.661317 \\ 0.964820 & -0.876953 \end{pmatrix} \tag{3.65}$$

$$\widetilde{\boldsymbol{C}}^T(\beta) = \begin{pmatrix} 0.524891 & 0.661317 \\ 0.964820 & -0.876953 \end{pmatrix} \tag{3.66}$$

and the PMOs are orthogonal in the overlap metric of the corresponding basis.

$$\mathbf{1} = \widetilde{\boldsymbol{C}}^T \boldsymbol{S}_{AA} \widetilde{\boldsymbol{C}}. \tag{3.67}$$

A general overlap matrix $\boldsymbol{S}_{AB}$ is defined by the inner products of all BFs of the bases A and B.

$$(\boldsymbol{S}_{AB})_{ij} = \langle \chi_{A,i} | \chi_{B,j} \rangle \tag{3.68}$$

For two primitive, normalized Gaussian basis functions $\bar{\chi}_i$ and $\bar{\chi}_j$ with orbital angular momentum quantum number $l = 0$, the integral can be evaluated analytically.

$$\langle \bar{\chi}_i | \bar{\chi}_j \rangle = s_{x,ij} s_{y,ij} s_{z,ij} = s_{ij} \tag{3.69}$$

$$s_{a,ij} = \left(\frac{2\sqrt{(\alpha_i \alpha_j)}}{\alpha_i + \alpha_j} \right)^{\frac{1}{2}} \mathrm{e}^{\frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j} (a_j - a_i)^2} \tag{3.70}$$

a_i is the ath component of the center of the primitive basis function i. α_i is its exponent. Each element of the overlap matrix is a sum over the overlap integrals of primitive basis functions, multiplied by the corresponding contraction coefficients d_{ki}.

$$S_{ij} = \sum_{kl} d^*_{ki} d_{lj} s_{kl} \tag{3.71}$$

[3] Orbitals are stored in $\boldsymbol{C}$ in column wise order. I.e., $\boldsymbol{C}^T$ contains the orbitals in row wise order, which is the preferred notation here in order to mirror the Mathematica implementation in App. D.

[4] The PMOs presented here are eigenfunctions to the NR HF Hamiltonian and were obtained using the program TURBOMOLE.

The basis indices AB were dropped here for simplicity. $\boldsymbol{S}_{AA}$ is

$$\boldsymbol{S}_{AA} = \begin{pmatrix} 1.000000 & 0.413617 \\ 0.413617 & 1.000000 \end{pmatrix}. \tag{3.72}$$

Calculation of AE AOs and Building the Intermediate AO Basis

In order to exploit similarity of Li AE AOs and PAOs, the next step is to construct a new basis for LiH. The basis consists of the AE AOs of Li and the basis functions of H. Two coefficient matrices describe the Li AOs.

$$\begin{aligned} \boldsymbol{C}_{\mathrm{Li}}^{T}(\alpha) &= \begin{pmatrix} 0.991884 & 0.031695 \\ -0.279110 & 1.029918 \end{pmatrix} \\ \boldsymbol{C}_{\mathrm{Li}}^{T}(\beta) &= \begin{pmatrix} 0.991884 & 0.031695 \\ -0.279110 & 1.029918 \end{pmatrix} \end{aligned} \tag{3.73}$$

Li has an open 2s shell. In order to avoid broken spin-symmetry of the Li orbitals, which would later spin-polarize the reconstructed orbitals, the orbitals were spin-averaged by occupying each spin-orbital equally (here: $n_{2\mathrm{s}}(\alpha) = n_{2\mathrm{s}}(\beta) = 0.5$).[5] Thus, due to symmetry, the example is only discussed for the reconstruction of the α electron.

The combined basis of Li AOs and the H BF is created by adding a third *orbital* to $\boldsymbol{C}_{\mathrm{Li}}$. The third orbital has a unit coefficient at the position of the normalized hydrogen basis function. E.g., for α spin one has

$$\boldsymbol{M}^{T}(\alpha) = \begin{pmatrix} & & 0 \\ \boldsymbol{C}_{\mathrm{Li}}^{T}(\alpha) & & 0 \\ & & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.74}$$

This matrix, however, is not orthonormal in the overlap metric of the basis B, i.e.,

$$\mathbf{1} \neq \boldsymbol{M}^{T}\boldsymbol{S}_{BB}\boldsymbol{M}. \tag{3.75}$$

In order to fulfill the orthonormality condition one needs to orthogonalize the last row of $\boldsymbol{M}^{T}$ by the Gram–Schmidt (GS) procedure in the metric of the basis

[5] As I show in Ch. 3.6.5.2, using restricted-open-shell HF orbitals works just as well.

functions (see, e.g., Ref. [103, Ch. 5.2.8]).

$$\begin{aligned} |\chi_i'^{\mathrm{H}}\rangle &= |\chi_i^{\mathrm{H}}\rangle - \sum_j |\phi_j^{\mathrm{Li}}\rangle \langle\phi_j^{\mathrm{Li}}|\chi_i^{\mathrm{H}}\rangle \\ &= |\chi_i^{\mathrm{H}}\rangle - \sum_{klj} |\phi_k\rangle \langle\chi_k|\phi_j^{\mathrm{Li}}\rangle \langle\phi_j^{\mathrm{Li}}|\chi_l\rangle \langle\chi_l|\chi_i^{\mathrm{H}}\rangle \\ &= |\chi_i^{\mathrm{H}}\rangle - \sum_{klj} |\phi_k\rangle C_{kj} C_{lj}^* (\boldsymbol{S}_{BB})_{li} \end{aligned} \tag{3.76}$$

Note that the result of the orthogonalization does not depend on the ordering of the vectors in $\boldsymbol{C}_{\mathrm{Li}}$. The overlap matrix $\boldsymbol{S}_{BB}$ of the example is

$$\boldsymbol{S}_{BB} = \begin{pmatrix} 1.000000 & 0.241137 & 0.077754 \\ 0.241137 & 1.000000 & 0.413617 \\ 0.077754 & 0.413617 & 1.000000 \end{pmatrix}, \tag{3.77}$$

which gives the orthogonalized matrix $\boldsymbol{M}_\perp$

$$\boldsymbol{M}_\perp^T = \begin{pmatrix} 0.991884 & 0.031695 & 0.000000 \\ -0.279110 & 1.029918 & 0.000000 \\ 0.025645 & -0.460624 & 1.098697 \end{pmatrix} \tag{3.78}$$

$\boldsymbol{M}_\perp$ fulfills the condition

$$\mathbf{1} = \boldsymbol{M}_\perp^T \boldsymbol{S}_{BB} \boldsymbol{M}_\perp. \tag{3.79}$$

$\boldsymbol{M}_\perp$ is divided in two rectangular parts, the core matrix $\boldsymbol{M}_{\perp\mathrm{core}}$, which is defined by the core orbitals which are contained in the PP and the valence matrix $\boldsymbol{M}_{\perp\mathrm{val}}$.

$$\boldsymbol{M}_\perp^T = \begin{pmatrix} \boldsymbol{M}_{\perp\mathrm{core}} \\ \boldsymbol{M}_{\perp\mathrm{val}} \end{pmatrix} \tag{3.80}$$

$$\boldsymbol{M}_{\perp\mathrm{core}}^T = \begin{pmatrix} 0.991884 & 0.031695 & 0.000000 \end{pmatrix} \tag{3.81}$$

$$\boldsymbol{M}_{\perp\mathrm{val}}^T = \begin{pmatrix} -0.279110 & 1.029918 & 0.000000 \\ 0.025645 & -0.460624 & 1.098697 \end{pmatrix} \tag{3.82}$$

In the following, I will call the bases orthogonalized atomic core orbital (OACO) and orthogonalized atomic valence orbital (OAVO).

Valence PMO Selection and Reconstruction of AE MOs

A change-of-basis matrix is used to calculate the coefficients of a vector in the basis B, originally represented in basis A. [104, Ch. 6.3] The change-of-basis

matrix is, in this example, given by the matrix $\boldsymbol{S}_{AB}$, which is defined by the inner products of the basis functions of A and B.

$$\boldsymbol{S}_{AB} = \begin{pmatrix} 0.241137 & 1.000000 & 0.413617 \\ 0.077754 & 0.413617 & 1.000000 \end{pmatrix} \tag{3.83}$$

The change-of-basis matrix for the transformation into the OAVO basis is obtained by multiplication of $\boldsymbol{S}_{AB}$ with $\boldsymbol{M}_{\perp\mathrm{val}}$.

$$\boldsymbol{S}_{A\mathcal{B}} = \boldsymbol{S}_{AB}\boldsymbol{M}_{\perp\mathrm{val}} \tag{3.84}$$

$\mathcal{B}$ denotes the basis spanned by the OAVO. Multiplication from the left with the PMO coefficients $\vec{\widetilde{C}}_1(\alpha)^T$ and with $\boldsymbol{M}_{\perp\mathrm{val}}^T$ from the right gives the PMO coefficients in the new basis B. If one reconstructs only the occupied PMOs, one obtains

$$\begin{aligned} \boldsymbol{C}_{B,\mathrm{val}}^T(\alpha) &= \vec{\widetilde{C}}_1^T(\alpha)\boldsymbol{S}_{AB}\boldsymbol{M}_{\perp\mathrm{val}}\boldsymbol{M}_{\perp\mathrm{val}}^T \\ &= \begin{pmatrix} -0.200214 & 0.518494 & 0.661317 \end{pmatrix} \end{aligned} \tag{3.85}$$

$\vec{\widetilde{C}}_1(\alpha)$ is the first row of $\widetilde{\boldsymbol{C}}(\alpha)$ in (3.65), i.e., the lowest energy (2s) eigenvector of the molecule with PP.

The new MO coefficients provide orbitals that are *orthogonal* to the core orbitals of Li, but not necessarily normalized: The reason is that the orthogonalization of the H basis function can remove components in H that overlap with the Li core. These components are, however, present in the PMOs. The change-of-basis matrix describes in this case a *mapping*, i.e., the spanned spaces are not identical. The mapped vectors are shorter than the original vectors, i.e., not normalized. By renormalization, here via Löwdin symmetrical orthogonalization[6] the normalized orbital coefficients are obtained.

$$\boldsymbol{C}_{B,\mathrm{val,norm}}^T(\alpha) = \begin{pmatrix} -0.204421 & 0.529391 & 0.675215 \end{pmatrix} \tag{3.86}$$

[6]Strictly speaking, the renormalization does not need the orthogonalization part: Orthogonal orbitals were mapped to an orthonormal basis, so only the length of the resulting orbitals can change but not the angle between them. Instead, one could calculate the norm of each reconstructed orbital and scale each orbital by the inverse of the norm. However, as this work aims only for a reference implementation I simply used the already implemented orthogonalization schemes for renormalization. Besides symmetrical orthogonalization, I could have chosen also other symmetrization schemes. For the GS method I found almost no numerical differences compared to the symmetric orthogonalization, although the result of the GS algorithm depends on the ordering of the initial vectors. I concluded that the obtained vectors are already nearly orthogonal after reconstruction and any orthonormalization scheme will generate here almost identical orbitals. For a more complete discussion of different orthogonalization schemes in the context of core-valence orthogonalization, I refer here to the works of Cook [105] and Baerends *et al.* [106]

Table 3.4.: Molecular properties of LiH. The table compares values obtained after reconstruction with values obtained from a set of reference MOs calculated in the same basis.

Property	Reconstructed MOs	Reference MOs	$\varepsilon_{\mathrm{rel}}/\%$
EDM$/e\,a_0$	3.949	3.836	2.93
Kinetic energy$/E_\mathrm{h}$	8.170	8.155	1.93×10^{-1}
Mass velocity$/E_\mathrm{h}$	-0.002	-0.002	7.75×10^{-2}
$\left\langle -\frac{Z}{r} \right\rangle / E_\mathrm{h}$	-20.613	-20.590	1.13×10^{-1}
Contact density at Li$/{a_0}^{-3}$	8.204	8.215	1.32×10^{-1}
1e$^-$-Darwin corr.$/E_\mathrm{h}$	0.002	0.002	6.38×10^{-2}

Recombination with the OACO provides the full set of new MO coefficients.

$$\boldsymbol{C}^T_{B,\mathrm{norm}}(\alpha) = \begin{pmatrix} \boldsymbol{M}^T_{\perp\mathrm{core}} \\ \boldsymbol{C}^T_{B,\mathrm{val,norm}}(\alpha) \end{pmatrix} \tag{3.87}$$

$$= \begin{pmatrix} 0.991884 & 0.031695 & 0.000000 \\ -0.204415 & 0.529393 & 0.675219 \end{pmatrix} \tag{3.88}$$

The reconstructed orbitals are in good agreement with converged[7] AE orbitals.

$$\boldsymbol{C}^T_{\mathrm{converged}}(\alpha) = \begin{pmatrix} 0.991671 & 0.031910 & 0.001655 \\ -0.209620 & 0.547965 & 0.658809 \end{pmatrix} \tag{3.89}$$

A full working example is programmed in a MATHEMATICA document `example_protocol.m` provided in App. D and in the supplemental material of this thesis.

Property Calculation with Reconstructed MOs

After obtaining the coefficients in the new basis, molecular properties can be calculated. Tab. 3.4 compares properties calculated with reconstructed orbitals to those calculated with HF orbitals. Only small deviations are observable.

3.5. Analysis of Reconstructed Orbitals

The choice of reconstruction used in this work has some implications on the distribution of electrons within the molecule. In the following section I will

[7] The energy change in subsequent SCF cycles are below $10^{-9}\,E_\mathrm{h}$.

discuss the effects under the assumption that an AE orbital and a PO overlap exactly outside a core radius r_{core}.

In general, the POs $\widetilde{\psi}_i$ are not *exactly* representable in the new basis, as the occupied PAOs contain contributions from core- and virtual orbitals that have not been included in the reconstruction. Hence the norm of the reconstructed orbitals ψ_i is smaller than one.

$$\begin{aligned} \langle\widetilde{\psi}_i|\widetilde{\psi}_i\rangle &= 1 \\ \langle\psi_i|\psi_i\rangle &< 1 \end{aligned} \tag{3.90}$$

Subsequent renormalization of the MOs is thus necessary, which however amounts to a redistribution of electron density in the molecule. Referring to the numerical example, similarity of the Li PAOs and AOs in the valence region indicates that the major source of MO differences comes from the core region at the Li nucleus.

$$\langle\widetilde{\psi}_i|\widetilde{\psi}_i\rangle = \langle\widetilde{\psi}_i|\widetilde{\psi}_i\rangle_0^{r_{\text{core}}} + \langle\widetilde{\psi}_i|\widetilde{\psi}_i\rangle_{r_{\text{core}}}^{\infty} \tag{3.91}$$

$$\langle\psi_i|\psi_i\rangle = \langle\psi_i|\psi_i\rangle_0^{r_{\text{core}}} + \langle\psi_i|\psi_i\rangle_{r_{\text{core}}}^{\infty} \tag{3.92}$$

$$\langle\psi_i|\psi_i\rangle_{r_{\text{core}}}^{\infty} \approx \langle\widetilde{\psi}_i|\widetilde{\psi}_i\rangle_{r_{\text{core}}}^{\infty} \tag{3.93}$$

$$\implies \langle\psi_i|\psi_i\rangle_0^{r_{\text{core}}} < \langle\widetilde{\psi}_i|\widetilde{\psi}_i\rangle_0^{r_{\text{core}}} \tag{3.94}$$

(Sub- and Superscripts on the brackets denote the radial integration limits, where the origin has been set to the Li atom and r_{core} is a radius around the Li atom.) Renormalization will thus shift electron density from the outer part of the wave function to the inner part. The LiH bond of the example becomes more ionic and, as a consequence, the dipole moment increases. This effect is observed for most examples discussed in Ch. 3.6.

A few further points should be noted here: First, the procedure relies on the fact that the PAOs and AO do overlap. If the SC assumption fails, then the scheme is not applicable anymore. However, to my experience, the greater risk is that the PAO basis is not representable in the AE AO basis, which degrades the quality of reconstructed orbitals (see Ch. 3.6.3.2).

Second, the unoccupied AOs used to construct the new basis $\boldsymbol{M}$, have to be selected with care. If one chooses an insufficient number of unoccupied orbitals, then one cannot properly account for bond the polarization in the molecule and the MOs will rather resemble linear combinations of *occupied* atomic orbitals. The same holds, if the selected unoccupied orbitals are orthogonal to the PMOs. In Ch. 3.6.3.3 I investigate the influence of the number of included unoccupied orbitals.

Third, at the same time one should make sure that the selected virtual orbitals do not reconstruct the (reduced) nodal structure of the PMOs (see Ch. 3.2.8). The success of the reconstruction method thus relies somewhat on the error cancellation between an incomplete PMO expansion in the core part of the orbitals and a proper expansion of their valence part.

Fourth, I did not discuss the picture-change-error: The modern EC PP of the Stuttgart group are developed using a Dirac–Hartree–Fock–Coulomb–Breit (DHF-C/B) Hamiltonian. I.e., the effective Hamiltonian used to obtain POs contains an effective DHF-C/B potential (or better: The effective Hamiltonian contains a spin-averaged version of the effective DHF-C/B potential, as all POs in this work were obtained using a NR one-component (1c) implementation of TURBOMOLE). In all reconstructions presented herein, the AE AOs were obtained using a 1c ZORA implementation of a modified TURBOMOLE, i.e., I used de facto a SR Hamiltonian. However, the errors introduced through the straightforward reconstruction scheme are expected to be larger than the errors due to the picture change. Furthermore, the picture change is likely to be most recognizable very close to the core — exactly the region which is anyway badly described by POs. Thus I expect further that the picture, in which the AE AOs are obtained, has the largest effect on the reconstructed orbitals.

3.6. Numerical Tests

In this chapter, I show the performance of my reconstruction method on some examples. Atomic, diatomic and polyatomic systems are discussed. The influence of reconstruction parameters on the orbitals is studied and provides estimates for the implementation of an automated AE orbital reconstruction. All reconstruction parameters are tabulated in App. B.

3.6.1. Expectation Values as Quality Probes

The protocol derived in this work shall be applied to the calculation of molecular properties that have a strong contribution from the wave function close to nuclei. Hence I define a set of test properties that is used as marker for the quality of the scheme. Schwerdtfeger [24] defines core properties as

$$r^{-n}\frac{\partial^m}{\partial r^m}, \qquad m+n>0 \tag{3.95}$$

The test properties in this work are

- the kinetic energy operator $\hat{p}^2$, $\quad m = 2, n = 0$
- the mass-velocity correction $\hat{p}^4$, $\quad m = 4, n = 0$
- the one-electron Darwin term[8] $4\pi Z\delta(r)$, $\quad m = 0, n \approx 2$
- the Coulomb attraction $\frac{Z}{r}$, $\quad m = 0, n = 1$

Additionally, I chose the electric dipole moment (EDM) as a property to test the deformation of the electronic wave function in the valence space. Lastly, the principal axes of the hyperfine coupling tensor, $A_{\parallel}$ and $A_{\perp}$, as well as the electronic structure parameter W_a of the parity-violating hyperfine coupling, were chosen as test properties. The hyperfine coupling tensor parameters $A_{\parallel}$ and $A_{\perp}$ can be specified for any molecule with a symmetry axis C_3 or higher, if a principal component of the tensor is chosen along the molecular axis. [108, p. 45] They are connected to the *isotropic* and the *dipole* term of the tensor by

$$\begin{aligned} A_\mathrm{iso} &= \frac{A_{\parallel} + 2A_{\perp}}{3} \\ A_\mathrm{dip} &= \frac{A_{\parallel} - A_{\perp}}{3} \end{aligned} \tag{3.96}$$

The isotropic component is directly linked to the Fermi-contact interaction

$$A_\mathrm{iso} = \frac{8\pi}{3} \frac{\mu_\mathrm{e}}{S_\mathrm{e}} \frac{\mu_N}{I_N} \rho(\vec{r}_N) \tag{3.97}$$

and thus sensitive to the electron density $\rho(\vec{r}_N)$ at the position of nucleus N. The nuclear magnetic moment μ_N and the nuclear spin quantum number I_N are state-dependent properties of the nucleus N. μ_e and S_e are the electron magnetic moment and the electron spin quantum number. The dipole component of the hyperfine interaction tensor is calculated by

$$A_\mathrm{dip} = \frac{\mu_\mathrm{e}}{S_\mathrm{e}} \frac{\mu_N}{I_N} \left\langle \frac{3\cos^2(\theta) - 1}{2r^3} \right\rangle \tag{3.98}$$

where θ is the angle between the instantaneous position of the electron and the molecular axis. A_dip is sensitive to the rotationally averaged expectation value of the $\frac{1}{\hat{r}^3}$ operator. Both components are thus core-like properties. The reader is referred to Ref. [108, pp. 54 f.] for further information about the hyperfine coupling tensor.

Usually, when using NR methods, nuclei are well approximated as point-like charges. But a finite solution to the Dirac equation requires nuclei to be treated

[8] According to Ref. [107, p. 77], the Dirac delta function can be approximated by $\delta_n(r) = \frac{n}{\pi} \frac{1}{1+n^2r^2}$.

as finite charge distributions. A multipole expansion of the nucleus' electrostatic and magnetostatic fields gives rise to corresponding interactions between electron and nucleus. Zel'dovich [109] showed in 1957 that parity non-conservation, predicted by Lee and Yang [110], is caused in nuclei by an additional contact term to the magnetic quadrupole moment, the so called *anapole* moment k_a. [111] The interaction between electrons and nucleus inside the nucleus violates parity through the nuclear spin-independent weak interaction mediated through exchange of Z^0 bosons and through the nuclear spin-dependent anapole interaction. In linear open-shell molecules, the Z^0 boson exchange is suppressed by the electronic structure [112] and parity violation comes solely from the anapole interaction. In the effective spin-rotation-Hamiltonian of a linear molecule, the contribution of the nuclear spin-dependent parity-odd interaction to the splitting of hyperfine levels is parametrized by [112, 113]

$$W_\mathrm{a} k_{\mathcal{A},N} \left(\vec{\lambda} \times \vec{S}^\mathrm{eff} \right) \cdot \vec{I}. \tag{3.99}$$

W_a is approximately determined by

$$W_\mathrm{a} = \frac{1}{k_{\mathcal{A},N} \left[\vec{\lambda} \times \vec{S}^\mathrm{eff} \right]_{x,y}} \langle \Psi | \frac{\partial z_\mathrm{sd}^{(1,1)}}{\partial \vec{I}} | -\Psi \rangle \tag{3.100}$$

Ψ is an electronic state, $-\Psi$ the corresponding state of opposite parity and time (Kramers partner), $\vec{\lambda}$ the unit vector along the molecular axis pointing in the direction of the lighter nucleus, $\vec{S}^\mathrm{eff}$ the effective electronic spin, I the spin of the nucleus and $z_{sd}^{(1,1)}$ the nuclear spin-dependent parity violating term.

$$z_{sd}^{(1,1)}(i) = \sum_A \frac{G_\mathrm{F}}{\sqrt{2}} k_{\mathcal{A},N} \left[\vec{\sigma}_i \cdot \vec{p}_i, \frac{\tilde{\omega}}{c} \vec{\sigma}_i \cdot \vec{I}_A \rho_A(\vec{r}_i) \right]_+, \tag{3.101}$$

where i denotes the electron i, $\vec{\sigma}_i$ are the Pauli matrices, $\vec{p}_i$ is the linear momentum, $\vec{I}_A$ the nuclear spin of nucleus A, $\tilde{\omega} = 1/\left(2m_\mathrm{e} + \tilde{V}/c^2\right)$ the ZORA factor including the core model potential [114, 115], c the speed of light and G_F Fermi's constant. $k_{\mathcal{A},N}$ is the anapole moment of nucleus N and ρ_N the charge density of nucleus N. As the wave function contributes to W_a only in the spatial regions where the nuclear charge density is not vanishing, a proper description of W_a enforces a good description of the electronic wave function close to the nucleus.

3.6.2. Atoms

3.6.2.1. Noble gases

As a first test set I chose the noble gases Kr, Xe, Rn and Og. Their closed shells and the thus spherical symmetry should allow for a facile mapping of the PP orbitals onto their corresponding AE counterparts. In Fig. 3.4 the relative errors for the kinetic energy, the mass-velocity (MV) term, the electron-nucleus Coulomb interaction, the contact density $|\Psi(0)|^2$ and the Darwin term are shown. The EDM is not analyzed since the EDM of a closed-shell atom vanishes anyway and numerical imprecision of the calculation mounts up to (meaningless) relative errors a few magnitudes larger than the other quantities. The relative error ε is defined as

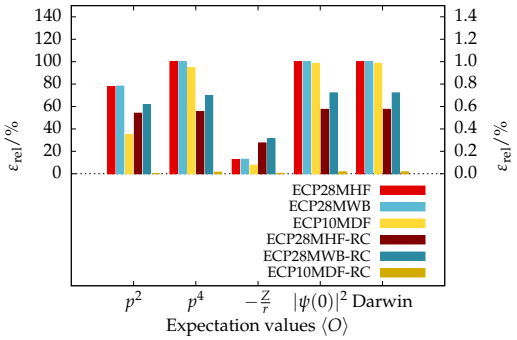
$$\varepsilon_{\text{rel}} = \left| \frac{x - x_0}{x_0} \right| \cdot 100\,\% \tag{3.102}$$

where x is a calculated value and x_0 is the reference value. Shown are the errors for the PP calculations w.r.t. an AE calculation as well as the errors after reconstruction of the wave function. The calculation of the expectation values included only those orbitals that are present in the PP system and not the full HF determinant.

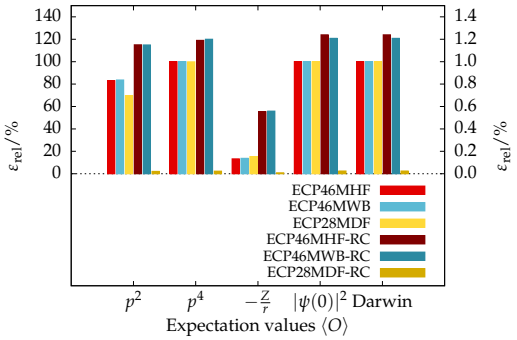
Fig. 3.4 shows that the values of core-like properties differ greatly from the reference when using PAOs. On the other hand, properties as, e.g., the EDM can be obtained rather accurately. The reconstruction, however, introduces the correct nodal structure into the core region of the valence orbitals (see Fig. 3.5). Hence the expectation values are in very good agreement with the AE results. Slight differences can be noted between the different PPs. These are likely to stem from core size effects: A small core PP is likely to give better results because the core shells are more atom-like. The outer core shells in large core PPs are more likely affected by the valence orbitals. For Rn, e.g., the small core PP gives much better results than the large core PPs. Picture change errors may also play a role here.

3.6.3. BaF

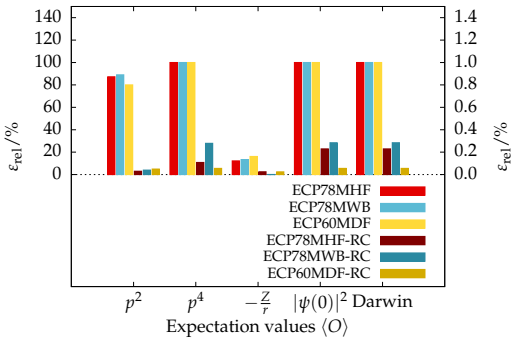
In the following subsection 3.6.3.1 I investigate the different PPs' influence on the quality of reconstructed AE wave functions for the diatomic molecule BaF. Then, in Sec. 3.6.3.2 I examine AE atomic basis sets for Ba w.r.t. the representability of the PMO. Sec. 3.6.3.3 shows how the choice of virtual orbitals in the reconstruction influences its quality whereas Sec. 3.6.3.4 explores the



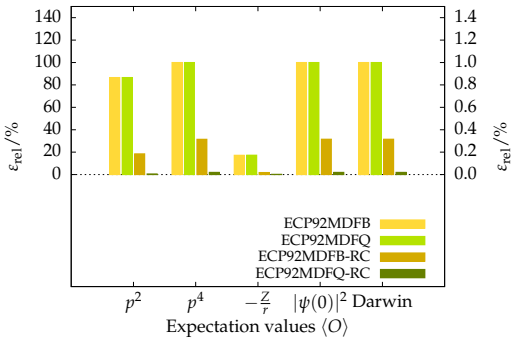
(a) Kr: ECP28MHF with NR reference, ECP28MWB and ECP10MDF with SR reference.



(b) Xe: ECP46MHF with NR reference, ECP46MWB and ECP28MDF with SR reference.



(c) Rn: ECP78MHF with NR reference, ECP78MWB and ECP60MDF with SR reference.



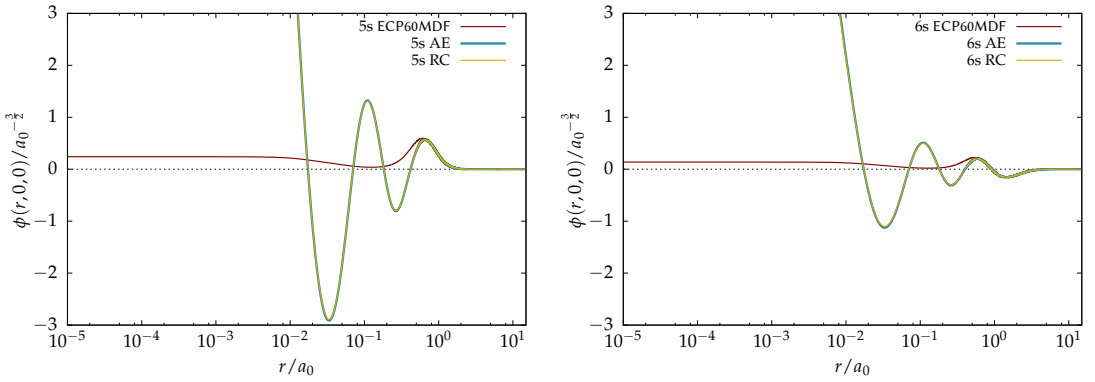
(d) Og: ECP92MDFB and ECP92MDFQ with SR reference.

Figure 3.4.: Comparison of core-like expectation values for Kr, Xe, Rn and Og. The bars calculated in bright colors are calculated with PAOs and the relative error w.r.t. the corresponding AE calculation is shown on the left sides. Bars in muted colors show the same expectation values calculated with reconstructed AOs and are denoted by the suffix *RC*. The corresponding relative errors are shown on the right sides and are one to two orders of magnitude smaller than the relative errors for the POs. In each subtitle, the employed PPs and the type of reference are given. The nomenclature of the PPs is explained in App. B.

convergence of the hyperfine structure (HFS) constants with SCF iterations subsequent to the reconstruction.

3.6.3.1. PP Influence

Many EC PPs have been developed up to today. Different PPs may lead to different reconstructed AE wave functions. I examined the Stuttgart type small



(a) 5s **(b)** 6s

Figure 3.5.: 5*s* and 6*s* orbitals of Rn. The red line depicts the orbital as obtained from a PP calculation, the blue line below the yellow line belongs to an SR AE calculation and the yellow line corresponds to the orbital after the reconstruction.

and large core PPs fitted to NR and 4c relativistic calculations [116–118], a shape-consistent Hay and Wadt PP fitted to a Cowan–Griffin relativistic Hamiltonian [119] and a PP from Stevens *et al.* [120]. The expectation values for core-like properties are calculated orbital-wise and summed over all molecular valence orbitals (MVOs). Then the relative error w.r.t. a corresponding NR or SR reference is calculated. The results are shown in Fig. 3.6.

In general, all properties of small core PPs agree with their AE counterparts within a 5 % margin. The relative error is much larger for large core PPs. Multiple issues can be the reason: Polarization of the outer core shells is not included in large core PP calculations. Furthermore, core-core and valence-core penetration may reduce the accuracy, as the core effectively pushes electrons out. [26]

3.6.3.2. Basis Set Influence

When reconstructing an AE wave function, each PP calculation is first performed using a certain basis set b_1. The representation of the wave function in b_1 is then transformed into a representation in a different basis set b_2. If the basis b_1 is not fully representable in b_2 then information is lost in the transformation process. This may have major effects on the properties, since the renormalization in basis b_2 redistributes electrons within the system.

A basis set completeness profile shows the overlap of a test basis function with a specific basis set. The profiles were developed by Chong [121] in order to

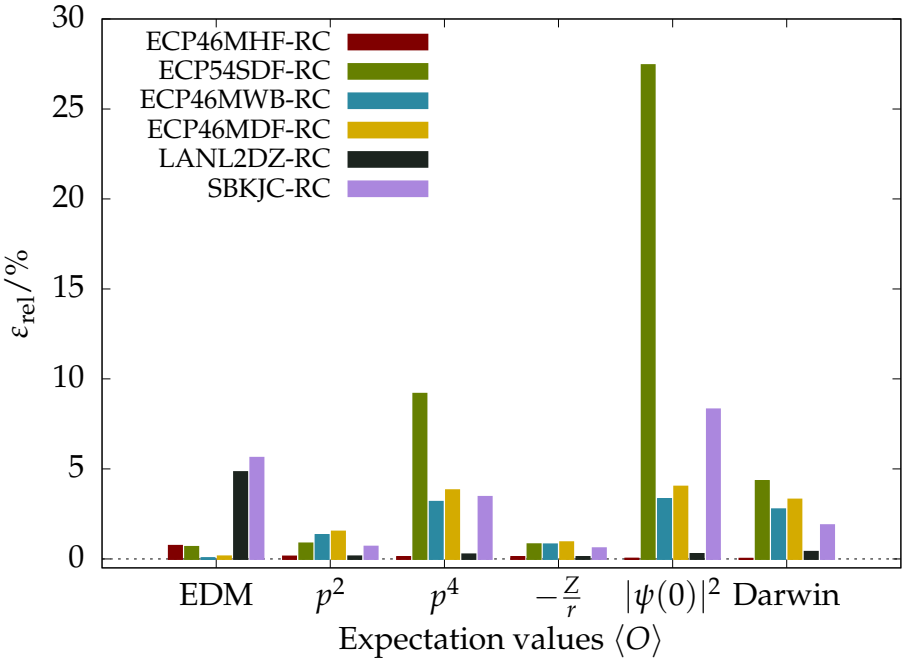
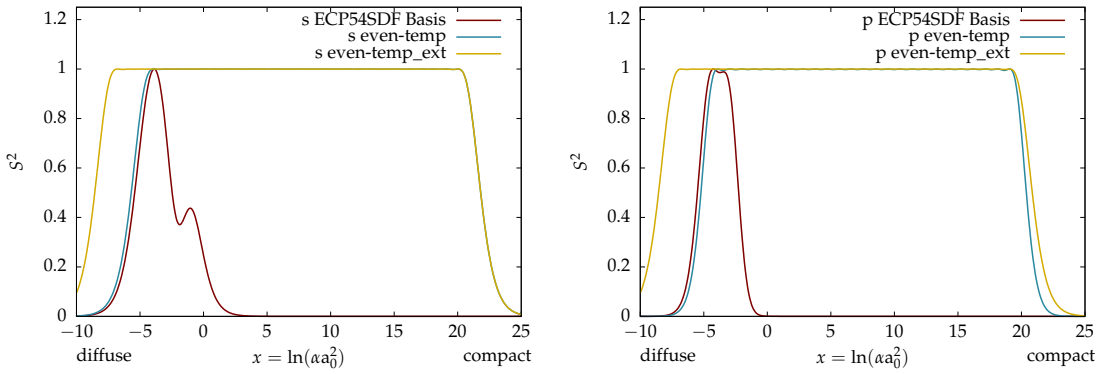


Figure 3.6.: Expectation values of core-like properties and EDM calculated with several reconstructed AE wave functions of BaF. ECP46MHF is a NR small core PP fitted to multi-electron states. ECP54SDF is a large core PP obtained by fitting single electrons to a Dirac–Fock wave function. ECP46MWB denotes a small core PP fitted to multi-electron states employing a Wood-Boring Hamiltonian and ECP46MDF (also small core) is generated by a multi-electron fit using a DC Hamiltonian.

compare different basis sets with respect to their covered space. The working equation, adapted from Ref. [121], is a sum over squares of overlap integrals between a test basis function $\chi(\alpha_{\mathrm{test}})$ and the basis set $\{\chi_0, \chi_1, \ldots\}$.

$$S^2(\alpha_{\mathrm{test}}) = \sum_{i=1}^{N_{\mathrm{basis}}} \langle \chi(\alpha_{\mathrm{test}}) | \chi_i \rangle \langle \chi_i | \chi(\alpha_{\mathrm{test}}) \rangle \tag{3.103}$$

The discrepancies in the completeness profiles in Fig. 3.7 show that this is indeed a source of error for the ECP54SDF calculation. The profile shows that the s-type basis functions can be represented by the basis functions of the even-temp basis set (see Appendix for this and the following bases), while the first peak of the p-type basis functions is outside the area of corresponding functions of the even-temp basis. In the specific case of Fig. 3.7b the ECP54SDF-p basis functions with a coefficient of about 0.0025 are not included in the even-temp-p basis. The problem is circumvented by extending the even-temp basis with additional (in this case diffuse) functions (even-temp_ext). Similar effects can be seen for the ECP46MWB (Fig. 3.8) where the even-temp_ext also gave much better results.



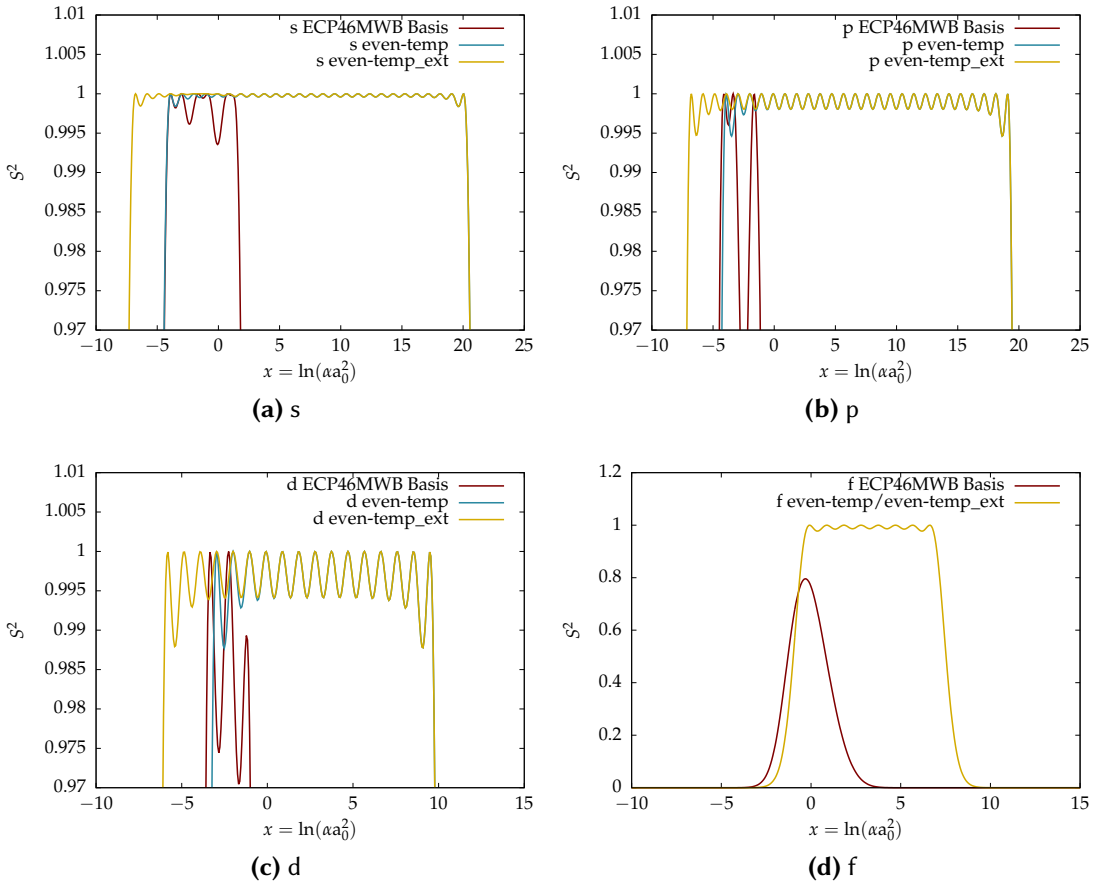
(a) s **(b)** p

Figure 3.7.: Basis set profiles for the Ba ECP54SDF valence basis set, the general even-tempered basis set and the extended even-tempered basis set. Each image shows a specific angular momentum.

In order to represent the basis set `ECP46MDF` correctly (Fig. 3.9), an even larger extended basis had to be created (`even-temp_ext2`).

Fig. 3.10 compares the results for the original `even-temp` basis used in Fig. 3.6 together with the results for the new basis sets `even-temp_ext` and `even-temp_ext2`. For all core-like properties, the values improve significantly. At a first glance that might seem strange since the extension took place in the diffuse region. The core-like properties are less influenced by this part of the orbitals and one would suggest that the main effect is on the EDM instead. However, one has to consider that the effect is of indirect nature: Overlap integrals between the PP orbitals and AE orbitals determine the weight of each atomic orbital in the reconstructed wave function. Overlap integrals are indeed strongly influenced by the tail of the wave function and thus the diffuse coefficients (see Fig. 3.3). An error in the tail region (large r) thus will have a stronger effect than an error close to the core (small r). If one improves the weights of the atomic orbitals in the reconstructed solution, one improves the overall wave function as well and hence the core-like properties are better described. This does not explain the slight worsening of the EDM, which can at this stage only be attributed to fortuitous error cancellation when using the basis set `even-temp`. As a general remark one should note that despite the unrepresentability of the basis `ECP46MWB` and especially `ECP46MDF` in `even-temp`, the calculation of core-like properties still gives remarkably good results (error smaller than 5 %).

I developed another basis named `even-temp-v7` (`v2` to `v6` failed to give the desired results) for an additional purpose: Nuclear spin-dependent parity-violating properties depend on the electronic wave function within the nucleus. [113] I



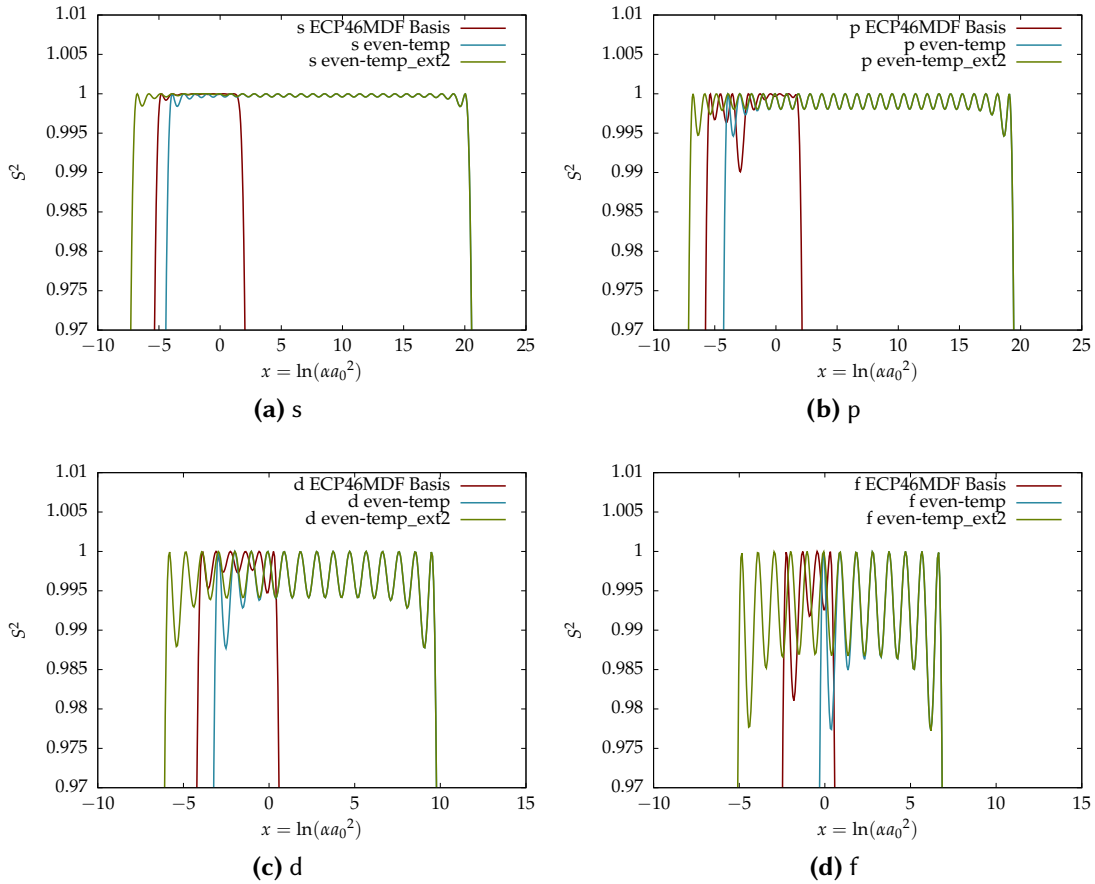
(a) s **(b)** p **(c)** d **(d)** f

Figure 3.8.: Basis set profiles for the Ba ECP46MWB valence basis set, the general even-tempered basis set and the extended even-tempered basis set. Each image shows a specific angular momentum.

devised a simple prescription in order to generate a basis that contributes within the nucleus as well as being diffuse enough to be used in molecular calculations:

1. Define the atomic charge Z.
2. Calculate the approximate mass of the nucleus with $m = 2.5Z$.
3. Calculate the root mean square radius of the nuclear charge distribution [122]

$$\begin{aligned} r_{\text{rms}} &= \left(0.863m^{\frac{1}{3}} + 0.57\right) \text{fm} \cdot 10^{-15}\text{m}\,\text{fm}^{-1}\frac{a_0}{a_0} \\ \frac{r_{\text{rms}}}{a_0} &= \left(0.863m^{\frac{1}{3}} + 0.57\right) 10^{-15}\frac{1}{\frac{a_0}{\text{m}}}. \end{aligned} \tag{3.104}$$



(a) s **(b)** p

(c) d **(d)** f

Figure 3.9.: Basis set profiles for the Ba ECP46MDF valence basis set, the general even-tempered basis set and the extended even-tempered basis set. Each image shows a specific angular momentum.

where $\{a_0\}$ denotes the numerical value of a_0 (see Ref. [123, Ch. 1.1]). The last factor ensures that r_{rms} is given in atomic units.

4. Calculate the exponential of a Gaussian nuclear charge distribution $\alpha_{\text{nuc}} = \dfrac{3}{2r_{\text{rms}}^2}$, where $\rho_{\text{nuc}}(r) = Z\rho_0 e^{-\alpha_{\text{nuc}} r^2}$.

5. Choose the largest exponential coefficient of the basis to be $\alpha_{\text{max}} = 1000\alpha_{\text{nuc}}$. This last step ensures that at least a few basis functions have a contribution within the nucleus. α_{max} generated in this manner is one order of magnitude larger than the values published by Knecht *et al.* [124] and Mastalerz *et al.* [125], but did not result in any (quasi-) linear dependencies of the Gaussian basis.

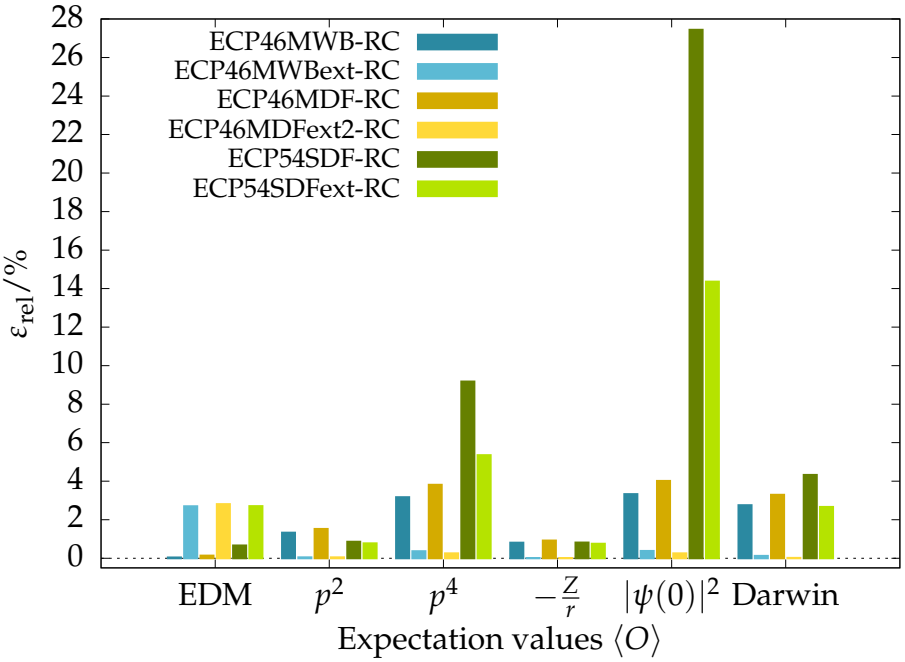


Figure 3.10.: Influence of the basis set on different core-like and valence-like properties. A properly chosen basis set significantly improves the results for all core-like properties. The deterioration of the EDM is attributed to error cancellation in case of the smaller basis.

6. Generate $n-1$ additional basis functions using the even-tempered relation

$$\frac{\alpha_i}{\alpha_{i+1}} = \frac{1}{\beta} \Leftrightarrow \alpha_i = \alpha_{\max} \left(\frac{1}{\beta}\right)^n \tag{3.105}$$

n and β are free parameters. My choices are found in App. A. For the application here, $\beta = 2.6$ was sufficient. n can be limited by computational resources, especially when dealing with high angular momentum basis functions. However, one requires the basis to incorporate the corresponding PP basis set. Thus the minimal n is determined by

$$n = \left\lceil \frac{\ln(\alpha_{\min}) - \ln(\alpha_{\max})}{\ln(\frac{1}{\beta})} \right\rceil \tag{3.106}$$

A basis set profile corresponding to this procedure is shown in Fig. 3.11 together with the exponential coefficient of the nuclear charge distribution of a Ca nucleus. Since the nuclear charge distribution of Ca is the most compact of the alkaline-earth-metals treated in this work, I designed the basis for Ca but used it also for the heavier metals. It is given in App. A.

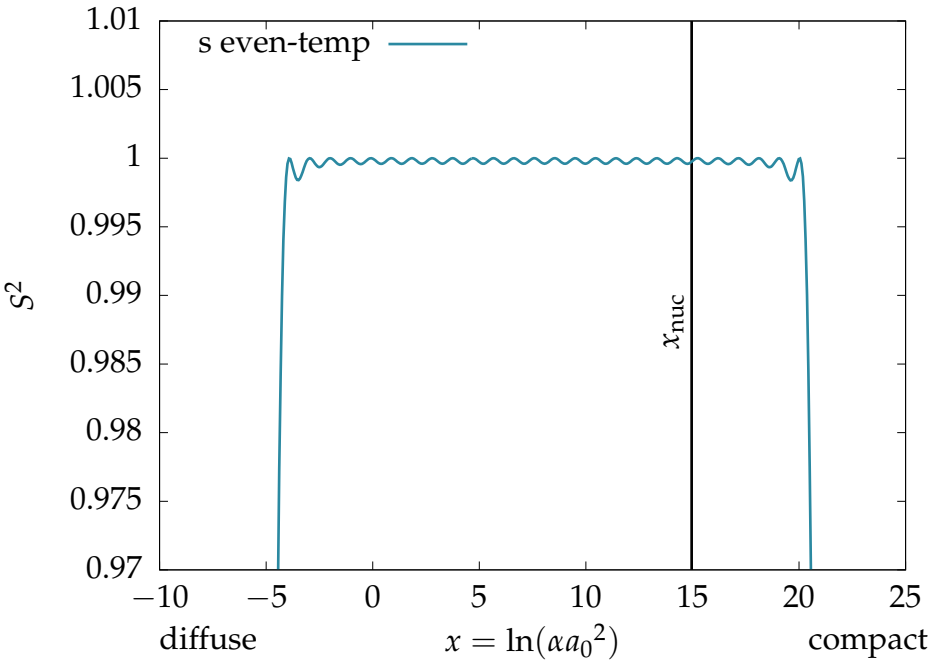


Figure 3.11.: Basis set profile of the `even-temp_v7` basis. The line denoted x_{nucleus} shows the position of the Ca nucleus' Gaussian charge distribution coefficient in the profile. The number of maxima on the right hand side (RHS) of the black line correspond to the number of basis functions with exponential coefficients larger than that of the nuclear charge distribution.

3.6.3.3. Virtual Orbital Influence

Apart from the effect in the previous section, there is another source of error in the reconstruction scheme: In step 4 of the protocol, I select the AE valence and virtual AOs to construct the mapping basis $\boldsymbol{M}$. My analysis of the PO in Sec. 3.2.7 showed further that high energy AOs can smooth out the core tail of valence AOs. If one would include these orbitals in $\boldsymbol{M}$, the reconstructed orbital would have a significantly altered core tail compared to the AE reference. Thus one needs to restrict the set of AOs which is used for the reconstruction. A simple choice might be the set of occupied AOs. Although this works for ground state atoms (see Figs. 3.4) it is not imperatively working for molecules: As explained in Ch. 3.5, the admixture of virtual AOs to occupied AOs accounts for the proper polarization of the occupied AOs in MOs. Including a few low lying virtual orbitals into the reconstruction proved sufficient. Fig. 3.12 shows how the valence and core-like properties change with the number of virtual AOs included in the reconstruction.

The figure shows that the EDM improves significantly with the use of

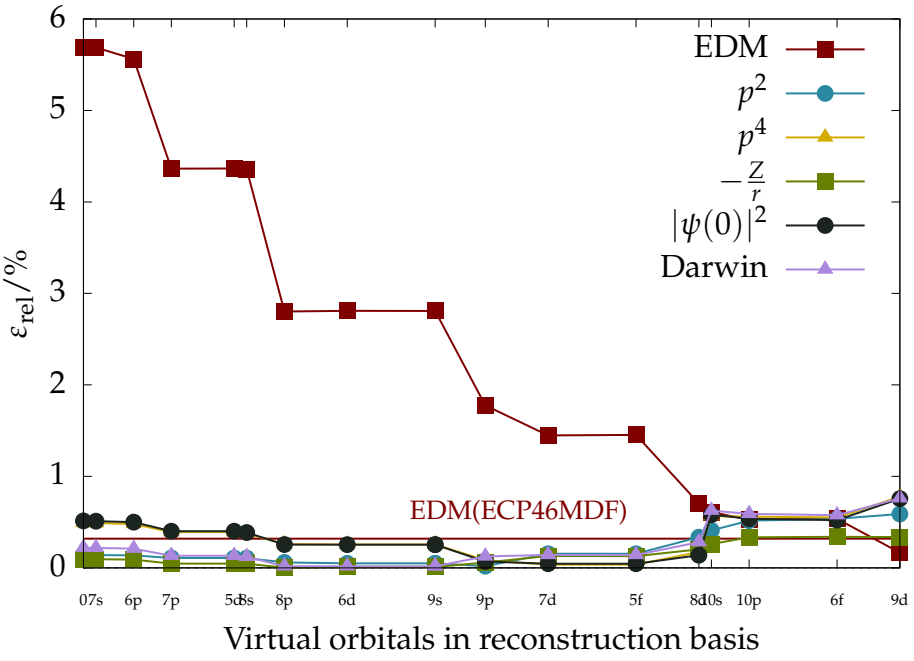


Figure 3.12.: Dependence of relative error of expectation values on the number of virtual AOs used in the reconstruction. The expectation values are calculated for the valence electrons only and referenced by a corresponding AE calculation. The x-axis labels correspond to the order and type of the orbital as given in the output of a HF calculation. The Aufbau principle does not apply for virtual orbitals.

virtual AOs in the reconstruction. The core-like properties, however, do not show a strong dependence on the virtual AO number. Only when the 10s shell is included, the values of the core-like properties start to worsen. It is likely that at this point virtual AOs are added that smooth the core tails of the MOs.

3.6.3.4. Comparison to Alternative Methods

So far, I calculated all expectation values with a self-written computer code. However, one of the goals of this thesis is to derive a method that can be easily attached to other computational tools and methods. Thus I used reconstructed SR orbitals as an initial guess for a 2c ZORA calculation and subsequent calculation of properties. The properties of choice are here the parallel and perpendicular component of the hyperfine coupling tensor as implemented in a modified [112, 113, 126–128] version of TURBOMOLE [95]. In Tab. 3.5 the tensor components are shown, when calculated directly after reconstruction and after full convergence of the SCF equations ($\approx$ 10–12 steps). The values are compared to the GRECP

Table 3.5.: ^{137}BaF hyperfine structure constants for the ^{137}Ba atom. The magnetic moment $\mu_{^{137}\mathrm{Ba}} = 0.93734$ and the nuclear spin $I_{^{137}\mathrm{Ba}} = \frac{3}{2}$ where taken from Ref. [131]. The relative errors were calculated with the values of Ref. [130] as reference.

Method	$A_\parallel$/MHz	$\varepsilon_{\mathrm{rel}}(A_\parallel)$/%	$A_\perp$/MHz	$\varepsilon_{\mathrm{rel}}(A_\perp)$/%
exp[129]	2453(9)		2401(9)	
exp[130]	2376		2301	
GRECP/SCF/NOCR[28][1a]	1479	37.3	1446	37.2
GRECP/RASSCF/NOCR[28][1b]	1488	37.4	1455	36.8
GRECP/SCF/NOCR/EO[28][1c]	2264	4.7	2186	5.0
GRECP/RASSCF/NOCR/EO[28]	2272	4.4	2200	4.4
ECPP/UHF/RC0[2a]	2022	14.9	1898	17.5
ECPP/GHF-ZORA/RC[2c]	1934	18.6	1863	19.0

[1a] GRECP=*Generalized relativistic effective core potential*, NOCR=*non-variational one center reconstruction*
[1b] RASSCF=*restricted active space SCF*
[1c] EO=*effective operator*
[2a] ECPP=*energy consistent pseudo potential*, UHF=*unrestricted HF wave function*, RC0=*reconstruction and direct property calculation (w/o SCF)*
[2b] GHF-ZORA=*generalized HF with ZORA*
[2c] RC=*reconstruction and full SCF before property calculation*

scheme of Titov [28]. All of the reconstructed values agree very well with the AE values. Also, they are in good agreement with the results of Titov and coworkers and in the same order of magnitude as the experimental results of Knight [129] and Ryzlewicz [130]. The relative shift of $A_\parallel$ w.r.t. $A_\perp$ is in good agreement with the experimental results. It should be emphasized that the PPs used for the reconstruction have not been derived for this purpose. That the reconstruction gives results with reasonable accuracy (especially when comparing to PPs that have been designed for this purpose), is very remarkable.

3.6.4. Nuclear Charge Scaling of HFS Constants

A simple feature of quantum chemistry is that trends throughout the periodic table are easily calculated. The scaling of a certain quantity with the atomic charge Z can be extrapolated from a set of calculations where one atom is changed at a time. Relativistic effects also follow a certain Z-scaling, which depends on the analytic form of the wave function. The factor that incorporates this scaling is called *relativistic enhancement factor* R_A. R_A usually incorporates all non-polynomial contributions of Z and the fine-structure constant α to the expectation value of

an operator. Dividing the expectation values by the enhancement factor, the least-squares fit of a double logarithmic plot of the expectation values against the atomic charge can be used to determine the scaling with Z: The slope of the fit line determines the polynomial degree. E.g., Isaev and Berger predicted in Ref. [113] that the parity violating parameter W_a of the hyperfine structure scales as $R_A Z^2$ along groups and as $R_A Z^k$ along the fourth ($k \approx 4$) and fifth ($k \approx 6$) period of the periodic table.

In the year 1933 Fermi and Segrè estimated the scaling of hyperfine structure parameters with Z. [132, 133] They used an analytic expression for the hyperfine coupling constant of a single s electron

$$a_A = \frac{16\pi}{3} \frac{\mu_A}{I_A} |\psi_{\text{eff}}(0)|^2 \tag{3.107}$$

and inserted an analytical formula for the wave function at $r = 0$.

$$|\psi_{\text{eff}}(0)|^2 = \frac{R_A}{\pi {a_0}^3} \frac{Z\alpha}{E_{\text{h}}} \frac{\partial E(n)}{\partial n} \tag{3.108}$$

where E_{h} is the Hartree energy, $E(n)$ is the energy as a function of the quantum number n and α is the fine structure constant. They obtained an empirical relativistic scaling factor R_A from experimental data and provided it as the analytical expression

$$R_A = \frac{1}{\left(1 - (Z\alpha)^2\right)^2}. \tag{3.109}$$

Assuming the energy derivative to be independent of Z, the contact density $|\psi|^2$ divided by R_A should then scale linearly with Z, i.e.,

$$\frac{|\psi_{\text{eff}}(0)|^2}{R_A} = cZ^k, \qquad k \approx 1. \tag{3.110}$$

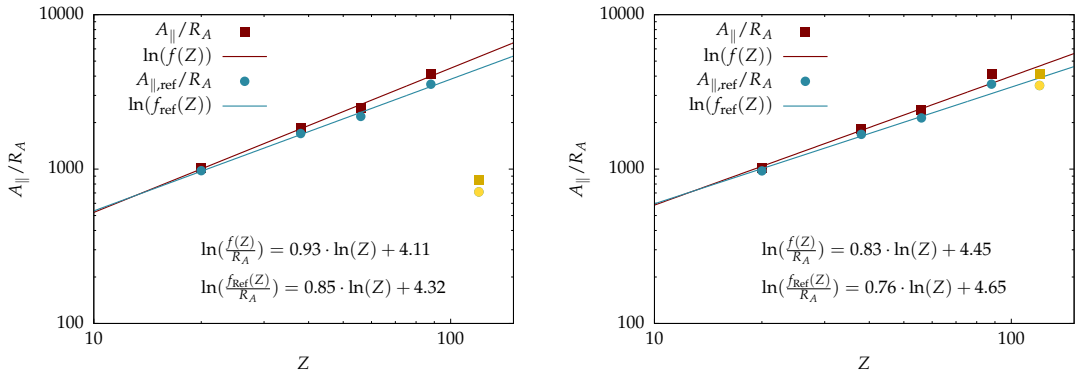
c incorporates all the constants.

Sushkov, Flambaum and Khriplovich [134] took up the theory of Fermi and Segrè and accounted for the mixing of configurations with $l > 0$. They find the same Z-scaling law but a different relativistic scaling factor for an $s_{\frac{1}{2}}$ state, in fact recovering the enhancement factor published by Racah [135].

$$R_A = \frac{3}{\gamma(4\gamma^2 - 1)} \tag{3.111}$$

$$\gamma = \sqrt{\chi^2 - (Z\alpha)^2} \tag{3.112}$$

$$\chi = (-1)^{j-l+\frac{1}{2}} \left(j + \frac{1}{2}\right) \tag{3.113}$$



(a) R_A according to Racah and Sushkov, Flambaum, Khriplovich.

(b) R_A according to Fermi, Segrè.

Figure 3.13.: Nuclear charge scaling of $A_\parallel/R_A$ for XF. A double logarithmic plot of $A_\parallel/R_A$ for XF, where X $\in \{\mathrm{Ca}, \mathrm{Sr}, \mathrm{Ba}, \mathrm{Ra}, \mathrm{Ubn}\}$, was used to determine the scaling of $A_\parallel$ with Z. The red line shows values obtained after two SCF iterations after the reconstruction. The blue line shows $A_\parallel$ as obtained from wave functions generated through an SCF procedure.

I chose the Z-scaling of the hyperfine coupling tensor component $A_\parallel$ as a test case. It is a core-like property and, after dividing by R_A, should scale linearly with Z. One should keep in mind, though, that the scaling law was derived for atomic systems. The parameter was calculated for the alkaline earth metals in XF where X $\in \{\mathrm{Ca}, \mathrm{Sr}, \mathrm{Ba}, \mathrm{Ra}, \mathrm{Ubn}\}$ ($Z_{\mathrm{Ubn}} = 120$). Since the ZORA model potential (see Ref. [114]) of Ubn was not available, I assumed the potential to be similar to the potential of Og and renormalized it to $Z = 120$.

In Fig. 3.13 I show a double logarithmic plot of the $\frac{A_\parallel}{R_A}$ as a function of Z. k of (3.110) is then obtained as the slope of the line. Fig. 3.13a shows the Z-scaling of $A_\parallel$ using the (absolute value of the) relativistic correction factor of Racah while Fig. 3.13b shows the same data using the R_A of Fermi and Segrè.

Firstly, I should note that the linear regression in Fig. 3.13a has been done only with the elements Ca through Ra. The scaling factor for Ubn according to (3.111) not only was unreasonably high, it also showed a negative sign. This is clearly an artifact of formula (3.111) as can be seen in Fig. 3.14.

The curve of R_A is continuous and smooth up to Og. From element 119 on the curve continues erratically. Analysis of (3.111) shows that the equation has

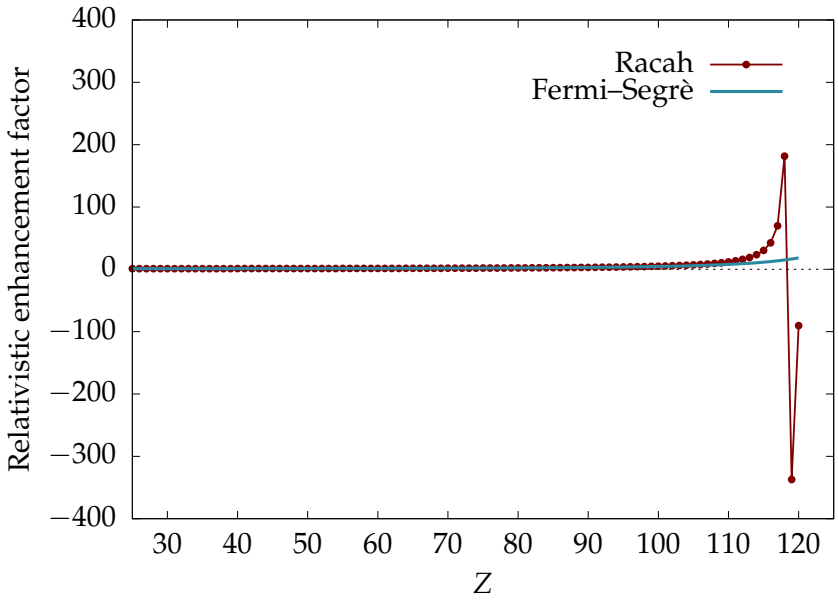


Figure 3.14.: Relativistic enhancement factors of the hyperfine coupling matrix elements for elements H through Ubn. The red line is calculated according to the formula given in Ref. [134] with $j = \frac{1}{2}$ and $l = 0$, whereas the blue line shows the relativistic enhancement factor of Fermi and Segrè. [132, 133].

four different roots:

$$Z_{1,2} = \pm\frac{1}{\alpha} \tag{3.114}$$

$$Z_{3,4} = \pm\frac{\sqrt{3}}{2\alpha} \tag{3.115}$$

The first two roots are present in (3.109) as well, whereas the roots three and four are a feature of (3.111). When using $\alpha = \frac{1}{137}$ the positive roots appear at $Z = 118.645,137$ and hence the formula is not applicable beyond Og.

The root at $Z = 137$ has its origin in the point-like nucleus approximation which fails for super heavy elements. Hence the relativistic enhancement factors are wrong for these nuclei. Dinh *et al.* observe in Refs. [136, 137] the failure of the factor (3.111) beyond $Z = 118$. According to the authors the treatment of finite nuclear size as small correction to the hydrogen-like wave functions is not a good approximation for elements beyond $Z = 100$.

For the elements Ca through Ra, Fermi's and Racah's relativistic corrections calculated from reconstructed wave functions show only minor differences. The latter regression shows a scaling with $Z^{0.93}$ while the former, including Ubn,

shows a scaling with $Z^{0.83}$. Both values are close to 1 which would be the expected scaling for atoms. Furthermore, both values are close to the reference results obtained with a converged 2c ZORA wave function. For these wave functions the scalings are $Z^{0.85}$ and $Z^{0.76}$, respectively.

Despite the failure of relativistic enhancement factors beyond $Z \approx 125$ even for (3.109), Gaul *et al.* showed in Ref. [138] that the ratio of expectation values of parity- and time-violating properties maybe predicted for super-heavy elements as well. However, to solve the problem entirely, one would need to derive the proper scaling laws from first principles using a finite nucleus model. Alternatively, one could also assume that the relativistic scaling laws hold for super-heavy elements. Then, the relativistic enhancement factors could be extracted from numerical calculations of atoms in employing a finite nucleus model. A polynomial fit to the data could provide at least a numerical relativistic enhancement model, applicable to the periodic table proposed by Pyykkö [27].

3.6.5. Reconstruction of Larger Molecules

3.6.5.1. Influence on Parity Violating Properties

Strontium methoxide, $CH_3OSr^{\bullet}$, has recently been used in laser cooling experiments by Kozyryev *et al.* [139], following a general analysis of polyatomic molecules for laser cooling in Ref. [140]. The authors state that isotopic substitution in the methyl group, as proposed in Ref. [140], would give rise to chiral targets that are interesting in the study of the violation of fundamental symmetries in physics as well as paving the way for the study of "[...] ultra cold chemistry with chiral molecules." Since the study of parity-violation needs exact wave functions close to nuclei and should incorporate relativistic effects, a treatment with PPs becomes difficult.

I tested the reconstruction scheme on $CH_3OSr^{\bullet}$. For the study, I chose the `even-temp-v7` basis developed in Sec. 3.6.3.2. The employed PP at the Sr atom was of the Stuttgart-type with 28 electrons included in the core of Sr. [118] The basis set was taken from the same reference. On all other atoms the Ahlrichs basis `def2-TZVP` was employed. All properties were calculated directly without further optimization of the wave function, except for $A_{\parallel}$ and W_a, which were calculated after one additional SCF iteration. The AE reference calculation was fully converged with a SR-ZORA UHF scheme. $A_{\parallel}$ and W_a were then calculated with a 2c-ZORA generalized Hartree–Fock (HF) program after additional SCF iterations. The results are shown in Tab. 3.6.

Table 3.6.: Core-like properties of $CH_3OSr^{\bullet}$. The reconstructed wave function was obtained from a Sr-ECP28MDF non-relativistic calculation for the molecule in combination with an Sr-even-temp-v7 scalar-relativistic calculation for the atom. The reference wave function was obtained using converged scalar-relativistic orbitals. $A_{\parallel}$ and W_a were calculated after two iterations using an GHF two-component relativistic algorithm.

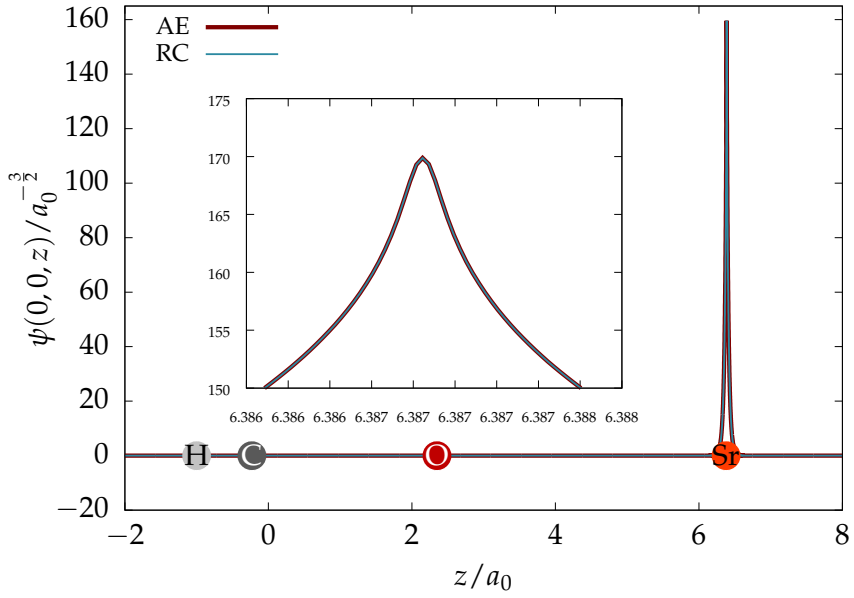
Property	RC	Reference	$\varepsilon_{\mathrm{rel}}/\%$
Kinetic energy$/E_{\mathrm{h}}$	3480.862	3481.448	1.68×10^{-2}
Mass velocity$/E_{\mathrm{h}}$	−563.523	−563.527	7.88×10^{-4}
$\langle \frac{Z}{r} \rangle / E_{\mathrm{h}}$	−8298.279	−8299.709	1.72×10^{-2}
$\lvert\Psi(\mathrm{Sr})\rvert^2 / {a_0}^{-3}$	64 660.140	64 660.289	2.31×10^{-4}
1-e$^-$-Darwin corr.$/E_{\mathrm{h}}$	205.791	205.793	8.85×10^{-4}
$A_{\parallel} / \frac{\mu_{\mathrm{Sr}}}{I_{\mathrm{Sr}}}$	1893.161	1824.767	3.75
W_a/Hz	−40.869	−39.177	4.32

Fig. 3.15 shows the Sr 1s and 2s core orbitals together with the corresponding $CH_3OSr^{\bullet}$ $1A_1$ and $2A_1$ orbitals. Visually, the orbitals compare very well and underline the observation from Tab. 3.6 that the frozen-core approximation holds. Also, the valence orbital $10A_1$ is reconstructed very well (see. Fig. 3.16). However, the reconstructed highest-occupied molecular orbital (HOMO) is only approximately identical to its AE counterpart. The exact features are not reconstructed, but the relative contribution of this orbital to the one-electron expectation values is relatively small and hence the relative error of AE orbital expectation values compared to reconstructed (RC) orbital expectation values is small as well. W_a is dominated by the orbital of the unpaired electron inside the Sr nucleus. Although this orbital does not show the exact features of the AE reference orbital, it reproduces the nodal structure and amplitudes at the Sr nucleus very well. Thus, the expectation values are in good agreement.

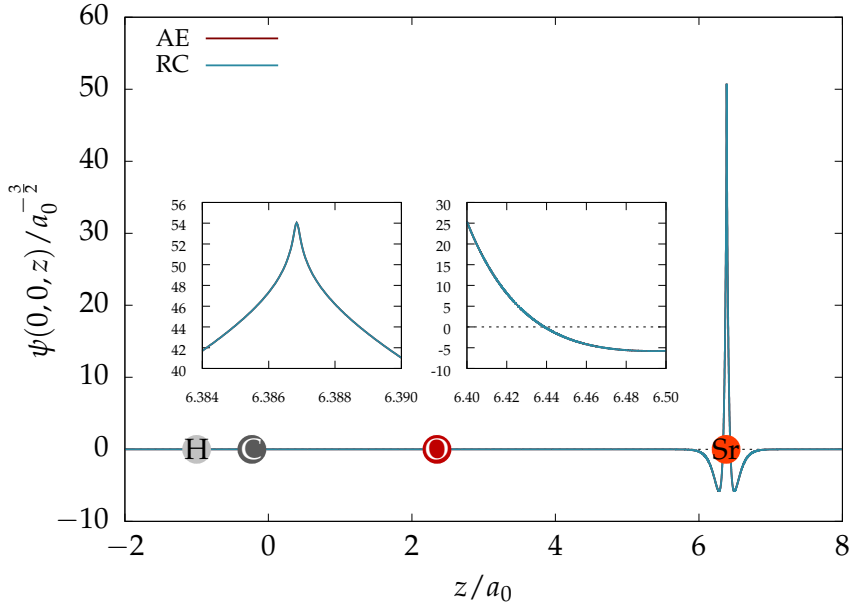
3.6.5.2. Influence of Molecular Symmetry

In this section I show that the reconstruction can be used independently of the symmetry of the molecules under investigation. Two test systems are therefore calculated: The tetrahedral molecule $PbCl_4$ and the C_1 symmetric (though tetrahedral-like) molecule NWHClF. While the former was chosen purely for didactic purposes, the latter has been discussed among other halides for its applicability in the observation of parity violation in molecules. [1, 141]

For the NR calculation of $PbCl_4$ I used the ECP60MDF on Pb, together with the basis set ECP60MDF_VTZ on Pb and def2-TZVP on Cl. For NWHClF I used the

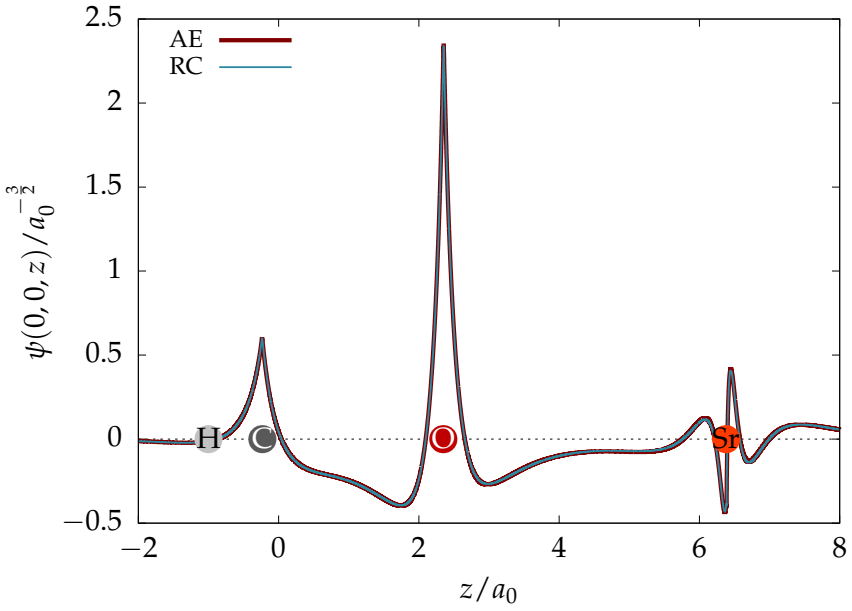


(a) 1s

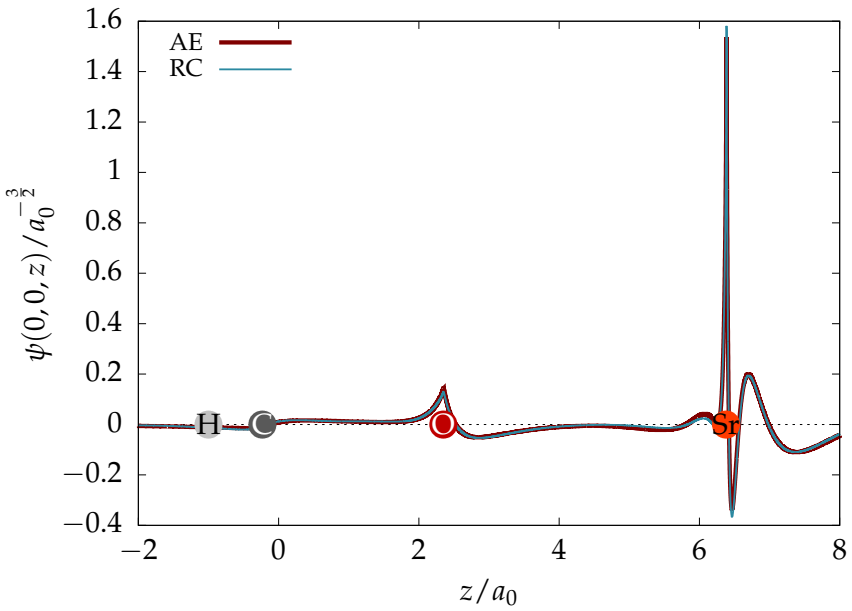


(b) 2s

Figure 3.15.: z-component of the 1s- and 2s-like orbitals of Sr in $CH_3OSr^{\bullet}$. The orbitals have been phase matched for better visualization. The red line shows the AE reference orbital while the blue line shows the orbital as obtained from a reconstruction. The red line is barely visible, as it lies directly behind the blue line.



(a) $10A_1$



(b) $14A_1$

Figure 3.16.: Comparison of orbital $10A_1$ and $14A_1$ of $CH_3OSr^{\bullet}$, RC vs. AE. The z-component is shown only. The orbitals have been phase matched for better visualization. The red line shows the AE reference orbital whereas the blue line shows the orbital as obtained from a reconstruction.

Table 3.7.: Core-like properties of $PbCl_4$ and NWHClF. The reconstructed $PbCl_4$ wave function was obtained from a Pb-ECP60MDF non-relativistic calculation for the molecule in combination with an Pb-`even-temp-v7` scalar-relativistic calculation for the atom. The reconstructed NWHClF wave function was obtained from a W-ECP60MDF non-relativistic calculation for the molecule in combination with an W-`even-temp-v7` scalar-relativistic calculation for the atom. The reference wave functions were obtained using converged scalar-relativistic orbitals.

Molecule	Property	RC	Reference	ε_{rel}/%
$PbCl_4$	Kinetic energy/E_h	30 387.001	30 394.517	2.47×10^{-2}
	Mass velocity/E_h	−236 379.512	−236 383.576	1.72×10^{-3}
	$\langle -\frac{Z}{r} \rangle / E_h$	−61 228.350	−61 251.290	3.75×10^{-2}
	$\lvert\Psi(\text{Pb})\rvert^2 / {a_0}^{-3}$	3 699 519.749	3 699 866.901	9.38×10^{-3}
	1-e^--Darwin corr./E_h	25 392.854	25 395.259	9.47×10^{-3}
NWHClF	Kinetic energy/E_h	21 117.681	21 120.805	1.48×10^{-2}
	Mass velocity/E_h	−87 494.319	−87 487.314	8.01×10^{-3}
	$\langle -\frac{Z}{r} \rangle / E_h$	−43 871.848	−43 890.212	4.18×10^{-2}
	$\lvert\Psi(\text{W})\rvert^2 / {a_0}^{-3}$	1 821 741.360	1 821 607.632	7.34×10^{-3}
	1-e^--Darwin corr./E_h	11 280.916	11 280.089	7.33×10^{-3}

ECP60MDF PP on W together with the `def2-TZVP` basis set and a the `def2-SVP` basis set on all other atoms. The atomic reference calculation was obtained within the `even-temp-v7` basis published herein. The core-like properties were calculated with the converged and the reconstructed SR HF wave function. All values are given together with the relative error in Tab. 3.7. They agree very well with the reference values.

3.6.6. Substituent Influence on Core-like Properties

Most interesting is the application of the reconstruction method to the prediction of substituent influences on core-like properties. Only if this is possible, the reconstruction scheme will be of value as absolute energy values rarely play a role in experiments. The test systems are NWHBrF and NWHFI, two compounds previously investigated in the context of parity violation by Figgen *et al.* [1]. I used molecular equilibrium structures from this reference. Furthermore, I employed on Br and I the PPs from Ref. [142], which were removed through sequential application of the reconstruction method. In this setup, the order of the reconstruction changes the final orbitals. I found that the influence of the reconstruction order on core-like properties is relatively small and the results differ by at most 0.01 % (see Tab. 3.8). This is approximately the error range w.r.t. comparable AE expectation values (see App. C) and I thus assume both

Table 3.8.: Influence of reconstruction order on core-like properties. The values were obtained by reconstructing first the MOs at the position of I and then at the position of W (first data column) and vice-versa (second data) column. The relative error of the first reconstruction order w.r.t. the second reconstruction order is very small.

Property	1.I, 2.W	1.W, 2.I	$\varepsilon_{\mathrm{rel}}/\%$
Kinetic energy$/E_{\mathrm{h}}$	28 615.514	28 614.654	3.00×10^{-3}
Mass velocity$/E_{\mathrm{h}}$	−92 971.585	−92 981.451	1.06×10^{-2}
$\langle -\frac{Z}{r} \rangle / E_{\mathrm{h}}$	−61 376.156	−61 372.908	5.29×10^{-3}
$\lvert\Psi(\mathrm{W})\rvert^2 / {a_0}^{-3}$	1 821 739.646	1 821 748.476	4.85×10^{-4}
$\lvert\Psi(\mathrm{I})\rvert^2 / {a_0}^{-3}$	280 475.890	280 465.046	3.87×10^{-3}
1-e$^-$-Darwin corr.$/E_{\mathrm{h}}$	12 520.193	12 520.199	5.28×10^{-5}
EDM$/e\,a_0$	552.773	552.845	1.30×10^{-2}

reconstructed orbitals to be identical within the scope of this work. Alternative methods for the orthogonalization of the basis functions of all atoms without PP in a molecule w.r.t. to AE ACOs have been proposed by Cook [105] and Baerends *et al.* [106]. In Fig. 3.17 I show the percental change of reconstructed and AE core-like properties w.r.t. the molecule NWHClF. The shifts are in very good agreement, except for the electron density at the W nucleus. However, the contact density at W changes only very little between the chosen test molecules (see Tabs. C.1–C.3). The electron density change at the tungsten atom in these systems may simply be negligible and thus is not a good quantifier for chemical substituent effects.

The one-electron Darwin correction, which is directly derived from the contact density, is in contrast very well reconstructed. But this is attributed mainly to the significant change of electron density at the Br and I nucleus (see Tabs. C.2 and C.3).

3.7. Conclusion

In this chapter, I explained and tested a scheme for the reconstruction of AE orbitals from POs, which were obtained using EC PPs. Most reconstruction approaches exploit the SC of POs, which is guaranteed by the specific design of the corresponding PPs. I gathered arguments that justify the shape consistency assumption for EC PP as well, leading to a straightforward reconstruction protocol based on overlap integrals.

Numerous tests on open- and closed-shell systems using PPs for group-2, -6, -14 and -17 atoms show the applicability of the protocol. Core and valence properties were calculated in the NR, SR and 2c relativistic approximations. Parity-conserving and parity-violating properties were equally well reconstructed, showing the broad application spectrum of the procedure. The most expensive step in the procedure is the Gram–Schmidt orthogonalization of the basis $\{\chi^{(\mathrm{R})}\}$ which scales usually as $\mathcal{O}(MN^2)$ (M being the number of reconstructed orbitals and N being the number of basis functions). Compared to the standard calculation of two-electron integrals in the HF SCF procedure, my method is thus *efficient*. It is likely that a prescreening of the overlap integrals can further improve the efficiency.

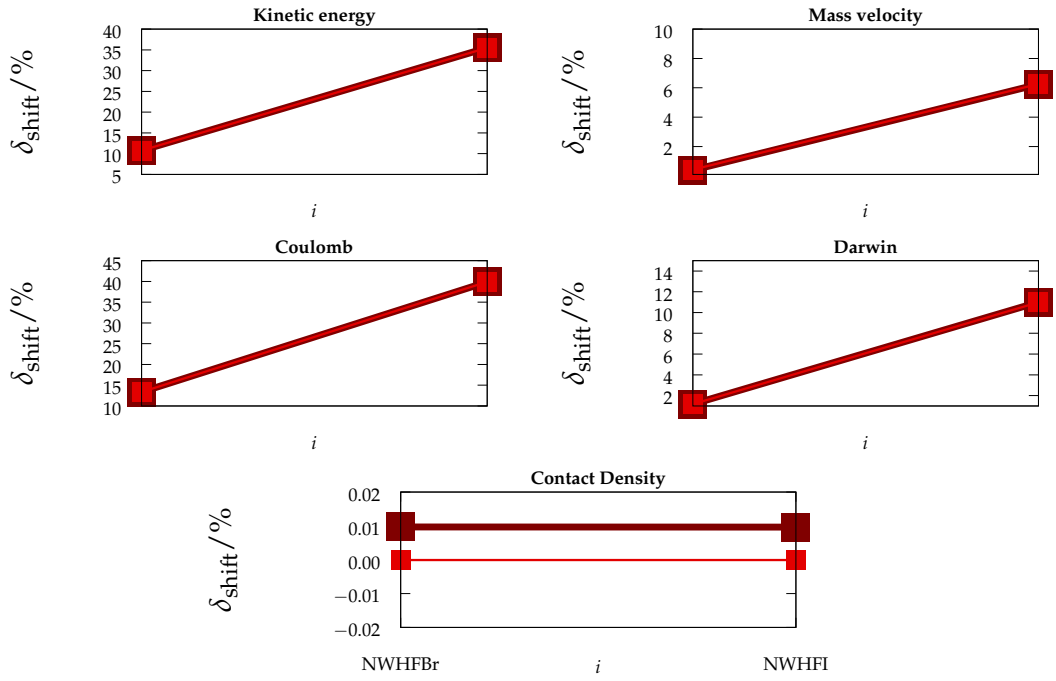


Figure 3.17.: Substituent influence on the expectation values of NWHFX with X = {Br, I}. Each left data point shows the relative shift of the expectation value of NWHBrF w.r.t. NWHClF and each right data points shows the shift for NWHFI w.r.t. NWHClF. The lines are shown to guide the eye. The dark red line connects the AE reference results whereas the bright red line connects the shifts calculated with reconstructed orbitals. My method correctly predicts the substituent influence for all properties but the contact density at the W nucleus. However, the electron density at W showed only very little deviation, thus rendering the systems inept for the determination of trends in the contact density.

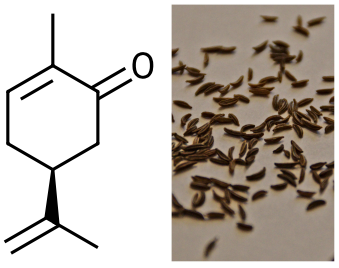
I could show that the error in the reconstruction is in general small but is reduced by including a number of unoccupied AE AOs in the reconstruction basis. Predefining the number of unoccupied AE used in the reconstruction as the active space used for energy-fitting the PP paves the way towards a black-box procedure. As the corresponding quantum chemical calculations for the atoms are cheap on modern computer hardware, the AE AOs can be constructed on the fly by a proper SR quantum chemistry software package. This can be done on top of converged PP SCF orbitals, leading to *backwards-compatibility* of the method. The orbitals obtained via this route can either be used for the calculation of approximate values of core-like properties or as initial guess for subsequent, more sophisticated calculations, e.g. 2c ZORA SCF iterations. In addition to the numerous tests on EC PPs I could also show the *practicality* of the approach by the successful reconstruction of orbitals of SC pseudo systems.

However, a prescription for an in principle exact reconstruction of AE MCSCF AOs from their PP counterparts by connecting both calculations is still desirable. Although I did not present such a rigorously *general* prescription, I developed a procedure fulfilling nearly all requirements set in the beginning and hence have the means available to tackle the AE orbital reconstruction in many calculations: My approach is indeed *applicable* to molecular systems. The AE AOs can be precalculated for the whole periodic table. This needs to be done only once and thus allows for a *black-box* implementation as easy as the generation of extended Hückel orbitals as start for SCF procedures in modern quantum chemistry programs. In comparison to similar AE calculations, the method is *accurate* and permits the calculation of chemical trends and trends throughout the periodic table.

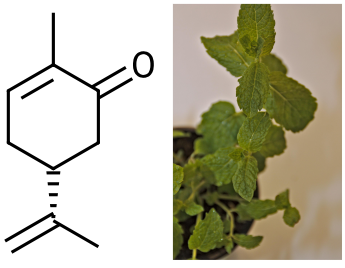
4. Direct Determination of Absolute Configuration

4.1. Introduction

The chiral compound NWHClF, which was discussed last in the previous chapter, and its structural analogs NWXYZ with $X, Y, Z = H, F, Cl, Br$ or I have been proposed in Refs. [1, 141] as candidates for the detection of parity violation in molecules. Figgen *et al.* predicted in Ref. [1] the gasphase structures and discussed the N−W stretching mode as a probe for parity violation. In Ref. [141], Nahrwold *et al.* predicted the parity violating effects on the chemical shielding constants of the tungsten nucleus. As already outlined in Ch. 1.3, some hypotheses connect the homochirality on planet earth to parity-violating effects that lower the energy of one enantiomer w.r.t. the other. However, the reasons for homochirality are not known and the study of chiral molecules and chiral interactions remains of great importance in chemistry and physics. Above all, the direct determination of absolute configuration of chiral centers is, as of yet, non-trivial. The major role of chiral centers in stereo-selective chemical reactions and for the biological activity of compounds only stresses the importance for methods that do not rely on exhaustive computation to identify the absolute configuration of stereocenters. The smell of the carvones is but one example for the distinct biological activity of chiral compounds: (*S*)-(+)-carvone has the odor of caraway, whereas (*R*)-(−)-carvone smells like spearmint.



(a) (*S*)-(+)-carvone



(b) (*R*)-(−)-carvone

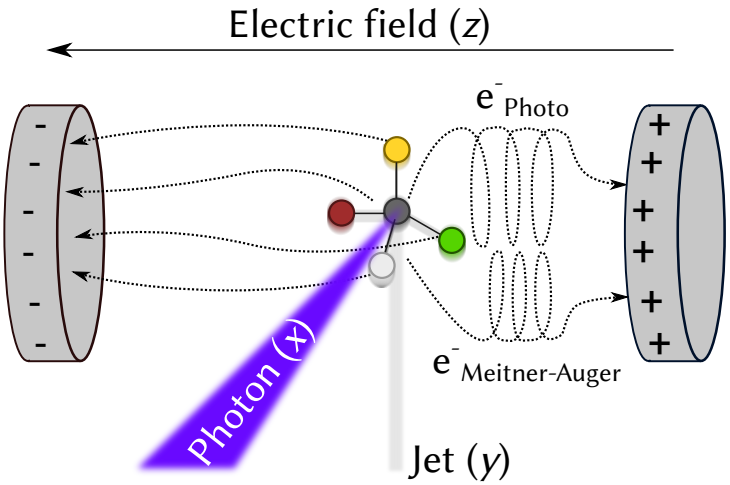


Figure 4.1.: Design of the COLTRIMS experiment. A beam of molecules enters the apparatus in y-direction. A short laser pulse in x-direction multiply ionizes molecules in the beam, which subsequently explode. An electric field in z-direction projects the fragments onto an MCP detector, where the time and position of the hits are recorded in coincidence.

The knowledge of the absolute configuration is usually a prerequisite for the study of the interaction of chiral systems. However, only a few methods exist for a direct determination of chirality. One of the approaches is Coulomb explosion imaging (CEI) [143], e.g., with the COLTRIMS reaction microscope [144]: A beam of gas phase molecules is injected in y direction into the spectrometer and crossed with a pulsed laser beam or high energy synchrotron radiation (x-direction). The photons "kick out" multiple electrons, either from the valence shells (laser) or atomic core shells, e.g., the L-shell of Br (synchrotron). Coulomb explosion of the multiply charged molecule follows. The atomic or molecular fragment ions of this explosion are projected by an electric field in z-direction onto an multi-channel plate (MCP) detector. The design of the MCP allows a position- (x, y) and time-sensitive measurement (see Fig. 4.1) of the fragment impacts. These data can be used in order to reconstruct the linear momenta $\vec{p_i}(t_i)$, which the fragments obtained during the Coulomb explosion. For this purpose one equates Newton's force equation for an accelerated object with the force applied on a point charge by a homogeneous electric field E_z in direction z of the laboratory frame.

$$\vec{F}_{i,z} = m_i a_{i,z} = m_i \frac{\mathrm{d}^2 r_{i,z}}{\mathrm{d}t^2} = q_i E_z \tag{4.1}$$

Integration of the equation of motion leads to a time-of-flight (TOF) equation

$$\begin{aligned} t_{\mathrm{r},i} &= \frac{p_{0,i,z}}{E_z q_i} \pm \frac{\sqrt{m_i \left(m_i v_{0,i,z}^2 + 2 E_z q_i (r_{i,z}(t_{\mathrm{r},i}) - r_{i,z}(t_{0,i})) \right)}}{E_z q_i} \\ t_i &= t_{0,i} + t_{\mathrm{r},i} \end{aligned} \tag{4.2}$$

where $t_{0,i}$ and $r_{0,i,z}$ are the TOF and position of fragment i with mass m_i and charge q_i, before its movement is governed by the force due to the electric field. For practical purposes these are usually assumed to be close to zero. Knowing the TOF, one can thus solve (4.2) for the linear momentum $p_{0,i,z}$ in z-direction. The other two components of the linear momentum are given by the conservation of linear momentum. [71, Suppl.]

$$r_{i,x} = \frac{p_{0,i,x}}{m_i} t_{\mathrm{r},i} \tag{4.3}$$

$$r_{i,y} = \left(\frac{p_{0,i,y}}{m_i} - v_{\mathrm{jet}} \right) t_{\mathrm{r},i} \tag{4.4}$$

The initial positions and jet velocity are usually kept as parameters and fitted by data on helium and nitrogen gas. Naturally, this reconstruction ignores any quantum effects that may dominate the Coulomb explosion process directly after the ionization. Therefore one assumes that the handedness of the molecular structure is conserved in the linear momenta of the fragments, which are induced by the Coulomb interaction of the charged fragments. As long as five ions are produced in the experiment, this seems to be a good approximation and from the reconstructed linear momenta the absolute configuration of the molecule is available. One possibility to determine the absolute configuration is to compute the scalar triple product of the linear momenta of Cl, Br and F (see also Fig. 4.2):

$$\cos(\theta) = \frac{\vec{p}_{\mathrm{F}} \cdot (\vec{p}_{\mathrm{Cl}} \times \vec{p}_{\mathrm{Br}})}{|\vec{p}_{\mathrm{F}}| \cdot |\vec{p}_{\mathrm{Cl}} \times \vec{p}_{\mathrm{Br}}|} \tag{4.5}$$

A positive value of $\cos(\theta)$ indicates the (S)-configuration, while a negative value indicates the (R)-configuration.

In the following I discuss briefly three publications, which introduced CEI using a COLTRIMS reaction microscope as valid means to study the absolute configuration of gas phase molecules. I have contributed to the first two publications and am, together with Sabrina Marquardt, main author of the last publication.

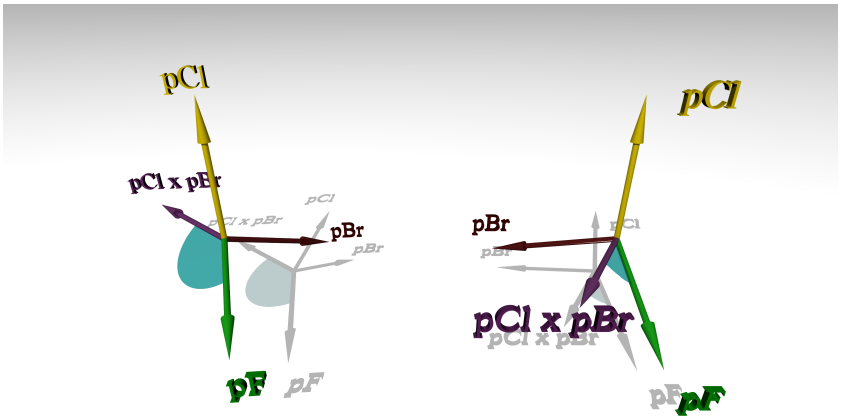


Figure 4.2.: Determination of enantiomers by momentum triple products. The cosine of the angle θ, here shown as blue disk section, is calculated by the triple product of the three momentum vectors $\frac{\vec{p}_F \cdot (\vec{p}_{Cl} \times \vec{p}_{Br})}{|\vec{p}_F| \cdot |\vec{p}_{Cl} \times \vec{p}_{Br}|}$. The left linear momentum arrangement belongs to the Coulomb explosion of the (*R*)-enantiomer, with to $\cos(\theta) < 0$, while the right arrangement belongs to the (*S*)-enantiomer with $\cos(\theta) > 0$. All linear momentum vectors in the picture are normalized to unity for better visualization.

4.2. Direct Determination of Absolute Molecular Stereochemistry in Gas Phase by Coulomb Explosion Imaging (DOI:10.1126/science.1240362)

4.2.1. Summary

In Ref. [71] we used the COLTRIMS technique to simultaneously detect the fragments of laser induced Coulomb explosions of CHBrClF molecules that were brought into gas phase by vaporization of a racemic liquid mixture. We preselected all events where five particles had been detected in coincidence and the linear momentum sum in x-direction for a specific mass and charge configuration was approximately zero. For each of these preselected events we reconstructed the linear momentum of each fragment due to Coulomb explosion. The absolute configuration of the parent molecule was then calculated for each event. We could show that this technique is also applicable to the isotopically chiral molecule $CHBr^{35}_{17}Cl^{37}_{17}Cl$. However, certain orientations of the molecule w.r.t. the detector plane lead to nearly identical TOF for the isotopes $^{35}_{17}Cl$ and $^{37}_{17}Cl$. Distinction of the isotopes was then not possible and the events were discarded.

4.2.2. Contribution

Providing enantiopure or enantiomerically enriched solutions of CHBrClF is a nontrivial but nonetheless important task, as it allows to verify the analysis routines of the COLTRIMS experiment: Let a mixture be composed of n_R molecules of enantiomer type (R) and n_S molecules of type (S). The probabilities to detect an (R)- or (S)-compound in one COLTRIMS experiment are then

$$\begin{aligned} x_R &= \frac{n_R}{N} \\ x_S &= \frac{n_S}{N} \\ N &= n_R + n_S \end{aligned} \tag{4.6}$$

Now let a random error in the experiment or in the analysis of the data lead to false assignment of the chiral designator in half of the cases. Then the observed probabilities x'_I to detect an R- or S-compound are

$$x'_R = \frac{1}{2}\frac{n_R}{N} + \frac{1}{2}\frac{n_S}{N} \tag{4.7}$$

$$= \frac{1}{2}\left(\frac{n_R}{N} + \frac{n_S}{N}\right) \tag{4.8}$$

$$= \frac{1}{2} \tag{4.9}$$

$$= x'_S \tag{4.10}$$

The mixture *appears* to be racemic and this systematic error goes unnoticed. In the general case, where the probability of an erroneous assignment is p_{err}, the probability to detect the R-compound is

$$x'_R = p_{\mathrm{err}} + (1 - 2p_{\mathrm{err}})x_R \tag{4.11}$$

I.e., the detected composition can differ significantly from the true composition.

I designed and implemented a program in the computer algebra system (CAS) MATHEMATICA that represents the CEI experiment as a classical molecular dynamics (CMD) algorithm in combination with a set of trajectory analysis routines. The program simulate the experiment and outputs data that mimics experimental data. These data can be analyzed in the same manner as the experimental outcome. Furthermore, the trajectories can be stored and analyzed to gather information about the Coulomb explosion process.

At the heart of the CMD module lies a Runge–Kutta (RK) algorithm of fourth order (RK4) [145, 146][147, p. 592] that is used to integrate Newton's equations of

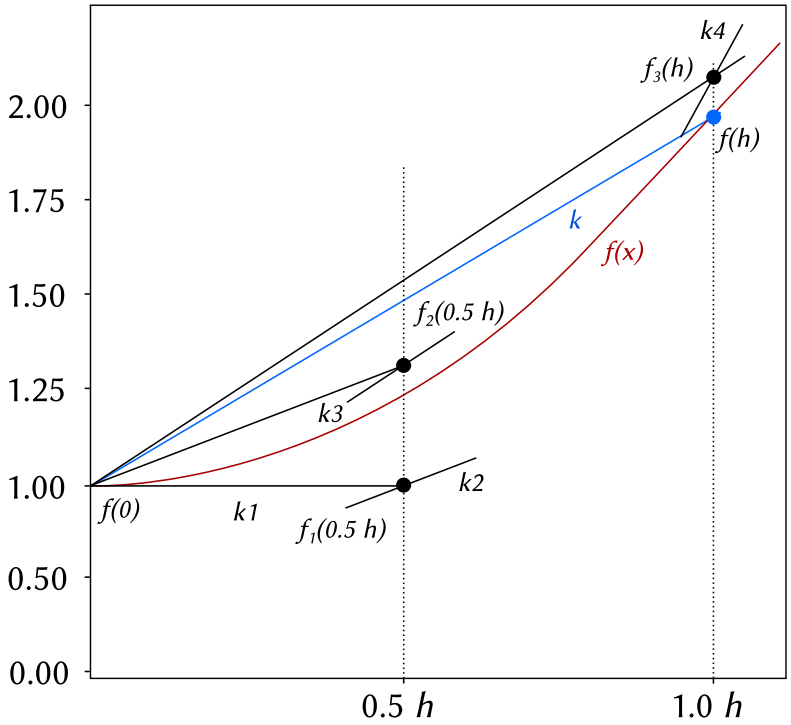
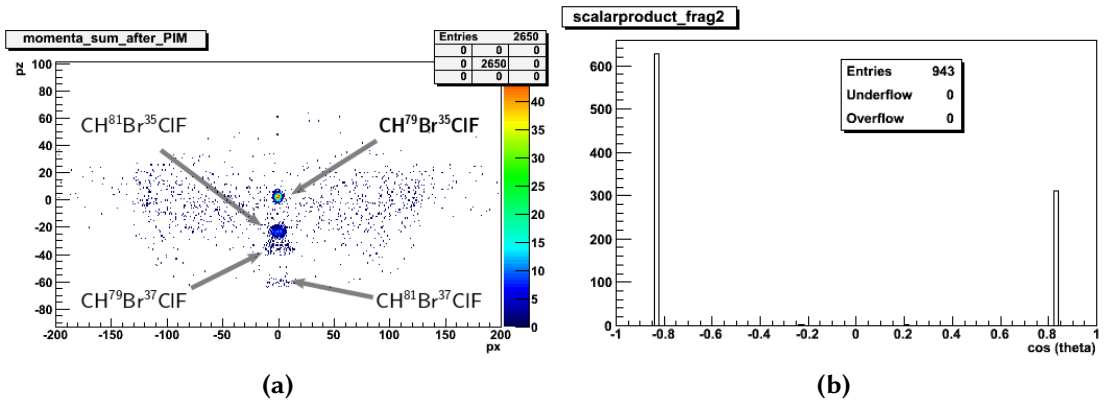


Figure 4.3.: Sketch of a RK4 integration step. Using the initial condition at $x = 0$, the function and its slope are evaluated at $x = 0.5h$ where h is the step width. From the function value and slope at $x = 0$ and $x = 0.5h$ a corrected function value and slope at $x = 0.5h$ are obtained. These numbers are in turn used to determine the function value and slope at $x = h$. The combination of all those values and slopes is then used to determine the new values $f(h)$ and $f'(h)$. The exact scheme is described, e.g., in Ref. [147, p. 592].

motion for all particles in time. The particle positions and velocities at the time $t = 0$ have to be known. In x- and z-direction, the velocity is usually assumed to be zero, while in y-direction the velocity takes the value of the jet-velocity. In each step, the initial conditions (at $t = t_0$) are used to predict the values at $t = t_0 + \Delta t$. In total one needs four function and slope evaluations for the prediction of the new values. A sketch of one integration step of the RK4 scheme is shown in Fig. 4.3. A classical electrostatic force field was used to describe the potentials in which the particles move.

In order to check, if the implementation of the experiment analysis is capable to identify non-racemic mixture correctly, I used the CMD module to simulate a COLTRIMS experiment with CHBrClF and an *ee* of -0.4, which was ensured by randomly inverting the atom coordinates before each simulation such that the proper ratio was obtained. Approximately 2500 simulations were performed with the molecule randomly oriented at the start of each trajectory. Also, the initial speed, isotopic composition and charge distributions were randomly



(a) **(b)**

Figure 4.4.: Momentum analysis and distribution of cos (θ) for simulated CEI data. The raw data was obtained by a classical molecular dynamics simulation and processed by the routines used for experiment analysis.

chosen. The uniform distributions of the molecular orientations relative to the detector were generated by quaternion transformations as described in Ref. [148]. Velocities of each molecule were normal distributed corresponding to a jet temperature of $T = 125\,\mathrm{K}$. Isotopes were selected according to their natural abundance. Furthermore, I allowed for a minor fraction of charge states higher than one to test the analysis routines. Absolute configuration of the molecules was ensured to satisfy a predefined enantiomeric excess. A set of analyzed simulation data is shown in Fig. 4.4. The enantiomeric excess of $ee = -0.4$ was recovered in good agreement. Differences in the observed ee are rooted in the rather large amount of physically improbable data points, caused by the parameter choice for the ensemble generation. E.g., I freely chose a 10 % probability of multiple ionization of the atomic fragments Br/Cl. In an experiment, this probability is likely much lower or the data points are not even recorded due to the spectrometer setup. I concluded that the experiment analysis works as expected.

The study was initiated by Prof. Dr. Robert Berger, Prof. Dr. Horst Schmidt-Böcking and Dr. Markus Schöffler and prospered in subsequent discussions with Prof. Dr. Reinhard Dörner, Prof. Dr. Michael Reggelin and Dr. Martin Pitzer. Dr. Martin Pitzer performed the experiment and its analysis. He, together with Prof. Dr. Robert Berger, Prof. Dr. Reinhard Dörner and Dr. Markus Schöffler were the primarily responsibles for writing the article. I contributed to the article through reviewing the essay and the supporting data presented here.

The permission from *The American Association for the Advancement of Science* to reprint the author's version of the following article M. Pitzer *et al.*, *Science* **2013**, *341*(6150), 1096–1100, DOI:10.1126/science.1240362 is gratefully acknowledged.

Direct determination of absolute molecular stereochemistry in gas phase by Coulomb explosion imaging

Martin Pitzer[1], Maksim Kunitski[1], Allan S. Johnson[1,2], Till Jahnke[1], Hendrik Sann[1], Felix Sturm[1], Lothar Ph. H. Schmidt[1], Horst Schmidt-Böcking[1], Reinhard Dörner[1], Jürgen Stohner[3], Julia Kiedrowski[4], Michael Reggelin[4], Sebastian Marquardt[4], Alexander Schießer[4], Robert Berger[4*], Markus S. Schöffler[1*]

[1]Institute for Nuclear Physics,
Johann Wolfgang Goethe-University Frankfurt,
Max-von-Laue-Str. 1, 60438 Frankfurt, Germany
[2]University of Ottawa, Ottawa, ON, Canada K1N 6N5
[3]Institute of Chemistry and Biological Chemistry,
Zurich University of Applied Sciences, Campus Reidbach,
Einsiedlerstr. 31, 8820 Wädenswil, Switzerland
[4]Clemens-Schöpf Institute,
TU Darmstadt,
Petersenstr. 22, 64287 Darmstadt, Germany

[*]To whom correspondence should be addressed;
E-mail: schoeffler@atom.uni-frankfurt.de,
robert.berger@tu-darmstadt.de

Bijvoet's method, which utilizes anomalous X-ray diffraction or dispersion, is the standard means of direct determination of the absolute (stereochemical) configuration of molecules, but it requires crystalline samples and often proves challenging in structures exclusively comprising light atoms. Herein

we demonstrate a mass spectrometry approach that directly images the absolute configuration of individual molecules in gas phase by cold target recoil ion momentum spectroscopy (COLTRIMS) following laser ionization-induced Coulomb explosion. The technique is applied to the prototypical chiral molecule bromochlorofluoromethane (CHBrClF) and the isotopically chiral methane derivative bromodichloromethane ($CHBr^{37}Cl^{35}Cl$).

A molecule that cannot be superposed with its mirror image by pure translation and rotation is termed chiral, with the non-identical mirror-images denoted as enantiomers. If it were not for the predicted tiny contributions due to parity-violating weak interactions (*1*), which are currently searched for in high-precision molecular physics experiments on chiral molecules (*2, 3*), the energy levels of both enantiomers would be equal. They can be distinguished by their interaction with other chiral objects, for instance left- or right-handed molecules and left- or right-handed circularly polarized photons. The latter type of interaction led Louis Pasteur more than one and a half centuries ago to the first discovery of molecular chirality by observing optical rotation in aqueous solutions of manually separated enantiomorphic crystals from double salts of tartaric acid (*4*). Van't Hoff (*5*) and Le Bel (*6*) independently ascribed Pasteur's observation to an underlying three-dimensional structure of molecules, which can result in two non-identical mirror image structures.

Enantiomers can be distinguished comparatively easily by their chiroptical signals, such as optical rotation, that are of (nearly, due to parity-violation) equal magnitude but opposite sign. This is manifested by the frequently used phenomenological $(+/-)$ terminology. The microscopic structure, on the other hand, is classified with the systematic *R*/*S* or *P*/*M* stereodescriptors (*7*). Assigning the absolute (stereochemical) configuration, however, i.e. establishing which of the two possible mirror-image spatial structural models gives rise to optical rotation with positive or negative sign, still poses a challenge.

The standard approach to directly determine absolute configuration is Bijvoet's method (*8*) of 1951, but after the technique of Coulomb explosion imaging (CEI) had been established (*9*), it was in 2001 pointed out as a potential means to determine the handedness of chiral molecules (*10*). Before 1951 chemical or biochemical conversions were used that relate compounds of unknown configuration to others with known configuration, a method which is still applied today. For this purpose Fischer (*11*) had arbitrarily assigned a given three dimensional structural model (denoted as the D-form) of saccharic acid to the compound that is weakly $(+)$ rotating in aqueous solution and its mirror image (L-form) to the $(-)$ rotating counterpart. Subsequently (see also Ref. (*12*) for the historical development), catalogues of the D and L series could be established, the components of which were chemically related either directly to D or L saccharic acid or to other chiral molecules already filed in the catalogue. Bijvoet finally confirmed (*8*) Fischer's arbitrary choice by studying the sodium rubidium double salt of tartaric acid with anomalous X-ray diffraction. Heavy elemental scatterers induce a pronounced phase and intensity shift when irradiated with X-rays near their absorption edge, which allows the determination of absolute configuration. While typically quite conclusive, Bijvoet's method is limited by requiring crystalline samples. A promising new approach using X-ray diffraction has recently been presented by Inokuma et al. (*13*). They inserted a chiral liquid sample into a crystalline host framework containing heavy atoms. Due to the interaction, the framework's symmetry changed to a chiral spacegroup, thus allowing the application of anomalous X-ray diffraction. Nonetheless, conventional crystallographic challenges and flaws, such as misassigned atoms, symmetry problems and guest disorder, persist and thus require support from mass spectrometry or nuclear magnetic resonance (*13*).

The lack of versatile direct approaches has led to the active exploration of indirect physicochemical approaches based on optical rotation and circular dichroism. Indirect methods rely on quantum chemical calculations or empirical rules to interpret the experimental data. Al-

ternatively, liquid chromatography with an enantioselective stationary phase is a wide-spread method for chiral discrimination. Its drawback for assignment lies in the need for a suitable analogue with known stereochemical configuration. Recent activities focus on nuclear magnetic resonance (NMR) spectroscopy, either by seeking to turn NMR directly into a chiroptical method (*14*) or by exploring possibilities to use residual dipolar couplings in chiral non-racemic alignment media (*15*). Also photo-electron circular dichroism has received renewed interest (*16*) and promising three-wave mixing strategies to obtain chiroptical signals in microwave spectroscopy have been reported recently (*17*).

Herein we focus on direct determination of absolute configuration in the gas phase by a molecular imaging technique that displays the three-dimensional structure of individual chiral molecules on a detector and thereby permits assignment of absolute configuration on a single-molecule basis. Kitamura et al. used a similar approach, but with highly charged argon atoms from an ion source as ionizing agents, to detect dynamical chirality in perdeuterated methane (*10*) and pointed out the possibility of detecting molecular handedness. In 2008, Gagnon et al. employed a related variant to study the structure of achiral dichloromethane (CH_2Cl_2) (*18*) by CEI. For the direct assignment of the absolute configuration high count rates for 4-fold or higher fragmentation coincidence events are required. The laser systems commercially available back in 2008 were unable to produce such rates. To overcome apparative limitations for stereochemical assignments, we combine in this work latest high-power femtosecond lasers of 10 to 100 times higher repetition rates (100 kHz) with improved fast hexanode delay-line detectors to surmount otherwise prohibitively long data acquisition time. In addition, we employ high-performance data recording techniques together with an off-line analysis protocol to cope with the increasing complexity emerging for polyatomic molecules (*19*). These improvements allow to utilize CEI to determine the absolute configuration of the prototypical chiral compound bromochlorofluoromethane (CHBrClF) and for isotopically chiral methane

derivatives in natural abundance such as $CHBr^{37}Cl^{35}Cl$.

$CHBrCl_2$ is commercially available and was used without further purification as impurities are easily discarded in the analysis step of coincidence experiments. Racemic CHBrClF was synthesized as described in Ref. (*20*) by reacting $CHBr_2Cl$ with HgF_2. The spectrometer system employed was described in detail in Ref. (*21*). For the present study the setup was augmented by an assembly of cold traps to recycle the volatile sample compounds.

The approach for direct determination of absolute configuration employs the well established COLd Target Recoil Ion Momentum Spectroscopy (COLTRIMS) as sketched in Fig. S1 in the Supplementary Material. A supersonic gas jet of chiral molecules (y-axis in the laboratory system) crosses either a high-power femtosecond laser (laboratory x-axis) to induce multiple ionization and resulting the Coulomb explosion or a synchrotron radiation beam (results not reported here) or an ion beam. Charged fragments are projected by a static electric field along the laboratory z-axis onto a position and time sensitive multichannel plate detector (MCP) with hexagonal delayline readout (*22*) where all fragments are detected in coincidence (see Fig. S2). From the impact position on the detector (x, y), the known distance between ionization zone and MCP, as well as the measured time-of-flight t, the velocities of all cations (formed in coincidence) can be derived. By assigning masses and charges to the various fragments, corresponding linear momenta of all detected particles can be obtained. In molecular multiple ionization and fragmentation the momenta of photons and electrons are usually negligible compared to the momenta of the Coulomb exploding ionic fragments. Hence, in a cold molecular beam, the sum of the ion momentum vectors has to be close to zero due to momentum conservation. The mass assignment can therefore be confirmed by checking this computed total momentum of all ions. Assignment of absolute configuration is in principle already possible, once four charged fragments (e.g. Br^+, Cl^+, F^+ and CH^+ for CHBrClF) are detected in coincidence. Signatures of those break-ups could be found in the data and used for assign-

ment (results not reported here). Herein, however, we focus on the complete fragmentation into five singly charged ions because analysis and interpretation are more straightforward. Additionally the background can be suppressed quite efficiently. Mass resolution in COLTRIMS is sufficient to distinguish various isotopes of bromine and chlorine in natural abundance, thereby even allowing stereochemical characterization of isotopically chiral molecules.

Further details of the measurement method can be located in the supporting material (SM) (*23*).

Fig. 1 shows the sum of all ion momenta in the case where five fragments of CHBrClF are measured in coincidence. The peak at zero total momentum shows events with the correct mass assignment $^{79}Br^+$, $^{35}Cl^+$, $^{19}F^+$, $^{12}C^+$ and $^{1}H^+$. These events were used for the determination of the absolute configuration. The peaks at lower momentum in z-direction are contributions from other isotopologues.

Fig. 2 demonstrates the capability to distinguish enantiomers in our racemic sample. For this purpose, an angle θ is defined, indicating if the momenta of bromine, chlorine and fluorine form a right-handed or a left-handed coordinate system. As distinct peaks are obtained in the histogram, almost all events can be assigned clearly to one enantiomer or the other. This shows the robustness of our method against the laser pulse length: Being about $40\,\mathrm{fs}$ long, our pulses are not short enough to consider the hydrogen frozen during multiple ionization. The clear separation of enantiomers in the histogram indicates, however, that the motion of hydrogen during the laser pulse does not alter bond angles to an extent that would prevent identification of enantiomers. These results also show that reconstruction of the exact geometric structure is not necessary for the determination of absolute configuration. Classical molecular dynamics simulations confirm that the enantiomers are mapped unambiguously onto their momentum space analogues that are presented here.

Fig. 3 overlays measured linear momenta on a rigid structural model of neutral CHBrClF,

with the linear momentum of carbon fixed along the x-axis, the momentum sum of chlorine and bromine defining the x-y-plane and all other linear momenta being oriented relative to these. For better visibility, momenta are normalized with respect to the carbon momentum. The momentum of hydrogen, being very small due to the low mass, is expanded by a factor of two in this figure. It is evident that the configuration of the two enantiomers is directly imaged on the detector. As a racemic mixture of CHBrClF was employed, an equal ratio for *S*- and *R*-configuration was obtained within the statistical uncertainty (329 and 302 events for *S* and *R*, respectively, with $|\cos(\theta)| > 0.6$ in each case). At first sight one might be surprized that the carbon ion is detected in the same direction as the proton. This is due to the fact that the position of the center of mass is conserved and thus H^+ and C^+, as the lightest two ions of the system, are repelled from the slowly moving heavier ions, as has been confirmed with the help of molecular dynamics simulations.

For CHBrClF the direct assignment of absolute configuration works unequivocally in the majority of fivefold coincidence events due to the comparatively large mass difference between the ions. The situation is considerably more challenging for the case of isotopically chiral systems. In the case of $CHBrCl_2$ (see Fig. 4), not all five-fold fragmentation events allow an unambiguous assignment of the isotope masses to each fragment and hence no determination of absolute configuration is possible for such events. However, a subset of events that permit a conclusive assignment can be selected by a procedure detailed in the SM (Figs. S3-S4 and text). Again, as a racemic sample was used, an almost 50:50 ratio of *S*:*R* is detected (282:273 events with $|\cos(\theta)| > 0.6$, see Fig. S5).

The technique still has several limitations: Volatile molecules and a large amount of substance are required due to the molecular beam source. Comparatively simple, rigid structures were studied. In the case of more complicated molecules, the kinematic properties of the fragments may not directly illustrate the geometric structure, making the identification of the abso-

lute configuration less straightforward. In this case, geometrical reconstruction or comparison with simple molecular dynamics simulations might become necessary, especially when several stereogenic elements are present. On a technical level, the probability of multi-fragment detection decreases dramatically with the amount of fragments, as both the fragmentation yield and the detection efficiency diminish exponentially with the number of fragments.

Stepwise fragmentation is a limitation as well. Additionally, when multi-ionization is slow compared to vibrational time scales, assignment can be hampered or even completely prohibited. For this purpose, faster ionization schemes with shorter laser pulses and higher laser intensity are required.

In conclusion, the present imaging approach allows for determination of absolute configuration of gas phase molecules on a per-molecule basis. Apart from a rigid structural model, it does not require theoretical input. Besides structure determination as demonstrated herein, the coincidence technique creates unique opportunities to study chirality in single molecules.

References and Notes

1. R. Berger, *Relativistic Electronic Structure Theory, Part: 2, Applications*, P. Schwerdtfeger, ed. (Elsevier, Netherlands, 2004), pp. 188–288.

2. C. Daussy, *et al.*, *Phys. Rev. Lett.* **83**, 1554 (1999).

3. M. Quack, J. Stohner, M. Willeke, *Annu. Rev. Phys. Chem.* **59**, 741 (2008).

4. L. Pasteur, *Leçons de chimie professées en 1860 par MM. Pasteur, Cahours, Wurtz, Berthelot, Sainte-Claire Devile, Barral et Dumas* (Hachette, Paris, 1861), pp. 1–48.

5. J. H. van't Hoff, *Archives neerlandaises des sciences exactes et naturelles* **9**, 445 (1874).

6. J.-A. Le Bel, *Bull. Soc. Chim. Paris* **T22**, 337 (1874).

7. R. S. Cahn, C. Ingold, V. Prelog, *Angew. Chem. Int. Ed.* **5**, 385 (1966).

8. J. M. Bijvoet, A. F. Peerdeman, A. J. V. Bommel, *Nature* **168**, 271 (1951).

9. Z. Vager, R. Naaman, E. P. Kanter, *Science* **244**, 426 (1989).

10. T. Kitamura, T. Nishide, H. Shiromaru, Y. Achiba, N. Kobayashi, *J. Chem. Phys.* **115**, 5 (2001).

11. E. Fischer, *Ber. Dtsch. Chem. Ges.* **24**, 2683 (1891).

12. F. W. Lichtenthaler, *Angew. Chem. Int. Ed.* **31**, 1541 (1992).

13. Y. Inokuma, *et al.*, *Nature* **495**, 461 (2013).

14. A. D. Buckingham, *Chem. Phys. Lett.* **398**, 1 (2004).

15. R. Berger, *et al.*, *Angew. Chem. Int. Ed* **51**, 8388 (2012).

16. C. Lux, *et al.*, *Angew. Chem. Int. Ed.* **51**, 5001 (2012).

17. D. Patterson, M. Schnell, J. M. Doyle, *Nature* **497**, 475 (2013).

18. J. Gagnon, K. F. Lee, D. M. Rayner, P. B. Corkum, V. R. Bhardwaj, *J. Phys. B* **41**, 215104 (2008).

19. B. Wales, *et al.*, *Nuclear Instruments and Methods in Physics Research A* **667**, 11-15 (2012).

20. J. Hine, A. M. Dowell, J. E. Singley, *J. Am. Chem. Soc.* **78**, 479 (1956).

21. J. Ullrich, *et al.*, *Rep. Prog. Phys.* **66**, 1463 (2003).

22. O. Jagutzki, *et al.*, *IEEE Transact. on Nucl. Science* **49**, 2477 (2002).

23. Further details on the measurement principle with six additional figures and four movies are available in the supporting material.

24. A. S. Alnaser, *et al.*, *Phys. Rev. A* **70**, 023413 (2004).

Acknowledgements

The authors are indebted to David Avnir, Timur Isaev and Frieder Lichtenthaler for discussion. We acknowledge help by Moritz Meckel with some figures and support by Achim Czasch concerning data analysis. This work was supported by the State Initiative for the Development of Scientific and Economic Excellence (LOEWE) in the LOEWE-Focus ELCH. The momentum data used to draw the figures are provided in the SM. Raw data are archived at the University of Frankfurt (affiliation 1) and are available upon request.

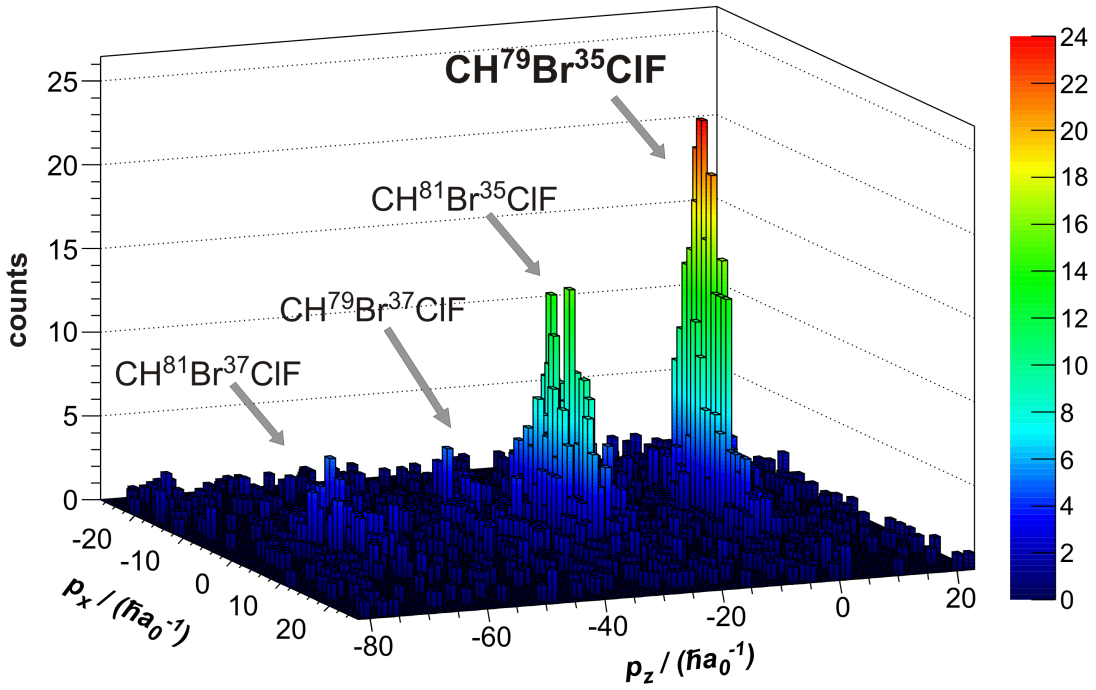


Fig. 1. Sum of linear momentum components in five-fold ionization of bromochlorofluoromethane, demonstrating the excellent resolution achieved in this experiment. Those events detected with p_x and p_z close to zero correspond to a fragment assignment to $^{79}Br^+$, $^{35}Cl^+$, $^{19}F^+$, $^{12}C^+$ and $^{1}H^+$. Fragments of other isotopologues can also be identified, but are not used in the present analysis. The atomic unit for momentum is defined as $\hbar\, a_0^{-1} \approx 1.992 \cdot 10^{-24}\,\mathrm{kg\,m\,s^{-1}}$.

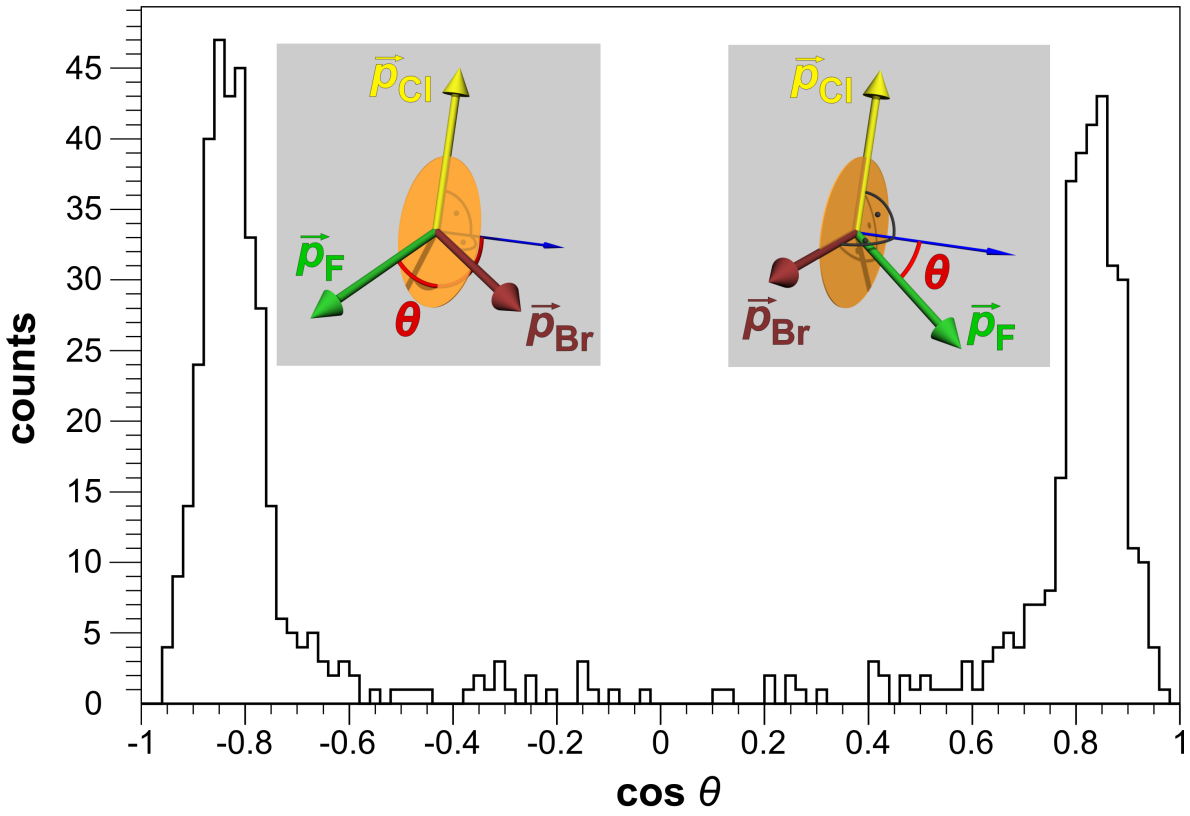


Fig. 2. Chiral discrimination for CHBrClF. The histogram for the cosine of the chirality angle θ shows the clear separation between enantiomers in our racemic sample of CHBrClF. As is illustrated in the inset, the angle is defined via $\cos(\theta) = \vec{p}_{\mathrm{F}} \cdot (\vec{p}_{\mathrm{Cl}} \times \vec{p}_{\mathrm{Br}}) \, (|\vec{p}_{\mathrm{F}}||\vec{p}_{\mathrm{Cl}} \times \vec{p}_{\mathrm{Br}}|)^{-1}$. The peak at negative values of $\cos(\theta)$ corresponds to the *S*-enantiomer, while the one at positive values corresponds to the *R*-enantiomer.

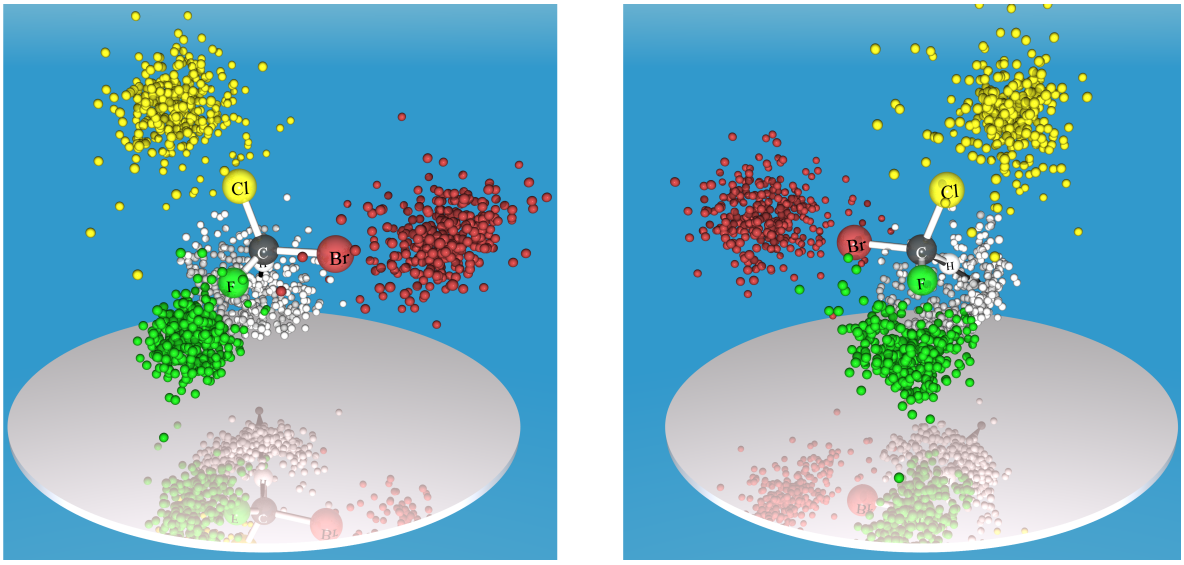


Fig. 3. Linear momenta in five-fold fragmentation of bromochlorofluoromethan enantiomers. Measured linear momenta in the five-fold fragmentation of (S)-CHBrClF (left, $\cos(\theta) < -0.6$) and (R)-CHBrClF (right, $\cos(\theta) > 0.6$) are indicated in the color codes C: black arrow, H: grey, F: green, Cl: yellow, Br: red. Momenta are rotated to the molecular frame of reference, defined by the momentum of the carbon ion and the momentum sum of bromine and chlorine ions. All momenta are normalized with respect to the carbon momentum. For better visibility, hydrogen momenta are expanded by a factor of 2. Whereas the substituents are expelled during the Coulomb explosion into the directions expected from the classical structural model, the central atom C is also accelerated away from the centre of mass and ejected in a similar direction as hydrogen. An animated version of the figure can be found in the SM (movies S1-S2).

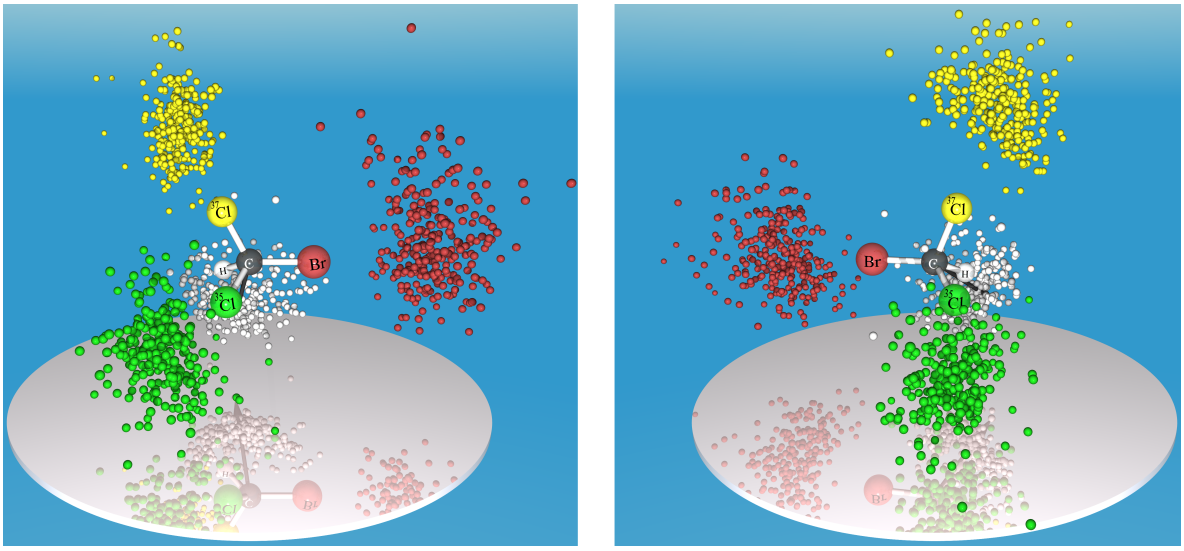


Fig. 4. Measured linear momenta in the five-fold fragmentation of (S)-$CH^{79}Br^{37}Cl^{35}Cl$ (left) and (R)-$CH^{79}Br^{37}Cl^{35}Cl$ (right). The color codes used correspond to C: black arrow, H: light grey, ^{35}Cl: green, ^{37}Cl: yellow, Br: red. Momenta are rotated to the molecular frame of reference, defined by the momentum of the carbon ion and the momentum sum of ^{79}Br and ^{35}Cl, and normalized with respect to the momentum of the carbon ion. Again, the hydrogen momentum is scaled by a factor of 2. Histogram and animated version of the figure can be found in the SM (Fig. S5, movies S3-S4). For a wider angle of vision, also a few outlier signals become clearly visible (Fig S7).

Supplementary content

Description of the working principle of the coincidence experiment with two additional figures (S1 and S2), details on the measurement, details on the calibration, procedure for selecting break-ups in isotopically chiral systems with two additional figures (S3 and S4), histogram for $CH^{79}Br^{37}Cl^{35}Cl$ break-ups, wider angle snapshots for figures 3 and 4 (S6 and S7), four movie files with animated versions of figures 3 and 4, two databases with the linear momentum data, Reference *(24)*.

4.3. Absolute Configuration from Different Multifragmentation Pathways in Light-Induced Coulomb Explosion Imaging (DOI:10.1002/cphc.201501118)

4.3.1. Summary

Ref. [72] describes the determination of absolute configuration through synchrotron X-ray photon induced Coulomb explosion of the chiral molecule CHBrClF. The event count rose significantly and the momentum distributions were sharper than those of Ref. [71], presumably due to a faster ionization and fragmentation process. We showed that enantiomeric distinction by the previously established method is still possible, if one of the fragments consists of more than one atom. However, the distinction uncertainty depends on the composition of the molecular fragment. For the CH^+ fragment we obtained a good enantiomeric distinction while the corresponding lines for CF^+ events were very close. Moreover, the line separation depends on the choice of the fragments used for the calculation of the scalar triple product and thus rendered the analysis more involved. The analysis of four fragments is then extended to an analysis of four (three) atoms, where one (two) atoms are not detected. It is shown that as long as the event produces at least four fragments, enantiomeric distinction is still possible. Three fragments alone are insufficient as their linear momenta must be coplanar due to total linear momentum conservation. Furthermore, reconstruction of the fifth particle's linear momentum from four momentum data points gives insight into the underlying fragmentation processes: E.g., the momenta of C^+, H^+, Br^+ and Cl^+ were used to reconstruct the momentum of F^+, showing a bimodal distribution. The linear momenta of $\approx$ 200 a.u. were attributed to the direct break-up into five particles where F^+ is not detected. The peak around 100 a.u. however is attributed to different fragmentation processes as, e.g., dissociation of a molecular fragment after the initial Coulomb explosion of the parent molecule. Calculation of all triple products available from four fragments allows cross-checks of the enantiomeric distinction. In the case of three detected particles, the cross-checks are not possible anymore. Furthermore, the background is significantly higher than that of the four-particle events. But still, a large number of events can be analyzed w.r.t. their absolute configuration.

4.3.2. Contribution

For this publication I provided quantum mechanical data from molecular structure optimizations performed with TURBOMOLE [95], employing DFT with the B3-LYP [11, 149–153] functional and cc-pVTZ-F12 basis on all atoms and a Stuttgart type PP on Br. The data showed that there exist metastable triply charged planar configurations of $CHBrClF^{3+\bullet}$ which could allow for an inversion of absolute configuration during Coulomb explosion. Furthermore, the simulation program written for Ref. [71] was improved by the implementation of harmonic and anharmonic bonding potentials. These can be used to simulate partial fragmentations as they can occur in quantum chemically based Born–Oppenheimer molecular dynamics (BOMD) simulations. The anharmonic bonding potential was approximated as a Morse potential. Also, we included the charge exchange model developed by Eland and Sheahan in Ref. [154], which minimizes the total electrostatic energy of the exploding system, consisting of the Coulomb interaction contribution and an ionization contribution. The latter is calculated for each atom j by the equation

$$I_j = qu_j + q^2 v, \tag{4.12}$$

where the first ionization potential is used as the work function u_j. v is a measure for the capacitative energy and was taken as 8 eV, as stated in the original publication. This scheme applies as long as all fragments lie within a sphere of 7 Å diameter. Outside of this sphere, the charges are fixed.

The initial draft of the work was provided by Dr. Martin Pitzer and significantly revised by Sabrina Marquardt and me. We both contributed in equal parts. The idea of this worked stemmed from discussions of Prof. Dr. Robert Berger, Prof. Dr. Reinhard Dörner, Prof. Dr. Horst Schmidt-Böcking, Dr. Markus Schöffler and Dr. Martin Pitzer. Experimental data was provided by the group of Prof. Dr. Reinhard Dörner. All authors of the publication contributed to the discussion of the experimental results.

The permission from *John Wiley and Sons* by license number 4355041341404 to reprint the following article M. Pitzer *et al.*, *ChemPhysChem* **2016**, *17*, 2465–2472, DOI:10.1002/cphc.201501118 is gratefully acknowledged.

ChemPubSoc Europe

DOI: 10.1002/cphc.201501118

CHEMPHYSCHEM Articles

Absolute Configuration from Different Multifragmentation Pathways in Light-Induced Coulomb Explosion Imaging

Martin Pitzer,[a] Gregor Kastirke,[a] Maksim Kunitski,[a] Till Jahnke,[a] Tobias Bauer,[a] Christoph Goihl,[a] Florian Trinter,[a] Carl Schober,[a] Kevin Henrichs,[a] Jasper Becht,[a] Stefan Zeller,[a] Helena Gassert,[a] Markus Waitz,[a] Andreas Kuhlins,[a] Hendrik Sann,[a] Felix Sturm,[a] Florian Wiegandt,[a] Robert Wallauer,[a] Lothar Ph. H. Schmidt,[a] Allan S. Johnson,[b] Manuel Mazenauer,[c] Benjamin Spenger,[c] Sabrina Marquardt,[d] Sebastian Marquardt,[d] Horst Schmidt-Böcking,[a] Jürgen Stohner,[c] Reinhard Dörner,[a] Markus Schöffler,*[a] and Robert Berger*[d]

The absolute configuration of individual small molecules in the gas phase can be determined directly by light-induced Coulomb explosion imaging (CEI). Herein, this approach is demonstrated for ionization with a single X-ray photon from a synchrotron light source, leading to enhanced efficiency and faster fragmentation as compared to previous experiments with a femtosecond laser. In addition, it is shown that even incomplete fragmentation pathways of individual molecules from a racemic CHBrClF sample can give access to the absolute configuration in CEI. This leads to a significant increase of the applicability of the method as compared to the previously reported complete break-up into atomic ions and can pave the way for routine stereochemical analysis of larger chiral molecules by light-induced CEI.

1. Introduction

Determination of the absolute configuration of a chiral species is a keystone of stereochemistry. The direct assignment of absolute configuration in crystalline samples is routinely performed with Bijvoet's method, which employs anomalous diffraction of X-rays.[1] For a vast range of substances that cannot easily be crystalized, determination of absolute configuration requires knowledge of the handedness of related compounds or extensive theoretical models. Recently, two research teams[2,3] have shown independently the use of Coulomb explosion imaging (CEI)[4] for the direct determination of absolute configuration of small chiral molecules in the gas phase.

When the bonding electrons are removed from a molecule instantaneously, the resulting atomic cations, which essentially remain in their equilibrium positions during ionization, experience strongly repelling Coulombic forces. The molecule subsequently undergoes a so-called Coulomb explosion. By performing a coincident measurement of the fragment ions' linear momenta, the microscopic structure of the molecule is imaged on the macroscopic detection device. For very small systems, Coulomb explosion imaging is capable of visualizing even fine details such as the nodal structure of the wavefunction[5] or the tunneling part of loosely bound helium molecules.[6,7]

First investigations on the structure of the CH_4^+ ion were already carried out in 1986 with foil-induced Coulomb explosion imaging.[8] In this approach, the molecular ions are accelerated and sent through a thin foil, stripping off the electrons. Kitamura et al.[9] employed highly charged argon atoms to multiply ionize perdeuterated methane CD_4 to demonstrate the dynamic chirality of this molecule. Herwig et al.[3,10] used foil-induced Coulomb explosion imaging to determine the absolute configuration of the monocation of 1,2-dideutero-oxirane. This approach, however, is limited to relatively small masses. Pitzer et al.[2] have employed a high-power Ti:sapphire femtosecond laser system with high repetition rate for the multiple ionization of CHBrClF (from a racemic mixture) and Cold Target Recoil Ion Momentum Spectroscopy (COLTRIMS)[11] for the detection of fragments. The results show that unambiguous de-

[a] Dr. M. Pitzer, G. Kastirke, Dr. M. Kunitski, Dr. T. Jahnke, Dr. T. Bauer, C. Goihl, F. Trinter, C. Schober, K. Henrichs, J. Becht, S. Zeller, H. Gassert, M. Waitz, A. Kuhlins, H. Sann, F. Sturm, F. Wiegandt, Dr. R. Wallauer, Dr. L. Ph. H. Schmidt, Prof. Dr. H. Schmidt-Böcking, Prof. Dr. R. Dörner, Dr. M. Schöffler
Institute for Nuclear Physics
Johann-Wolfgang-Goethe-Universität Frankfurt
Max-von-Laue-Straße 1, 60438 Frankfurt (Germany)
E-mail: schoeffler@atom.uni-frankfurt.de

[b] A. S. Johnson
University of Ottawa
Ottawa, ON, K1N 6N5 (Canada)

[c] M. Mazenauer, B. Spenger, Prof. Dr. J. Stohner
Institute of Chemistry and Biological Chemistry
Zurich University of Applied Sciences, Campus Reidbach
Einsiedlerstrasse 31, 8820 Wädenswil (Switzerland)

[d] S. Marquardt, S. Marquardt, Prof. Dr. R. Berger
Fachbereich Chemie
Philipps-Universität Marburg
Hans-Meerwein-Straße, 35032 Marburg (Germany)
E-mail: robert.berger@uni-marburg.de

Supporting Information for this article can be found under http://dx.doi.org/10.1002/cphc.201501118.

termination of absolute configuration is possible on the level of individual molecules from a racemic mixture. Limitations of laser-induced Coulomb explosion imaging for larger covalently bound systems have been discussed in Ref. [2]. Amongst those is the demand for a comparatively large amount of sample and the timescale of the laser-induced ionization. The latter leads in some cases, for example, for fast moving atoms such as hydrogen, to a broadening of the measured linear momentum distribution.

In this work, we show that a single X-ray photon (with energy $h\nu = 710$ eV) from a synchrotron light source can also be used to induce Coulomb explosion of CHBrClF into up to five atomic ions. First, we discuss the mechanisms contributing to the multiple ionization of the target molecule. We then demonstrate on the level of individual molecules from a racemic mixture the determination of absolute configuration for the complete fragmentation into five singly charged ions and compare the results to those previously obtained in laser experiments. This is followed by an investigation of partial break-ups of the target molecule, that is, break-ups including molecular ions like CH^+. Finally, the advantages and drawbacks of using events with an incomplete detection of the fragments are discussed.

2. Results and Discussion

2.1. Multiple Ionization Induced by a Single X-ray Photon

Once an electron is removed from an inner shell of an atom or a molecule, the excited system will relax to its ground state. An important relaxation process is the Auger decay, during which the excitation energy is transferred to another electron that can escape into the continuum, leaving a doubly charged atom or molecule behind.

With sufficient energy this process can occur in a cascade, leading to multiple ionization. Using photon energies slightly above the ionization threshold of the F(1s) state ($E_B = 688$ eV for atomic fluorine),[12] we observed fragmentation of CHBrClF into up to five singly charged ions.

In order to shed light on the ionization processes involved, we compared our findings to the extensive literature on multiple ionization of the rare-gas atoms from the same row of the periodic table as the halogen atoms in our sample, that is, neon (corresponding to fluorine), argon (corresponding to chlorine) and krypton (corresponding to bromine). All three halogens have energy levels that contribute to the photoionization cross section at the chosen photon energy of 710 eV (see Table 1 in the Supporting Information). As efficient Auger cascades leading to five-fold ionization have been reported for ionization of the krypton 3p or 3d states,[13,14] we expect excitation of the respective bromine state to be the prominent channel. For the four-fold ionization, the chlorine 2s state is expected to play an important role as well.[15,14] Recently, ultrafast charge rearrangement in a similar molecule, CH_3SeH, upon excitation of the Se L-shell by an X-ray free-electron laser has proven to be highly efficient.[16] A further efficient channel for additional charge removal besides Auger decay is the direct photoemission of electron pairs by shake-off or knock-off, known in atoms[17] and molecules.[18] The interplay between the different possibilities for photoionization and the various decay channels for the respective core holes makes it impossible to identify the exact ionization mechanism for single events.

Compared to the previous laser experiment, efficiency is increased significantly. Whereas the fraction of the five-fold fragmentation on the number of total events in the laser experiment was around 5×10^{-6}, it was increased in the synchrotron experiment by more than one magnitude to 7×10^{-5}.

2.2. Complete Fragmentation and Determination of Absolute Configuration

In analogy to our previous work,[2] we first investigate the break-up of CHBrClF into five singly charged ions. The occurrence of this fragmentation pathway can already be seen in the raw photo-ion coincidence spectra (Figure 1). In this graph, sums of the times-of-flight of several particles are plotted; due to linear momentum conservation, ions originating from the same molecule arrange on a line in this plot.

For details on the extraction of the ion momenta from the experimental data, we refer to the Experimental Section. By checking whether linear momentum conservation is fulfilled for a certain assignment, the correct isotopic mass is identified. Furthermore, this test allows to eliminate false coincidences and thus increase the signal-to-noise-ratio. For the remaining events, several triple products can be calculated from the linear momentum vectors of the respective three ions. In Figure 2a, we show the chirality parameter $\cos\theta_{F\cdot(Cl\times Br)} = \vec{p}_F \cdot (\vec{p}_{Cl} \times \vec{p}_{Br}) \cdot (|\vec{p}_F||\vec{p}_{Cl} \times \vec{p}_{Br}|)^{-1}$, abbreviated as F·(Cl×Br) in the following. A negative value of F·(Cl×Br) corre-

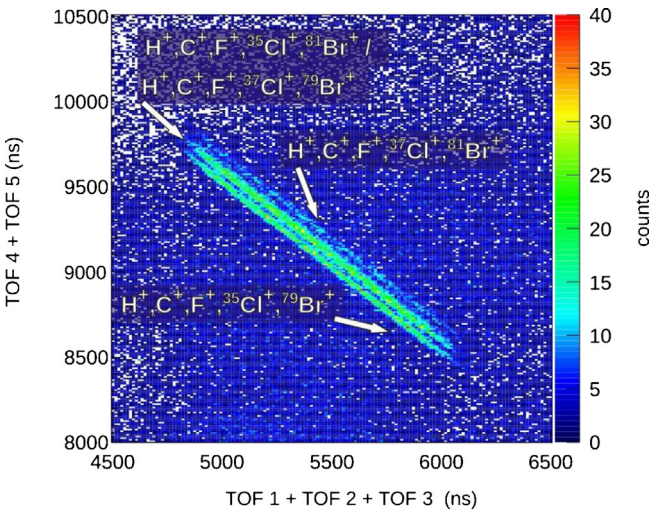


Figure 1. Coincidence spectrum of five measured ions. On the *x*-axis, the sum of the times-of-flight for the first three detected ions is plotted, on the *y*-axis the sum of the times-of-flight of the fourth and fifth ion. Ions from a Coulomb explosion yield a line in this representation due to linear momentum conservation. The substructure originates from the different isotopes of chlorine and bromine.

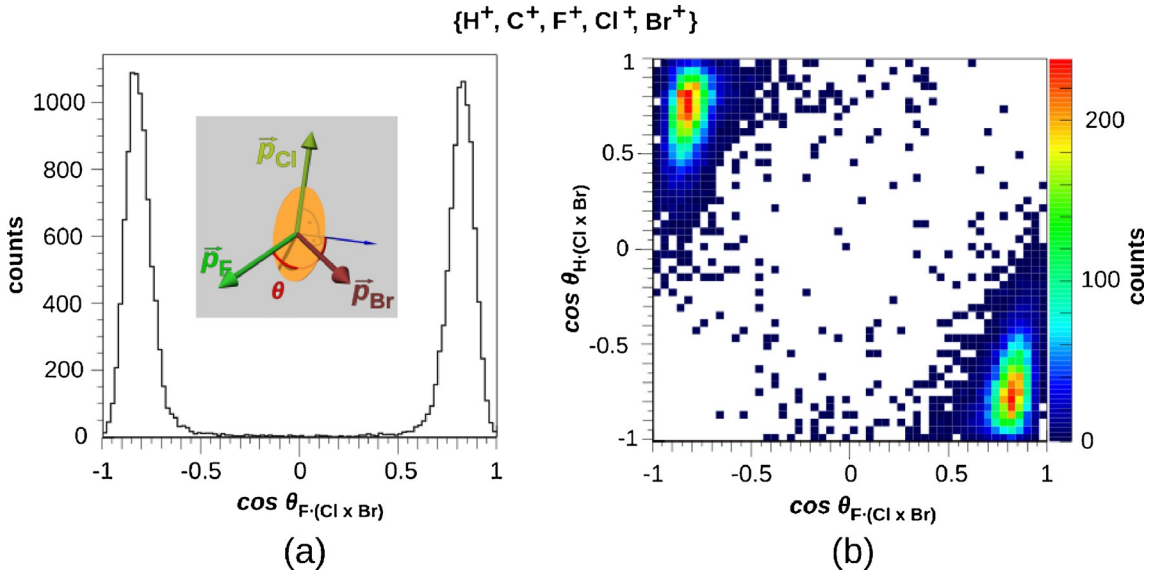


Figure 2. Chirality parameter cos θ for the fragmentation of CHBrClF into five singly charged atomic ions. The geometric definition of the angle θ is given in the inset. The two peaks in (a) indicate a clear separation of enantiomers. The correlation diagram (b) of the two triple products F·(Cl×Br) and H·(Cl×Br) demonstrates the consistency for almost all events.

sponds to an *S*-type arrangement of the momenta in the Cahn–Ingold–Prelog nomenclature[19] whereas a positive value corresponds to an *R*-type arrangement.

Figure 2b shows a correlation diagram of the triple products F·(Cl×Br) and H·(Cl×Br). The position of the peaks on the negative diagonal demonstrates that the assignment of absolute configuration is consistent for the proton and the fluorine ion: For events where a positive value of F·(Cl×Br) indicates an *R*-type molecule, the corresponding value for H·(Cl×Br) is negative. Proton migration to the other side of the carbon atom during the ionization process can thus be ruled out.

In our previous experiments on laser ionization, the linear momentum distribution of the light proton nevertheless proved to be rather broad. Figure 3 compares the triple product H·(Cl×Br) of the synchrotron experiment with the data set obtained with the laser. Despite the small number of events in the latter case, an improvement can clearly be seen. We attribute this to the typical Auger decay times being significantly below 10 fs whereas the laser pulse width was determined as 40 fs.

The three-dimensional representation of the linear momenta in the molecular frame of reference (Figure 4) confirms the

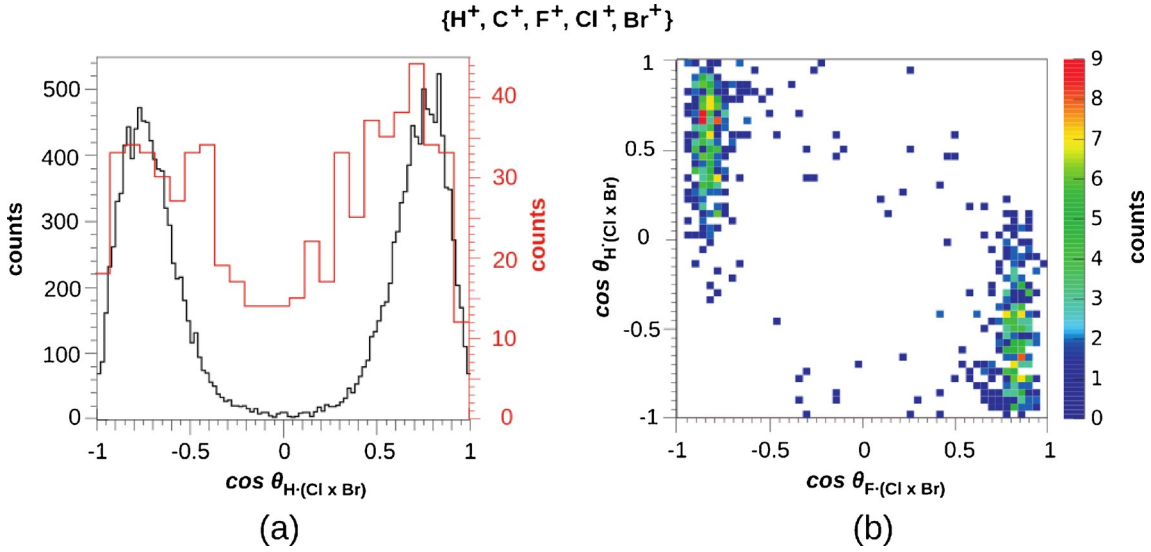


Figure 3. a) Comparison of the triple product H·(Cl×Br) for the laser experiment,[2] shown in red and the synchrotron experiment (black). The smaller spread in the synchrotron experiment is attributed to a faster ionization process, preventing broadening of the distribution due to proton motion. b) Correlation diagram for the laser ionization, in analogy to Figure 2(b), shows that this broadening does not lead to ambiguities in the assignment of enantiomers.

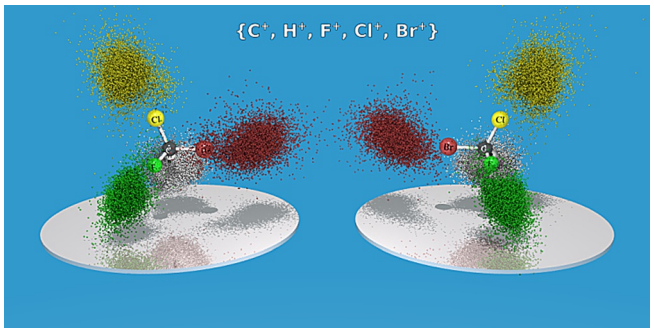


Figure 4. Three-dimensional representation of the linear momenta in the molecular frame for the fragmentation into five atomic ions, overlayed with a structure model. Color codes are white: H; black: C; green: F; yellow: Cl; red: Br. Transformation into the molecular frame is described in the Supporting Information.

good separation of enantiomers, suggesting an unambiguous assignment of absolute configuration on the level of individual molecules. The transformation into the molecular frame is described in the Supporting Information.

2.3. Partial Fragmentation into Molecular Ions

As can be seen from the definition of the chirality parameter, only three linearly independent momentum vectors are needed to determine the handedness of the system. It follows that the complete break-up of the molecule into atomic ions is not a prerequisite for the assignment of absolute configuration. Fragmentation pathways including molecular ions can thus be used to achieve this goal. This fact is of particular interest for the study of larger molecules where complete fragmentation cannot be expected.

In the case of CHBrClF, it is an obvious choice to investigate break-ups including a CX^+-fragment with X being H, F, Cl or Br. From these four possibilities, the fragmentation pathways with the ions $\{CF^+, H^+, Cl^+, Br^+\}$ (channel **II**) and $\{CH^+, F^+, Cl^+, Br^+\}$ (channel **III**) were found in the recorded data.

Figure 5 shows the triple products for different break-ups. While the assignment of absolute configuration works well with CH^+, the fragmentation channel **II** containing CF^+ barely allows to distinguish enantiomers. The graphs for **II** show that $\cos\theta_{H\cdot(Cl\times Br)}$ has a broad distribution and that $\cos\theta_{CF\cdot(Cl\times Br)}$ has two overlapping peaks close to 0, indicating that these three linear momenta are nearly coplanar due to the small linear momentum carried away by the proton.

As the detection probability for the cations is significantly lower than 100% (see Experimental Section), the detection efficiency is expected to be higher for the coincident measurement of four ions instead of five. Table 1, however, demonstrates that the yield in the case of the partial fragmentation

Table 1. Total ion yields for different fragmentation channels in the synchrotron experiment. Reaction products in square brackets remained undetected (see Section 3.4 for details). Whereas fragmentation into molecular ions does not increase the yield (channels II and III) significantly, the incompletely detected break-ups show a significantly higher efficiency. For channel III, an overlap with the pathway $[H]^{+,0}+\{C^+, F^+, Cl^+, Br^+\}$ cannot be eliminated completely.

	Fragmentation pathway	Yield (events)
I	$CHBrClF^{5+}\rightarrow\{C^+, H^+, F^+, Cl^+, Br^+\}$	4.7×10^4
II	$CHBrClF^{4+}\rightarrow\{CF^+, H^+, Cl^+, Br^+\}$	2.0×10^4
III	$CHBrClF^{4+}\rightarrow\{CH^+, F^+, Cl^+, Br^+\}$	1.1×10^5
IV	$CHBrClF^{5+,4+}\rightarrow[Br]^{+,0}+\{C^+, H^+, F^+, Cl^+\}$	5.9×10^5
V	$CHBrClF^{5+,4+}\rightarrow[F]^{+,0}+\{C^+, H^+, Cl^+, Br^+\}$	3.4×10^5
VI	$CHBrClF^{5+,4+}\rightarrow[C]^{+,0}+\{H^+, F^+, Cl^+, Br^+\}$	3.6×10^5
VII	$CHBrClF^{5+,4+,3+}\rightarrow[H,Br]^{2+,+,0}+\{C^+, F^+, Cl^+\}$	2.2×10^6
VIII	$CHBrClF^{5+,4+,3+}\rightarrow[H,Cl]^{2+,+,0}+\{C^+, F^+, Br^+\}$	1.1×10^6
IX	$CHBrClF^{5+,4+,3+}\rightarrow[C,F]^{2+,+,0}+\{H^+, Cl^+, Br^+\}$	1.5×10^6

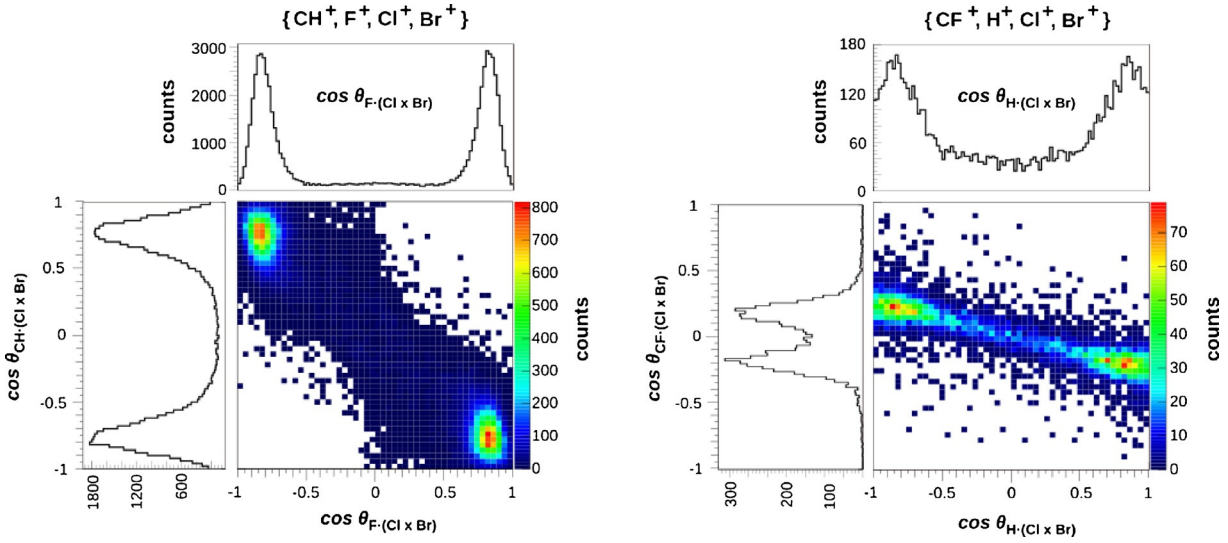


Figure 5. Triple products for fragmentation pathways including molecular ions, in analogy to Figure 2. The fragmentation into $\{CH^+, F^+, Cl^+, Br^+\}$ promises a good separation of enantiomers (left), despite a small background (see text). For the break-up into $\{CF^+, H^+, Cl^+, Br^+\}$, the separation is less clear, as the fragments CF^+, Cl^+ and Br^+ are nearly coplanar (right).

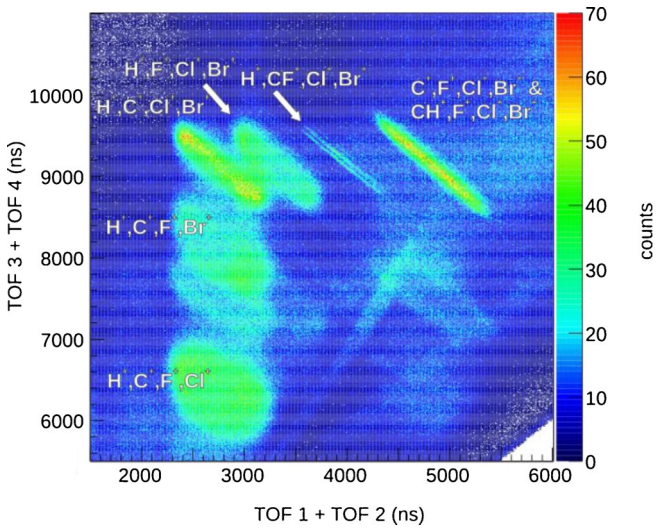


Figure 6. Coincidence plot for the identification of fragmentation channels. On the *x*-axis, the sum of the time-of-flight of the first and second ion is plotted; the *y*-axis displays the corresponding sum of the third and fourth ion. Fragmentation channels in which the detected ions have zero-sum momentum show up as narrow lines. Fragmentations with a neutral product are spread out in the plot due to the neutral particle's linear momentum. Different fragmentation pathways are identified from the ions' times-of-flight.

into four ions is not higher than in the case of complete fragmentation. Similar results for partial break-ups were extracted from the laser measurement,[2] and are shown in the Supporting Information.

2.4. Incomplete Detection of Ionic Fragments

Looking at the raw four-ion coincidence spectra (Figure 6), various features can be identified that correspond to fragmentation pathways with one of the fragment masses missing. Two possibilities contribute to this scenario: either a neutral dissociation product is involved or a charged fragment was not detected due to the limited detection efficiency of the setup (Experimental Section). As shown in Table 1, the yield for these break-up channels is considerably higher than for the channels investigated before. It is thus worthwhile to check if they can also be used to determine absolute configuration.

As one of the fragments is not detected, the linear momenta of the detected ions do not sum up to 0 as in the previous cases. Instead, the sum momentum is attributed to the missing fragment. To suppress some background, the linear momentum vectors of the detected ions and of the missing fragment are required to be smaller than the maximum linear momentum of the respective fragment in the case of the complete break-up.

Figure 7 shows reconstructed linear momenta for several fragmentation pathways, compared to the measured linear momenta for the fragmentation into five singly charged ions. The broadening of the distribution has several sources: Firstly, different fragmentation pathways can lead to the same observed reaction products. This is particularly evident for the undetected fluorine (Figure 7, middle). The shoulder at higher linear momentum values corresponds to the values of the fluorine linear momentum that is observed for the break-up into five ions. We conclude that this part of the distribution is caused by actual break-ups into five atomic ions for which the fluorine cation was not detected. Comparison of the values of the kinetic energy release (KER) for these events and the five-ion fragmentation supports this explanation. Secondly, the broadening is caused by the fact that background cannot be reduced as effectively because the sum momentum check cannot be applied as stringently. Thirdly, the isotopic mass of the detected fragments Cl^+ or Br^+ cannot be determined because the linear momentum distributions of the different isotopes overlap. For the calculation of linear momenta from the raw data, the mass of the lighter isotope was assumed, leading to an error of about 20 atomic units of linear momentum for the molecules containing the heavy isotope. This value, however, is not rooted in physical properties of the molecule but in the experimental parameters (Experimental Section).

Figure 8 (top row) shows triple products for different fragmentation pathways. A separation of the two enantiomers is clearly visible and the overall consistency can be demonstrated using different triple products (bottom row). The background is significantly higher than in the case where all fragments

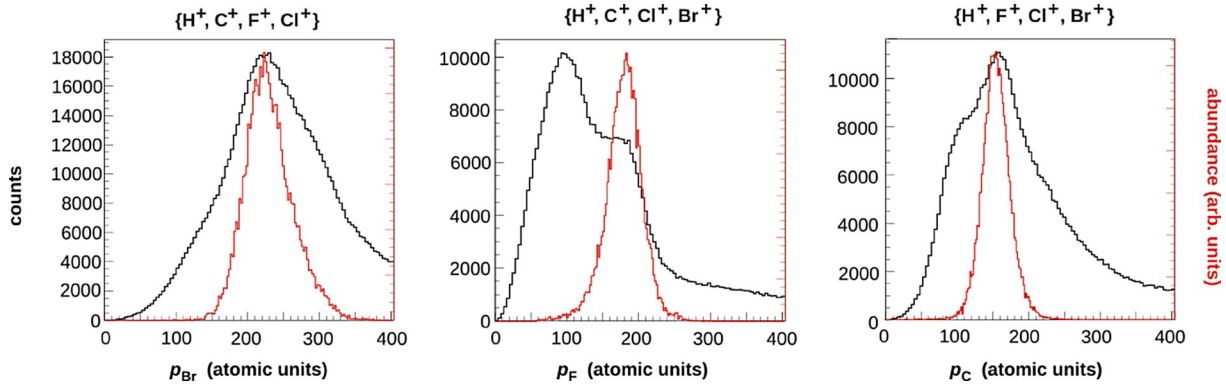


Figure 7. Distribution of reconstructed linear momenta (black) for different fragmentation pathways. The measured ions are given on top; their sum momentum is attributed to the missing mass. For comparison, the linear momentum distribution of the respective ion in the complete fragmentation is displayed (red).

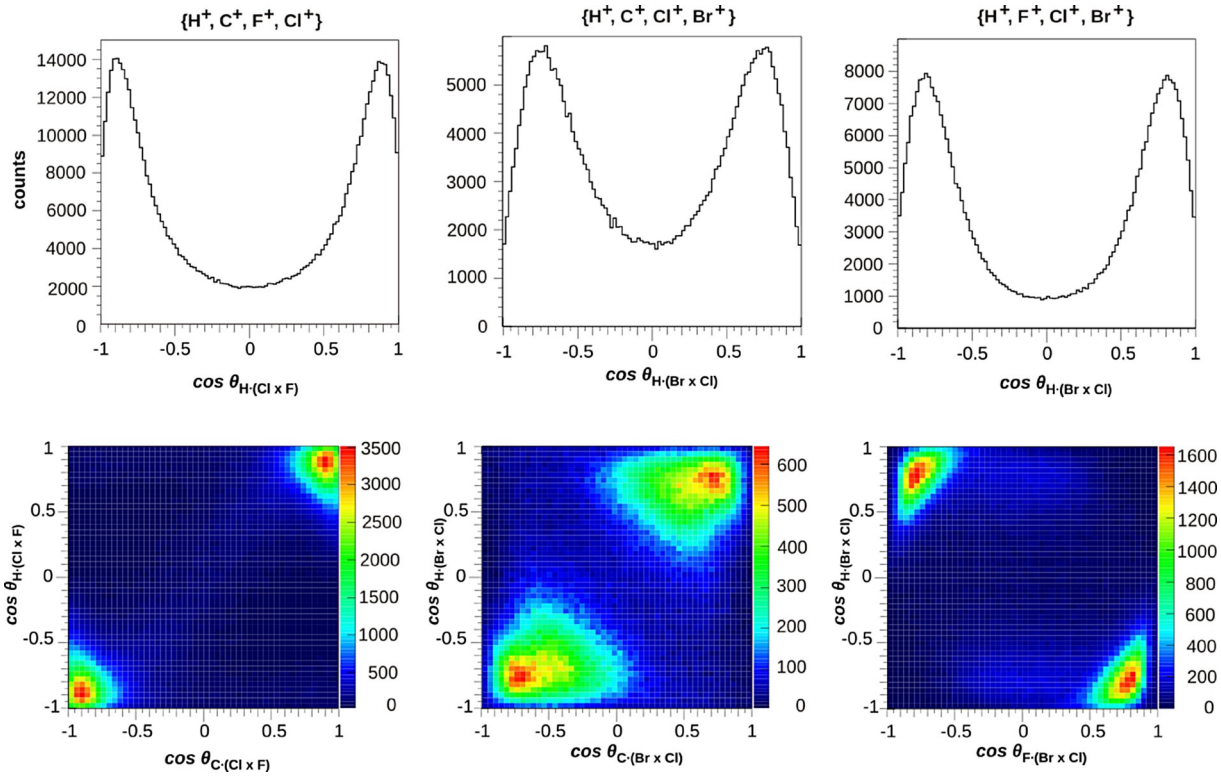


Figure 8. Separation of enantiomers using different fragmentation pathways with one fragment being undetected. The measured fragments are given on top of the column. The top row shows the triple products as defined in the text, the bottom row displays the correlation of two triple products. As is expected from a classical simulation, linear momenta of H^+ and C^+ point in the same direction, resulting in maxima on the main diagonal.

were detected. Contrary to the complete fragmentation, the assignment of absolute configuration is thus not certain on a singe-molecule level anymore.

In order to evaluate how reliable the assignment of enantiomers is, the consistency of seven possible triple products was checked (Table 2 in the Supporting Information). For the fragmentation pathways **IV** and **V**, about 75% of the events show a consistent sign of all seven triple products; for pathway **VI** with detected {H^+, F^+, Cl^+, Br^+}, the consistency lies below 20%. This discrepancy is supposedly due to the multiplicity of (sequential) pathways in which a neutral carbon atom can be ejected during Coulomb explosion.

The events with consistent sign were used to estimate the influence of the unknown isotopic composition of the individual molecules. To do so, linear momenta were calculated for all isotopic combinations of chlorine and bromine isotopes, and the resulting signs of triple products were compared (see Table 3 in the Supporting Information). For every isotopic combination, more than 90% of the events still showed consistency, meaning that for these events, the unknown isotopic composition does not affect the correct assignment of handedness. This value could probably be increased with an improved spectrometer.

In analogy, events with only three detected particles could as well be used to determine handedness, provided an undetected fragment is involved (if this were not the case, all linear momenta would be coplanar due to linear momentum conservation, and thus insufficient to determine the handedness).

Figure 9a shows that the background is even higher than in the four-particle case. As only one triple product can be calculated, the possibility of cross-checks as in the previous cases is lost. Assignment of handedness must thus be taken with caution as rearrangement during the fragmentation process cannot be excluded. Quantum chemical calculations reveal existence of a metastable planar isomer of the CHBrClF trication, which implies a concomitant loss of structural information in CEI. Even for small fragment dications such as $CHBr^{2+}$, isomers are known to exist[20] (see also the review[21] on (meta)stable highly charged molecular ions). Nevertheless, the two distinct peaks at same absolute value indicate a good separation of enantiomers for the fragmentation channels shown here, and the three-dimensional representation for the *R*-enantiomer in the molecular frame (Figure 9b) confirms this finding. Analogue plots for additional break-up channels can be found in the Supporting Information.

3. Conclusions

In this contribution we showed that a single X-ray photon from a synchrotron source can be used to induce multiple fragmentation of the chiral prototype CHBrClF, allowing determination of absolute configuration.

Taking literature on multiple photoionization of noble gas atoms[22,23] into account, we conclude that mostly initial holes in the bromine M-shell and chlorine L-shell contribute to the fragmentation into five and four ions. In order to achieve such

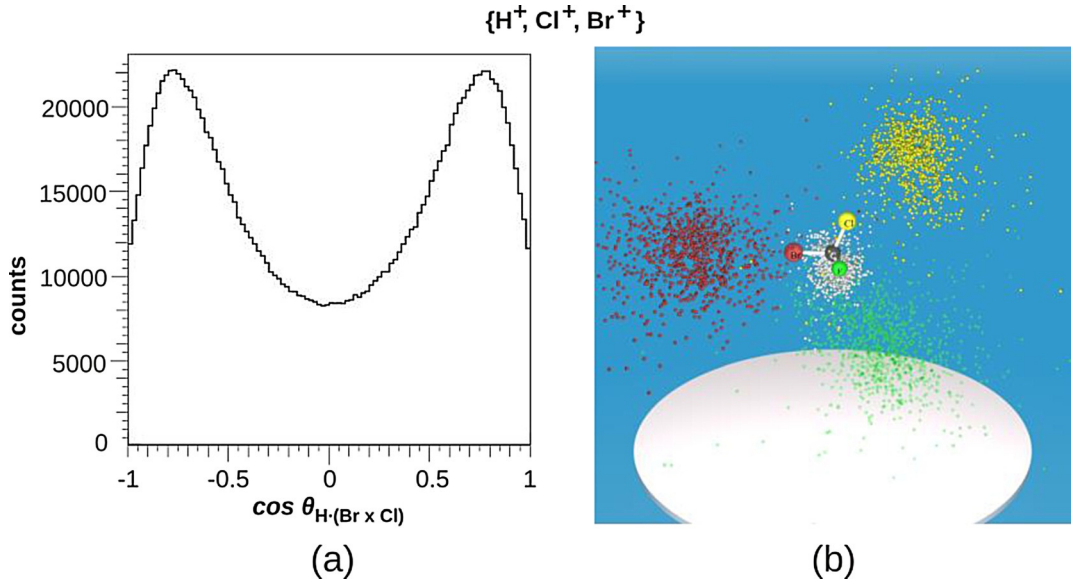


Figure 9. Fragmentation pathway with three detected ions H^+, Cl^+, Br^+. The triple product H·(Br×Cl) indicates a separation of enantiomers (a), albeit only on a statistical level. The three-dimensional distribution of linear momenta in the molecular frame (b) for the *R*-enantiomer confirms the separation. The reconstructed momentum is attributed to fluorine as indicated by the semi-transparent color. Several outliers are also visible. For better visibility, only 1000 events are shown. The transformation into the molecular frame of reference is described in the Supporting Information.

high charge states in organic molecules with only light atoms, energies well above the fluorine 1s state (or respectively carbon 1s or oxygen 1s states in the case of non-halogenated species) seem to be most promising. As work on multiple ionization of neon shows, the probability for quadruple ionization still increases several hundred eV above the neon 1s threshold, due to double core hole excitation[22] or to direct double ionization.[23]

With the photon energy chosen in the present case (710 eV), the efficiency for fragmentation into five singly charged ions was significantly increased compared to previous laser experiments. Moreover, the proton's linear momentum distribution was not as broad as in the latter case, presumably due to faster ionization and fragmentation processes.

Investigation of fragmentation pathways containing the molecular ions CH^+ and CF^+ shows that these channels can be used for the determination of absolute configuration as well. Although efficiency is not increased, this approach seems promising for more complex molecules for which complete fragmentation cannot be achieved anymore.

A significant increase in efficiency is obtained in cases where one of the fragments is not detected. When four atomic ions are collected and one atom is missing, the determination of handedness is still very reliable on a statistical level, and use of different fragments shows consistent assignment of enantiomers. The non-vanishing background, however, leaves an uncertainty in the enantiomeric determination of *individual* molecules. When only three ions are detected, a separation of enantiomers is possible, and efficiency is increased significantly. Together with a more detailed investigation of the fragmentation dynamics, these fragmentation pathways can thus be used to determine the absolute configuration of a sample.

Experimental Section

Measurements were performed with the COLTRIMS-technique that has been described in detail elsewhere.[11,24] The vapor pressure of the sample at room temperature was used to form a supersonic jet by expanding the gas through a 30 μm nozzle into a vacuum chamber. The jet is collimated by two skimmers (300 μm in diameter) and at right angle intersected with a femtosecond laser beam or synchrotron radiation.

The results presented here were obtained at the undulator beamline SEXTANTS[25] of the synchrotron SOLEIL in St. Aubin, France, with a photon energy of 710 eV. For this experiment, the racemic sample of CHBrClF was prepared as described by Swarts[26] via fluorination of $CHBr_2Cl$ with SbF_3 in presence of Br_2. If one considers atomic cross section data, the major contributions to ionization at this energy are expected to stem from the F(1s) state (23%), 21% from Br(3d), 18% from Br(3p) and 15% from Cl(2p) (see Table 1 in the Supporting Information). Formation of double core hole states from C(1s) ionization is considerably less likely, but expected to induce fast formation of higher charge states that are favorable in multiple fragmentation (see for example Ref. [27] for results on the parent compound methane).

Contrary to the laser experiment (Ref. [2] and the Supporting Information herein), lighter elements present in the residual gas are hardly ionized by synchrotron radiation which leads to a lower background signal.

After ionization, positive ions and electrons are separated by the homogeneous electric field of a time-of-flight (TOF) spectrometer. In the measurements presented here, electrons are not considered. Cations are guided onto a time- and position-sensitive detector. The ionic species can be identified because the square root of the mass-to-charge-ratio is proportional to the time-of-flight in an electric field of strength *E*. Contrary to conventional mass spectrometry, the electric field and the length of the spectrometer are tuned in such a way that linear momentum information can be retrieved

from the spread in time-of-flight. In the direction perpendicular to the TOF-axis, ions propagate with the velocity they gain from Coulomb explosion. Linear momenta in all three dimensions can then simply be calculated with the known time-of-flight and impact position on the detector. It should be noted that the effect of the linear momentum spread on the time-of-flight width decreases linearly with increasing field strength E while the separation of the masses decreases only with the square root of the field strength. A compromise between good mass resolution and good linear momentum resolution must thus be found. In the case of complete detection of all fragmentation products, different isotopes can then be separated by virtue of linear momentum conservation. This is not possible for incomplete detection of the fragments. Assuming the wrong isotope will result in an error in the calculated linear momentum; this error will aggravate with higher spectrometer field because the same mass difference corresponds to a bigger linear momentum difference than in case of a low electric field.

When performing multi-hit coincidences, the limited efficiency of the detector severely affects data acquisition rates. Depending on the mass we estimate an efficiency per ion in the order of 20% to 50% for 2 keV ion kinetic energy at impact on the MCP,[28] resulting in a probability of less than 4% to actually record a five-particle event and even lower probabilities for detecting complete breakups of larger molecules.

Acknowledgements

We gratefully acknowledge help from and discussion with Timur Isaev, Alexander Schießer, Guido Grassi, Julia Kiedrowski and Michael Reggelin. The experiments were performed at SOLEIL synchrotron (France) at the SEXTANTS beam line with the approval of the Soleil Peer Review Committee (Project No. 20130255 and 20140056). We are grateful to N. Jaouen for his help during the measurements as well as the SOLEIL staff for stable operation of the storage ring during the experiments. Sebastian Marquardt thanks the Fonds der Chemischen Industrie and Sabrina Marquardt thanks the Deutsche Studienstiftung for a scholarship. Markus Schöffler acknowledges support by the Adolf Messer foundation. This work was supported by the State Initiative for the Development of Scientific and Economic Excellence (LOEWE) in the LOEWE focus ELCH and the German Federal Ministry of Education and Research (BMBF).

Keywords: absolute configuration · chirality · cold target recoil ion momentum spectroscopy · Coulomb explosion imaging · synchrotron radiation

[1] J. Bijvoet, A. Peerdeman, A. van Bommel, *Nature* **1951**, *168*, 271 – 272.
[2] M. Pitzer, M. Kunitski, A. Johnson, T. Jahnke, H. Sann, F. Sturm, L. Schmidt, H. Schmidt-Böcking, R. Dörner, J. Stohner, J. Kiedrowski, M. Reggelin, S. Marquardt, A. Schießer, R. Berger, M. Schöffler, *Science* **2013**, *341*, 1096 – 1100.
[3] P. Herwig, K. Zawatzky, M. Grieser, O. Heber, B. Jordon-Thaden, C. Krantz, O. Novotný, R. Repnow, V. Schurig, D. Schwalm, Z. Vager, A. Wolf, O. Trapp, H. Kreckel, *Science* **2013**, *342*, 1084 – 1086.
[4] Z. Vager, R. Naaman, E. Kanter, *Science* **1989**, *244*, 426 – 431.
[5] L. P. H. Schmidt, T. Jahnke, A. Czasch, M. Schöffler, H. Schmidt-Böcking, R. Dörner, *Phys. Rev. Lett.* **2012**, *108*, 073202.
[6] M. Kunitski, S. Zeller, J. Voigtsberger, A. Kalinin, L. P. H. Schmidt, M. Schöffler, A. Czasch, W. Schöllkopf, R. E. Grisenti, T. Jahnke, D. Blume, R. Dörner, *Science* **2015**, *348*, 551.
[7] J. Voigtsberger, S. Zeller, J. Becht, N. Neumann, F. Sturm, H.-K. Kim, M. Waitz, F. Trinter, M. Kunitski, A. Kalinin, J. Wu, W. Schöllkopf, D. Bressanini, A. Czasch, J. B. Williams, K. Ullmann-Pfleger, L. P. H. Schmidt, M. S. Schöffler, R. E. Grisenti, T. Jahnke, R. Dörner, *Nat. Commun.* **2014**, *5*, 5765.
[8] Z. Vager, E. Kanter, G. Both, P. Cooney, A. Faibis, W. Koenig, B. Zabransky, D. Zajfman, *Phys. Rev. Lett.* **1986**, *57*, 2793 – 2795.
[9] T. Kitamura, T. Nishide, H. Shiromaru, Y. Achiba, N. Kobayashi, *J. Chem. Phys.* **2001**, *115*, 5 – 6.
[10] P. Herwig, K. Zawatzky, D. Schwalm, M. Grieser, O. Heber, B. Jordon-Thaden, C. Krantz, O. Novotný, R. Repnow, V. Schurig, Z. Vager, A. Wolf, O. Trapp, H. Kreckel, *Phys. Rev. A* **2014**, *90*, 052503.
[11] R. Dörner, V. Mergel, O. Jagutzki, L. Spielberger, J. Ullrich, R. Moshammer, H. Schmidt-Böcking, *Phys. Rep.* **2000**, *330*, 95 – 192.
[12] J. Yeh, I. Lindau, *Atom. Data Nucl. Data* **1985**, *32*, 1 – 155.
[13] V. Jonauskas, S. Kuas, R. Karazija, *Phys. Rev. A* **2011**, *84*, 053415.
[14] T. Carlson, W. Hunt, M. Krause, *Phys. Rev.* **1966**, *151*, 41 – 47.
[15] S. Brünken, C. Gerth, B. Kanngießer, T. Luhmann, M. Richter, P. Zimmermann, *Phys. Rev. A* **2002**, *65*, 042708.
[16] B. Erk, D. Rolles, L. Foucar, B. Rudek, S. W. Epp, M. Cryle, C. Bostedt, S. Schorb, J. Bozek, A. Rouzee, A. Hundertmark, T. Marchenko, M. Simon, F. Filsinger, L. Christensen, S. De, S. Trippel, J. Küpper, H. Stapelfeldt, S. Wada, K. Ueda, M. Swiggers, M. Messerschmidt, C. D. Schröter, R. Moshammer, I. Schlichting, J. Ullrich, A. Rudenko, *Phys. Rev. Lett.* **2013**, *110*, 053003.
[17] A. Knapp, A. Kheifets, I. Bray, T. Weber, A. L. Landers, S. Schössler, T. Jahnke, J. Nickles, S. Kammer, O. Jagutzki, L. P. Schmidt, T. Osipov, J. Rösch, M. H. Prior, H. Schmidt-Böcking, C. L. Cocke, R. Dörner, *Phys. Rev. Lett.* **2002**, *89*, 033004.
[18] K. Kreidi, D. Akoury, T. Jahnke, T. Weber, A. Staudte, M. Schöffler, N. Neumann, J. Titze, L. P. H. Schmidt, A. Czasch, O. Jagutzki, R. A. C. Fraga, R. E. Grisenti, M. Smolarski, P. Ranitovic, C. L. Cocke, T. Osipov, H. Adaniya, J. C. Thompson, M. H. Prior, A. Belkacem, A. L. Landers, H. Schmidt-Böcking, R. Dörner, *Phys. Rev. Lett.* **2008**, *100*, 133005.
[19] R. S. Cahn, C. Ingold, V. Prelog, *Angew. Chem. Int. Ed.* **1966**, *5*, 385 – 415.
[20] J. Roithová, J. Zabka, Z. Herman, R. Thissen, D. Schröder, H. Schwarz, *J. Phys. Chem. A* **2006**, *110*, 6447 – 6453.
[21] D. Schröder, H. Schwarz, *J. Phys. Chem. A* **1999**, *103*, 7385 – 7394.
[22] N. Saito, I. Suzuki, *Phys. Scr.* **1994**, *49*, 80 – 85.
[23] T. Carlson, M. Krause, *Phys. Rev.* **1965**, *140*, 1057 – 1064.
[24] J. Ullrich, R. Moshammer, A. Dorn, R. Dörner, L. Schmidt, H. Schmidt-Böcking, *Rep. Prog. Phys.* **2003**, *66*, 1463 – 1545.
[25] M. Sacchi, N. Jaouen, H. Popescu, R. Gaudemer, J. M. Tonnerre, S. G. Chiuzbaian, C. F. Hague, A. Delmotte, J. M. Dubuisson, G. Cauchon, B. Lagarde, F. Polack, *J. Phys. Conf. Ser.* **2013**, *425*, 072018.
[26] F. Swarts, *Bull. Acad. R. Belg.* **1893**, *26*, 102 – 106.
[27] J. H. D. Eland, M. Tashiro, P. Linusson, M. Ehara, K. Ueda, R. Feifel, *Phys. Rev. Lett.* **2010**, *105*, 213005.
[28] M. Krems, J. Zirbel, M. Thomason, R. D. DuBois, *Rev. Sci. Instrum.* **2005**, *76*, 093305.

Manuscript received: December 5, 2015
Final Article published: June 14, 2016

4.4. Theoretical Study of Gas Phase Fragmentation of Multiply Charged Bromochlorofluoromethane (CHBrClF)

4.4.1. Summary

In Ref. [73] we scanned the DFT potential energy hypersurface (PES) of $CHBrClF^{+\bullet}$, $CHBrClF^{2+}$, $CHBrClF^{3+\bullet}$, $CHBrClF^{4+}$ and $CHBrClF^{5+\bullet}$. We found stable isomers for the first three specimen, some of them being achiral (planar). The planar structures of doubly and triply ionized CHBrClF were shown to have geometrical features that are similar to those of formic acid. Dissociation energies of all metastable structures w.r.t. dissociation of single atoms were calculated and were used to further analyze experimental Photoion-Photoion-Coincidence (PI2CO) data. We could show that most predicted channels are present in data obtained from a laser experiment. Using statistical rate theory, we could show that some reaction channels were kinetically hindered and thus did not appear in the PI2CO spectrum. From a BOMD trajectory we determined the internal energy of a molecular fragment to be above the dissociation threshold, thus leading to subsequent fragmentation. Experimental fourfold photoion coincidence (PI4CO) data was analyzed with the previously established methods, where we focused on the fragmentation into three atomic and one molecular fragments. The predominant direct fragmentations into CX^+ and atomic fragments was explained by minimal structural rearrangements, as compared to the formation of XY^+ fragments. The event count per channel in the PI4CO spectrum, however, did not resemble the channel position in the energy diagram. Again, statistical rate theory was used to explain the branching ratios for the competing reaction channels as well as the experimentally predominant break-up into five particles: A large number of the recorded events stems from break-up into five particles where one particle was not detected and thus they form the major part of the PI4CO spectrum. The branching ratio at high energies is dominated by translational degrees of freedom of the fragments. This explained the ratio of experimentally recorded product counts. A simplified equation for the branching ratio was provided that allows estimates by knowledge of atomic masses only. The energy dependence of the branching ratio was also shown by calculating BOMD trajectories for different temperatures. Finally, it was shown that momentum reconstruction from BOMD data give very similar results as compared to the momentum reconstruction from CMD data.

4.4.2. Contribution

For this work, I performed the BOMD calculations and analyzed them w.r.t. fragment formation, charge distributions and phase-space properties. I extended our CEI program to read BOMD trajectories from TURBOMOLE in order to use the CEI reconstruction analysis routines on quantum chemical trajectories. Thus, I could also show that the error by assuming a radial Coulomb explosion is rather small. Using discrete convolution methods, I implemented efficient schemes to calculate the branching ratios of competing reaction channels from kinetic, vibrational, rotational and electronic degrees of freedom. Additionally, I devised an approximative formula for the mass-based estimate of branching ratios. A backtracking algorithm was implemented to name and place all possible signals in X-fold photoion coincidence (PIXCO) spectra given a previously defined set of particles and charges, isotopes. An HTML-based graphical user interface can then be used to identify signals appearing in experimental spectra. With these tools at hand I explained signals under investigation in the PI2CO and PI4CO spectra.

This work was proposed by Sabrina Marquardt and me and evolved further through discussions with Prof. Dr. Robert Berger, Prof. Dr. Reinhard Dörner, Dr. Markus Schöffler and Dr. Martin Pitzer. Sabrina Marquardt and I contributed in equal parts to the writing of the article. Experimental data was provided by Dr. Martin Pitzer and Kilian Fehre.

The following article comprises the unreviewed draft of the article at the time of print of this thesis.

Theoretical study of gas phase fragmentation of multiply charged bromochlorofluoromethane (CHBrClF)

Sebastian Marquardt [*] Sabrina Marquardt [*]
Martin Pitzer [†‡] Kilian Fehre [‡] Horst Schmidt-Böcking [‡]
Reinhard Dörner[‡] Markus Schöffler[‡] Robert Berger [*§]

May 24, 2018

1 Abstract

Structures and energetics of singly, doubly and triply charged ions of bromochlorofluoromethane (CHBrClF) are studied on a composite electronic structure level that combines structure energy optimizations on the density functional theory (DFT) level with single point energy calculations on the explicitly correlated coupled cluster singles, doubles and perturbative triples level (CCSD(T)-F12). Quintuply charged states are studied on the DFT level. For the di- and trications, geometrical features resembling formic acid and its monocation are found. Thus, several local minima on the Born-Oppenheimer potential energy hyper surface of $CHBrClF^{2+}$ and $CHBrClF^{3+\cdot}$ correspond to a planar arrangement of the nuclei and the relevance of these structures in the determination of the absolute configuration of CHBrClF in Coulomb Explosion Imaging experiments as described by Pitzer et al. [Science, **2013**, *341*, 1096–1100; ChemPhysChem, **2016**, *17*, 2465–2472] are discussed. Dissociation energies, branching ratios and Born-Oppenheimer molecular dynamics trajectories for the Coulomb explosion of several multications of CHBrClF are computed and compared to experimental results.

[*]Fachbereich Chemie, Philipps Universität Marburg, Hans-Meerwein-Strasse, 35032 Marburg, Germany
[†]Institut für Physik, Universität Kassel, Heinrich-Plett-Str. 40, 34132 Kassel, Germany
[‡]Institute for Nuclear Physics, Johann Wolfgang Goethe-University Frankfurt, Max-von-Laue-Straße 1, 60438 Frankfurt, Germany
[§]corresponding author

2 Introduction

Multiply charged ions are of great interest for fundamental research in chemistry and physics since Linus Pauling predicted the existence of metastable He_2^{2+} in the year 1933. [1] It took time, until the year 1984, when Guilhaus et al. reported the first experimental observation of the molecule. [2] In the year 1989 Belkacem et al. measured the bond length of He_2^{2+}. [3] The dications of small molecules already show an interesting interplay between chemical bonding and Coulomb repulsion forces. [4, 5] The accumulation of charge at small distances leads to thermodynamic instability, i.e. multiply charged molecules generally possess at least one exothermic fragmentation pathway. [6] The study of multiply charged ions can therefore give new insight into inter atomic interactions and, despite their unstable nature, new highly charged ions that are stable in the time frame of experiments are continuously found. Only recently Yatsuhashi et al. could generate the stable tri- and tetracations of C_2I_2. [7]

There are several methods that are well-established for the investigation of multiply charged ions: On the experimental side there are electrospray ionization, charge stripping or double-charge transfer spectroscopy and on the theory side *ab initio* molecular orbital methods are an essential part of this research field. [8]

In this study we investigated cations of CHBrClF to address the questions if (meta)stable highly charged structures exist and if any of these are achiral. Several other publications describe halide derivatives of methane: Maquin et al. [9] performed charge stripping of CH_3X^+ (X = F, Cl, Br, I). Apeloig et al. [10] found ion-dipole complexes of CHX^+ and HY where X and Y were fluorine or chlorine atoms. Guenat et al. (Ref. [11]) studied CCl_4^{2+}, Roithová et al. (Ref. [12]) looked at the reactivity of $CHBr^{2+}$ with molecular hydrogen and identified CH_3Br^{2+} as intermediates whereas Duflot et al. (Ref. [13]) found different isomers of CH_3Cl^{2+} by complete active space self-consistent field calculations. Grant and coworkers investigated in Ref. [14, 15] CH_3Br^{2+}, ${CH_2Br_2}^{2+}$ and ${CHBr_3}^{2+}$ as well as the corresponding chlorine-substituted species.

In contrast to all mentioned molecules, CHBrClF is chiral in its neutral electronic ground state. It has been the prototype molecule selected for the direct determination of absolute configuration of single molecules in the gas phase with the Cold Target Recoil Ion Momentum Spectroscopy (COLTRIMS) [16, 17] experiment. [18, 19]

During the COLTRIMS experiment CHBrClF is ionized by high energy photons from a synchrotron source or by intense laser pulses. Highly charged ions are ideal for the COLTRIMS momentum reconstruction procedure, e.g. $CHBrClF^{5+\cdot}$. In this experiment, the charged constituents of such a system repel each other and the molecule is said to undergo a Coulomb explosion. The charged products are accelerated within an electric field towards a detector. Their impact on the detector is measured in coincidence, i.e. ions which originate from the same molecule can be grouped together. From the time of flight of the ion and the position of impact on the detector one can reconstruct the initial ($t \to 0$; directly after the Coulomb explosion) linear momenta. If the chiral

structure of the intact molecule is conserved in the fragments' linear momenta, these can be used in order to determine the initial handedness of the exploded molecule on a *per event* basis.

The assignment of absolute configuration from linear momenta relies on the following assumptions in this simple form is possible only if one assumes instantaneous fragmentation after removal of five electrons or stepwise ionization without large structural changes i.e. without inversion at the stereogenic center. But in reality stepwise fragmentation could pose problems. Some of the initial limitations have been lifted recently, where also events invoking three and four fragments have been successfully analyzed.[19] However, in some cases this was possible only at the cost of giving up the *per event* analysis.

We used density functional theory calculations in order to scan the Born-Oppenheimer (BO) surfaces of $CHBrClF^{5+\cdot}$, $CHBrClF^{4+}$, $CHBrClF^{3+\cdot}$, $CHBrClF^{2+}$ and $CHBrClF^{+\cdot}$ for chiral and achiral equilibrium structures. The determination of starting points was guided by the $CHBrClF^{+\cdot}$ structures reported by He and Wang. [20] Furthermore, transition structures on the potential energy hyper surface (PES) were searched. The depths of the local minima were determined with respect to the dissociation into fragments on the explicitly correlated coupled cluster level of theory (CCSD(T)-F12b). Born–Oppenheimer molecular dynamics (BOMD) simulations were performed in order to investigate the molecular dynamics with respect to the charge and spin state of the molecule.

3 Results and Discussion

3.1 Doubly and triply ionized species

3.1.1 Most stable local minima

Stable local minima were found for the mono-, di- and trication of CHBrClF, whereas for the tetra- and pentacation of CHBrClF no minima were observed. The structural parameters are listed in Figs. 1 – 3. The structures are named by a numeral, that indicates the charge state, and a letter. The letters were assigned arbitrarily but similar structures for the different charge states share the same letter. Therefore, structures within one charge state might not be named in alphabetical order. In the case of $CHBrClF^{+\cdot}$, the most stable arrangement is a tetrahedral structure. In contrast, the most stable $CHBrClF^{2+}$ isomer is C_S symmetric with hydrogen attached to bromine and pointing towards the fluorine atom (2E), whereas for $CHBrClF^{3+\cdot}$ it is the isomer with hydrogen attached to chlorine and pointing towards the fluorine atom (3B). Tetrahedral equilibrium structures were located only for $CHBrClF^{+\cdot}$ and $CHBrClF^{2+}$.

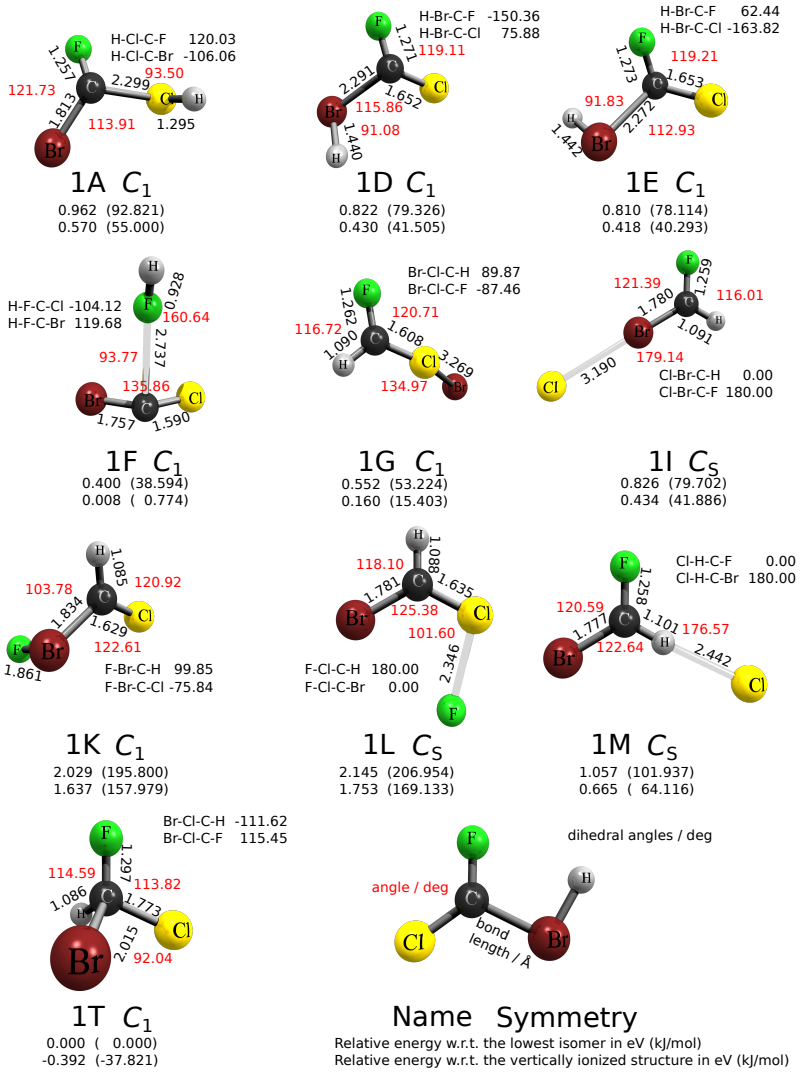


Figure 1: Local minimum structures of $CHBrClF^{+\cdot}$. Color codes are as follows: H (white), C (gray), Br (red), Cl (yellow), F (green). Relative energies without zero point vibrational energy correction are given to the lowest isomer and to the vertically excited structure (not shown). The energies were calculated at the CCSD(T)-F12b/cc-pVTZ-F12 level.

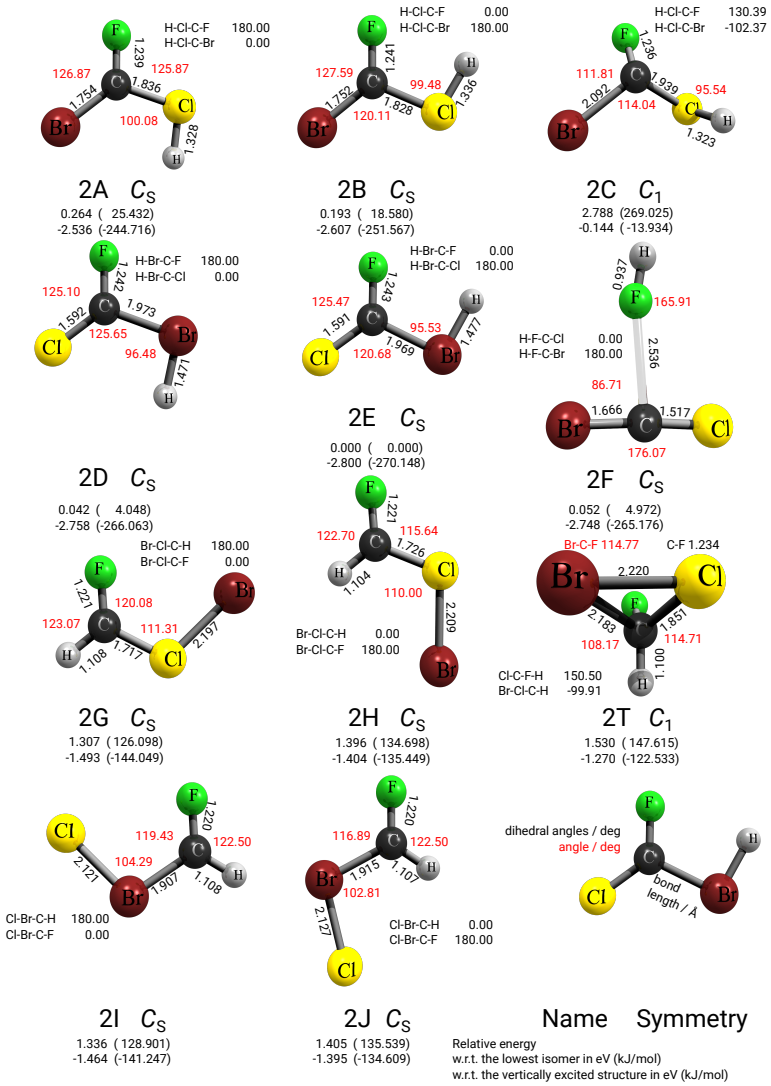


Figure 2: Local minimum structures of $CHBrClF^{2+}$. Color codes are as follows: H (white), C (gray), Br (red), Cl (yellow), F (green). Relative energies without zero point vibrational energy correction are given to the lowest isomer and to the vertically excited structure (not shown). The energies were calculated at the CCSD(T)-F12b/cc-pVTZ-F12 level.

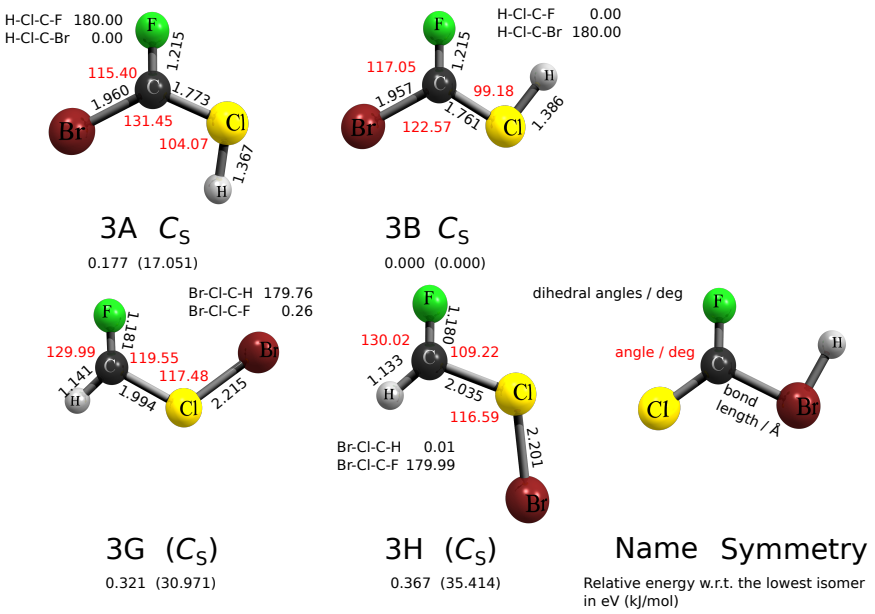


Figure 3: Local minimum structures of $CHBrClF^{3+\cdot}$. Color codes are as follows: H (white), C (gray), Br (red), Cl (yellow), F (green). Relative energies without zero point vibrational energy correction are given to the lowest isomer and to the vertically excited structure (not shown). The energies were calculated at the CCSD(T)-F12b/cc-pVTZ-F12 level.

3.1.2 Planar structures

Structures 1I, 1L, 1M, 2A, 2B, 2D, 2E, 2F, 2G, 2H, 2I, 2J, 3A and 3B are planar. So in any COLTRIMS experiment, in which those structures are populated during the sequential multiple ionization process, information about the initial (neutral) molecule's absolute configuration is lost. For the laser driven Coulomb explosion in Ref. [18] the fivefold ionization occurred within about 35 fsec, for the inner shell ionization induced fragmentation typical time scales are below 10 fsec. So far, the experiments with CHBrClF have been performed only with a racemic mixture that masks any statistical conversion of the absolute configuration. The existence of metastable planar isomers of CHBrClF has already been stated in Ref. [19] and Duflot et al. [13] found planar isomers of another methane derivative, CH_3Cl^{2+}. However, since no indication of planar structures was found for the break-up into five singly-charged ions, we conclude that these structures are not populated in this break-up.

3.1.3 Formic acid motif

The valence electronic structure of the twofold ionized CHBrClF resembles formic acid and the C_S symmetric structures also show similar geometrical features: There are two bonds from the central carbon atom to single atoms of which one should correspond to the C-O double bond in formic acid and therefore be shortened in comparison to the corresponding bond in the tetrahedral structure. The third group, which consists of two atoms, corresponds to the OH-group in formic acid. We can find the described bond shortening for the C-Br bond in 2A and 2B (2.183 Å in 2T, 1.754 Å in 2A, 1.752 Å in 2B) and for the C-Cl bond in 2D and 2E (1.851 Å in 2T, 1.592 Å in 2D, 1.591 Å in 2E). The comparison is shown in Fig. 4. The structure 2T is not strictly tetrahedral. The bromine and chlorine atoms moved closer together (2.220 Å) and form a bridged structure, which is also known for a constitutional isomer of formic acid - dioxirane. [21]

In structure 2F the C-F bond is very long, so that we would rather describe the structure as an ion-dipole complex between $CBrCl^{2+}$ and HF with the former being valence isoelectronic to CO_2. In accordance with that both bonds to Br and Cl are shortened (C-Br 1.666 Å, C-Cl 1.517 Å). In the high energy structures 2G, 2H, 2I and 2J, with F and H being the atomic residues bound to carbon, no significant alternation of the bond lengths can be observed.

As the electronic structure of the twofold ionized CHBrClF resembles formic acid, $CHBrClF^{3+\cdot}$ is expected to resemble the monocation of the formic acid. In the latter, the C-O double bond is elongated and the C-O single bond is shortened in comparison to the neutral species. [22, 23, 24] Indeed, for structures 3A and 3B we find elongated C-Br bonds and shortened C-Cl bonds if we compare them to 2A and 2B (C-Br 1.754 Å in 2A, 1.960 Å in 3A, 1.752 Å in 2B and 1.957 Å in 3B; C-Cl 1.836 Å in 2A, 1.773 Å in 3A, 1.828 Å in 2B and 1.957 Å in 3B, Fig. 4).

We could not find any references that reported a stabilized dication of formic

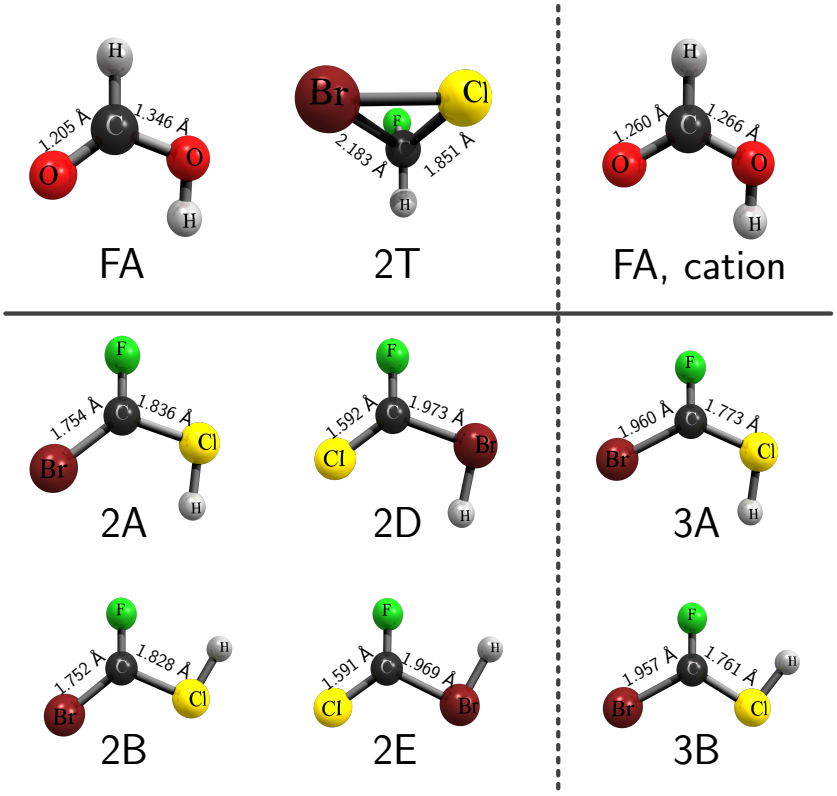


Figure 4: Structure comparison of dicationic and tricationic planar minimum structures of CHBrClF with formic acid (FA) and the cation of formic acid. The structures 2A and 2B show a shortened C-Br bond length in comparison to 2T. A similar bond shortening occurs in FA, where the C=O double bond is shortened with respect to the C-OH bond. In contrast, 2D and 2E show a shortened C-Cl bond in comparison to 2T. The ion 3A (3B) shows an elongation of the C-Br bond and a shortening of the C-Cl bond with respect to 2A. The corresponds to the structural changes when going from FA to the FA cation. Next to C-O, C-Br and C-Cl the corresponding bond lengths are typeset for clarification.

acid, but Wang et al. describe the dissociative double ionization of formic acid. [25, 26] This also matches our results, that there are no stable minima for $CHBrClF^{4+}$.

3.1.4 Transition to planar structure

Using the QST2 method, we could find a transition structure that connects the planar conformers 2G and 2H. However, a transition structure between the iso-

mers 2T and 2G, that one may consider necessary in order to invert the absolute configuration, was not observed. Hence we conclude that in stepwise ionization processes, planar structures of the dication are less likely to be formed directly out of the tetrahedral structure. Therefore it seems reasonable to assume that Coulomb explosion without rearrangements prevails as predominant channel for energy release.

3.1.5 Dissociation energies

Table 1: Dissociation energies for $CHBrClF^{+\cdot}$. The energies were calculated at the cc-pVTZ-F12/CCSD(T)-F12b level and are not corrected with vibrational zero point energies.

Structure	Dissociation reaction	Dissociation energy in eV
1A	$CHBrClF^{+\cdot} \rightarrow CBrClF^{+} + H^{\cdot}$	0.749
	$CHBrClF^{+\cdot} \rightarrow CBrF^{+\cdot} + HCl$	0.541
1D	$CHBrClF^{+\cdot} \rightarrow CBrClF^{+} + H^{\cdot}$	0.889
	$CHBrClF^{+\cdot} \rightarrow CClF^{+\cdot} + HBr$	0.958
1E	$CHBrClF^{+\cdot} \rightarrow CBrClF^{+} + H^{\cdot}$	0.902
	$CHBrClF^{+\cdot} \rightarrow CClF^{+\cdot} + HBr$	0.971
1F	$CHBrClF^{+\cdot} \rightarrow CBrClF^{+} + H^{\cdot}$	1.311
	$CHBrClF^{+\cdot} \rightarrow CBrCl^{+\cdot} + HF$	0.384
1G	$CHBrClF^{+\cdot} \rightarrow CHClF^{+} + Br^{\cdot}$	–0.037
	$CHBrClF^{+\cdot} \rightarrow CHF^{+\cdot} + BrCl$	2.804
1I	$CHBrClF^{+\cdot} \rightarrow CHBrF^{+} + Cl^{\cdot}$	0.182
	$CHBrClF^{+\cdot} \rightarrow CHF^{+\cdot} + BrCl$	2.530
1K	$CHBrClF^{+\cdot} \rightarrow CHBrCl^{+} + F^{\cdot}$	0.028
	$CHBrClF^{+\cdot} \rightarrow CHCl^{+\cdot} + BrF$	1.498
1L	$CHBrClF^{+\cdot} \rightarrow CHBrCl^{+} + F^{\cdot}$	–0.087
1M	$CHBrClF^{+\cdot} \rightarrow CHBrF^{+} + Cl^{\cdot}$	–0.049
	$CHBrClF^{+\cdot} \rightarrow CBrF^{+\cdot} + HCl$	0.446

Energies for dissociation into the various fragments listed in Tabs. 1, 2 and 3 and illustrated in Figs. 5 – 7 suggest that only some of the singly charged species can be termed *stable*. Most of the two- and threefold ionized CHBrClF isomers are metastable and have negative dissociation energies and all of them have at least one channel with a dissociation energy below 0.2 eV. Those ions are not expected to be stable for a longer time period in the experiment. In the case of $CHBrClF^{+\cdot}$, the planar isomers (1I, 1L, 1M) show low dissociation energies as well. The negative dissociation energies for 1G, 1L and 1M are not expected. Frenking and Koch et al. described radical cations and dications of methyl halogenides as ion-dipole complexes. [27, 28] Those ion-dipole complexes should be more stable than the separated fragments. The discrepancy can be explained by the methodology: The structures were optimized at the density functional level, but the energies are calculated as single points on the

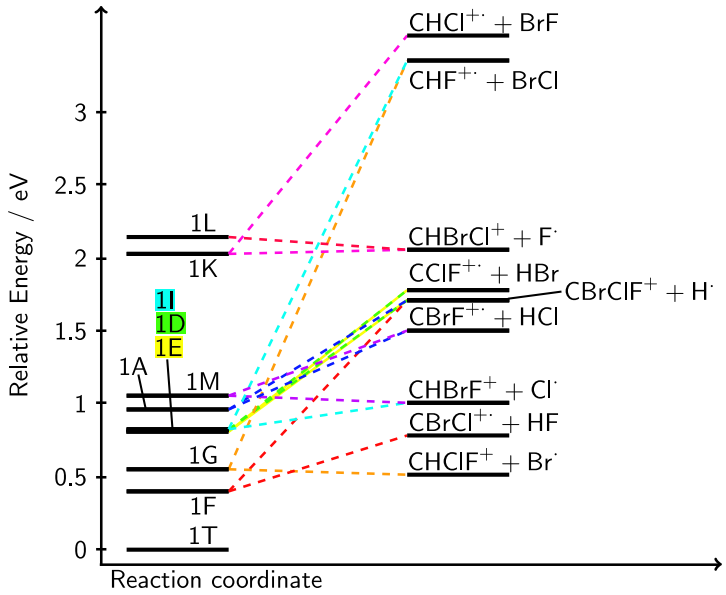


Figure 5: Dissociation pathways for the isomers of $CHBrClF^{+\cdot}$.

coupled cluster PES. The structures might not correspond to local minima at the coupled cluster level. On the density functional level of theory the mentioned isomers have positive dissociation energies. In the case of $CHBrClF^{2+}$ the fragmentation pattern follows the common rule that formation of two odd electron fragments is energetically disfavored as compared to formation of two even electron fragments.

Taking into account the excess energy in the $CHBrClF^{2+}$ system after vertical ionization, we conclude that dissociation follows ionization rapidly. In the case of $CHBrClF^{+\cdot}$ only the tetrahedral isomer has a lower energy than the vertically ionized molecule. We thus conclude that the planar structures are less likely populated. The vertical ionization energy of the tricationic structure could not be obtained: A time-dependent density functional calculation showed that the lowest excitation energy of $CHBrClF^{3+\cdot}$ is about 0.23 eV. We see two close electronic configurations for the trication and thus expect convergence issues due to multi-reference character of the wave function.

3.1.6 Photoion–photoion coincidence spectra

Experimental photoion photoion coincidence (PiPiCo) spectra were obtained by laser and synchrotron experiments. Different charge states can lead to break-up into two cationic fragments. As before, we consider here possible break-ups into from doubly and triply charged CHBrClF. From the calculated dissociation barriers we expected five intense channels on the two-ion coincidence spectrum:

Table 2: Dissociation energies for $CHBrClF^{2+}$. The energies were calculated at the cc-pVTZ-F12/CCSD(T)-F12b level and are not corrected with vibrational zero point energies.

Structure	Dissociation reaction	Dissociation energy in eV
2A	$CHBrClF^{2+} \rightarrow CBrClF^{2+\cdot} + H^{\cdot}$	4.220
	$CHBrClF^{2+} \rightarrow CBrClF^{+} + H^{+}$	–0.102
	$CHBrClF^{2+} \rightarrow CBrF^{2+} + HCl$	2.995
2B	$CHBrClF^{2+} \rightarrow CBrClF^{2+\cdot} + H^{\cdot}$	4.291
	$CHBrClF^{2+} \rightarrow CBrClF^{+} + H^{+}$	–0.031
	$CHBrClF^{2+} \rightarrow CBrF^{2+} + HCl$	3.066
2C	$CHBrClF^{2+} \rightarrow CBrClF^{2+\cdot} + H^{\cdot}$	1.695
	$CHBrClF^{2+} \rightarrow CBrClF^{+} + H^{+}$	–2.627
	$CHBrClF^{2+} \rightarrow CBrF^{2+} + HCl$	0.471
2D	$CHBrClF^{2+} \rightarrow CBrClF^{2+\cdot} + H^{\cdot}$	4.441
	$CHBrClF^{2+} \rightarrow CBrClF^{+} + H^{+}$	0.119
	$CHBrClF^{2+} \rightarrow CClF^{2+} + HBr$	4.288
2E	$CHBrClF^{2+} \rightarrow CBrClF^{2+\cdot} + H^{\cdot}$	4.483
	$CHBrClF^{2+} \rightarrow CBrClF^{+} + H^{+}$	0.161
	$CHBrClF^{2+} \rightarrow CClF^{2+} + HBr$	4.331
2F	$CHBrClF^{2+} \rightarrow CBrClF^{2+\cdot} + H^{\cdot}$	4.432
	$CHBrClF^{2+} \rightarrow CBrClF^{+} + H^{+}$	0.110
	$CHBrClF^{2+} \rightarrow CBrCl^{2+} + HF$	0.857
2G	$CHBrClF^{2+} \rightarrow CHClF^{2+\cdot} + Br^{\cdot}$	4.369
	$CHBrClF^{2+} \rightarrow CHClF^{+} + Br^{+}$	–2.796
	$CHBrClF^{2+} \rightarrow CHF^{2+} + BrCl$	6.895
2H	$CHBrClF^{2+} \rightarrow CHClF^{2+\cdot} + Br^{\cdot}$	4.280
	$CHBrClF^{2+} \rightarrow CHClF^{+} + Br^{+}$	–2.885
	$CHBrClF^{2+} \rightarrow CHF^{2+} + BrCl$	6.806
2I	$CHBrClF^{2+} \rightarrow CHBrF^{2+\cdot} + Cl^{\cdot}$	3.330
	$CHBrClF^{2+} \rightarrow CHBrF^{+} + Cl^{+}$	–2.570
	$CHBrClF^{2+} \rightarrow CHF^{2+} + BrCl$	6.866
2J	$CHBrClF^{2+} \rightarrow CHBrF^{2+\cdot} + Cl^{\cdot}$	3.261
	$CHBrClF^{2+} \rightarrow CHBrF^{+} + Cl^{+}$	–2.639
	$CHBrClF^{2+} \rightarrow CHF^{2+} + BrCl$	6.797

1. $\{CBrClF^{+},H^{+}\}$,
2. $\{CHClF^{+},Br^{+}\}$,
3. $\{CHBrF^{+},Cl^{+}\}$,
4. $\{CBrClF^{2+\cdot},H^{+}\}$ and
5. $\{CHClF^{2+\cdot},Br^{+}\}$.

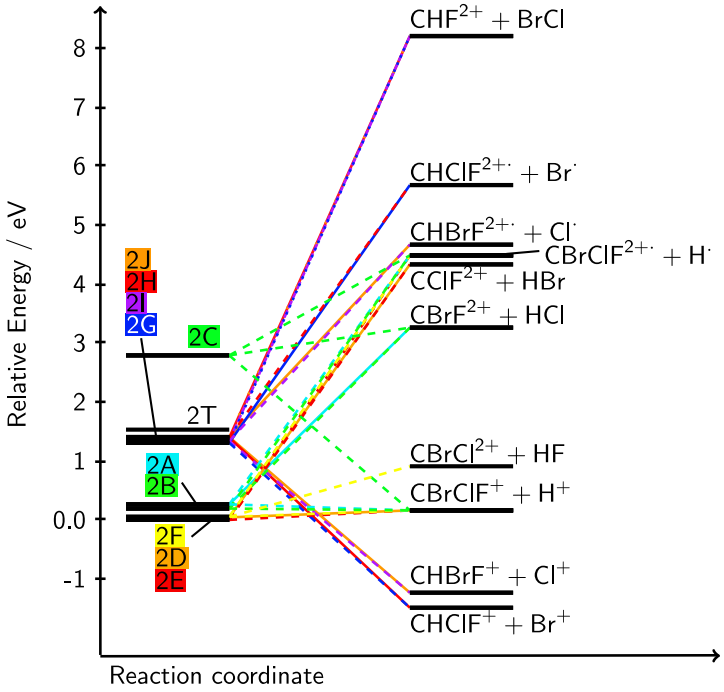


Figure 6: Dissociation pathways for the isomers of $CHBrClF^{2+}$.

Tab. 4 lists the total number of events for each channel out of a total of $4.7 \cdot 10^8$ events in the laser experiment and out of $5 \cdot 10^7$ events in the synchrotron experiment. The events were selected by restricting the momentum sum of coinciding particles. The data was not corrected for reduced detection efficiency of the heavier ions.

Fig. 8 shows the experimental spectrum with an overlay of predicted signals shown as black spots. The signals were simulated with a local computer program by generating all atom, isotope and charge combinations of CHBrClF and calculating the time-of-flight (TOF) of the resulting charged fragments in a specific spectrometer setup. As already indicated by Tab. 4, decomposition products of reactions 1, 2, 3 and 5 can be found. Reaction channel 4 is not open in the laser experiment. In contrast, the results of a similar synchrotron experiment show a significantly increased population of all channels where also channel 4 is populated.

The relative weight of channel 4 w.r.t. channel 5 is understood qualitatively by statistical rate theory: A unimolecular reaction rate constant is written as

$$k(E) = \frac{\sigma W^{\ddagger}(E - E_0)}{h\rho(E)} \tag{1}$$

Table 3: Dissociation energies for $CHBrClF^{3+\cdot}$. The energies were calculated at the cc-pVTZ-F12/CCSD(T)-F12b level and are not corrected with vibrational zero point energies.

Structure	Dissociation reaction	Dissociation energy in eV
3A	$CHBrClF^{3+\cdot} \rightarrow CBrClF^{3+} + H^{\cdot}$	5.506
	$CHBrClF^{3+\cdot} \rightarrow CBrClF^{2+\cdot} + H^{+}$	–6.203
	$CHBrClF^{3+\cdot} \rightarrow CBrF^{3+\cdot} + HCl$	7.367
3B	$CHBrClF^{3+\cdot} \rightarrow CBrClF^{3+} + H^{\cdot}$	5.682
	$CHBrClF^{3+\cdot} \rightarrow CBrClF^{2+\cdot} + H^{+}$	–6.026
	$CHBrClF^{3+\cdot} \rightarrow CBrF^{3+\cdot} + HCl$	7.544
3G	$CHBrClF^{3+\cdot} \rightarrow CHClF^{2+\cdot} + Br^{+}$	–5.608
3H	$CHBrClF^{3+\cdot} \rightarrow CHClF^{2+\cdot} + Br^{+}$	–5.654

Table 4: Number of two-ion coincidence events in the experimental spectra for fragmentation channels with negative dissociation energies. Relative channel intensities are shown with respect to the most intense break-up to correct for different ionization and detection efficiencies. Non-integer numbers in the event counts are due to background subtraction.

Molecule	Channel	Dissociation energies in eV	Event count laser (rel. strength)		Event count synchr. (rel. strength)	
2A-C	$\{CBrClF^{+},H^{+}\}$	$-0.031/-0.102/-2.627$	$6.5 \cdot 10^4$	(0.01)	3967.5	(0.17)
2G-H	$\{CHClF^{+},Br^{+}\}$	$-2.796/-2.885$	$4.9 \cdot 10^6$	(1.00)	23901	(1.00)
2I-J	$\{CHBrF^{+},Cl^{+}\}$	$-2.570/-2.639$	$1.05 \cdot 10^6$	(0.21)	11722	(0.49)
3A-B	$\{CBrClF^{2+\cdot},H^{+}\}$	$-6.026/-6.203$	–	(0.00)	3169.5	(0.34)
3G-H	$\{CHClF^{2+\cdot},Br^{+}\}$	$-5.608/-5.654$	$9.3 \cdot 10^5$	(0.19)	9201	(1.00)

where σ describes the reaction path degeneracy, $W^{\ddagger}(E-E_0)$ is the sum of states (SOS) of the activated complex and $\rho(E)$ is the density of states (DOS) of the reactant. E_0 is the activation barrier for this channel. In this equation, the quantum number J has already been integrated out. The branching ratio of two competing reactions is given by the fraction of the relevant rate constants.

$$R(E) = \frac{k_A(E)}{k_B(E)} = \frac{W_A^{\ddagger}(E - E_{0,A})}{W_B^{\ddagger}(E - E_{0,B})} \tag{2}$$

We assume that both reactions have identical reactants and do not continue, i.e., each product molecule is assumed to have infinite lifetime. Including rotational, vibrational, electronic and translational degrees of freedom (DOF), we obtain the branching ratio shown in Fig. 9. Note that $J_{\max} = 2 \times 10^5$ was chosen to ensure a sufficient decay of the population of rotational states. The vibrational DOS is obtained using a Morse oscillator in the dissociation modes and an harmonic oscillator for the remaining modes. Electronic excitations were obtained using the random phase approximation for the molecules (only the 15 lowest eigenstates) and NIST [29] data for Br^+ up to the ionization threshold. The energies are shifted to the onset of the lower lying channel (5). After vertical

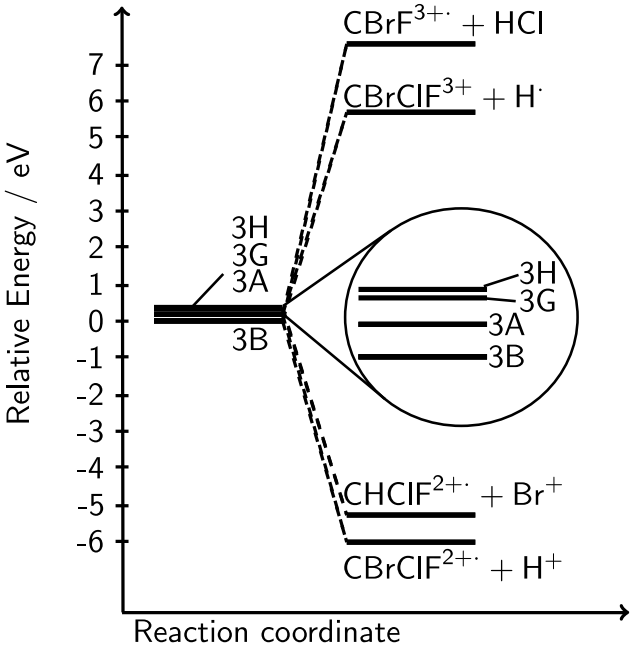


Figure 7: Dissociation pathways for the isomers of $CHBrClF^{3+\cdot}$.

ionization of CHBrClF to $CHBrClF^{3+\cdot}$, approx. 10.3 eV are available as internal energy in channel 5 which gives a ratio of $R(10.3\text{ eV}) \approx 0.13$. Our branching ratio estimate is thus in good agreement with experimental event count ratios.

3.1.7 Dynamic configuration conversion

Dynamical interconversion processes were studied by BOMD simulations of the Coulomb explosions of $CHBrClF^{2+}$ and $CHBrClF^{3+\cdot}$. Trajectories were calculated for vertically ionized $CHBrClF^{2+}$ and $CHBrClF^{3+\cdot}$ for the first 726 fs after ionization. A video of such a trajectory can be found in the supplementary material. In both cases we see a dissociation into a planar molecular fragment CHClF and Br. The kinetic energies in the final frame depend somewhat on the electronic spin state, but do not differ significantly (see Tab. 5). Although one of the frames in each simulation shows a planar arrangement of all atoms, no relaxation into the corresponding local PES minima occurred. Tab. 6 shows selected atom distances in the planar configuration during Coulomb explosion.

In these planar configurations the atoms carry already enough momentum so that Coulomb explosion prevails over relaxation into a local minimum. Further ionization and fragmentation from this structure would not necessarily lead to a loss of chirality information. However, a rotation of the planar CHClF fragment, as is shown in the video, before subsequent fragmentation, likely leads to arbitrary absolute configuration assignments. The probability of subsequent

fragmentation is strongly determined by the excess energy stored in internal DOF of the molecular fragment after the first Coulomb explosion. We calculated this for the molecular fragment $CHClF^{2+}$ after Coulomb explosion on the doublet electronic spin state surface. For that, we tracked the center of mass (COM) of the molecular fragment and calculated its velocity, linear momentum and kinetic energy. By subtracting the COM velocity from the atoms' velocities, we obtained the velocities of the atoms due to their internal (rovibrational) motion. Relaxing the structure of the fragment in a frame, for which the internal kinetic energy is known, we obtain the internal potential energy as the energy gain during relaxation. The overall internal kinetic energy in the last frame is

$$E_{\text{kin}}^{\text{int}} = 1.83 \text{ eV} \tag{3}$$

From the fragment relaxation, we get

$$E_{\text{pot}}^{\text{int}} = 0.22 \text{ eV} \tag{4}$$

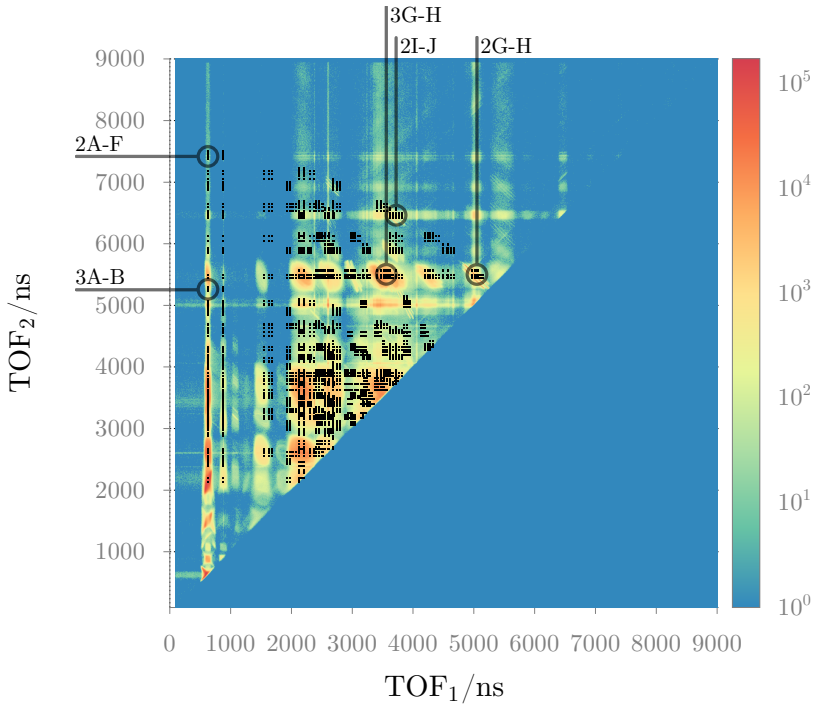


Figure 8: PiPiCo spectrum of CHBrClF overlaid with predicted appearance times. The spectrum is colored according to the ion count and the predicted appearance times are shown as black spots. The circles indicate the time-range in which specific reaction products should appear.

Summing up, we obtain the total energy stored in internal DOF.

$$E_{\mathrm{tot}}^{\mathrm{int}} = 2.05 \text{ eV} \tag{5}$$

The dissociation reaction

$$\mathrm{CHClF}^{2+} \rightarrow [\mathrm{CHF} + \mathrm{Cl}]^{2+}$$

has a barrier of 1.56 eV (using DFT). This is smaller than the energy remaining in the internal DOF and thus can lead to further dissociation of the molecular fragment.

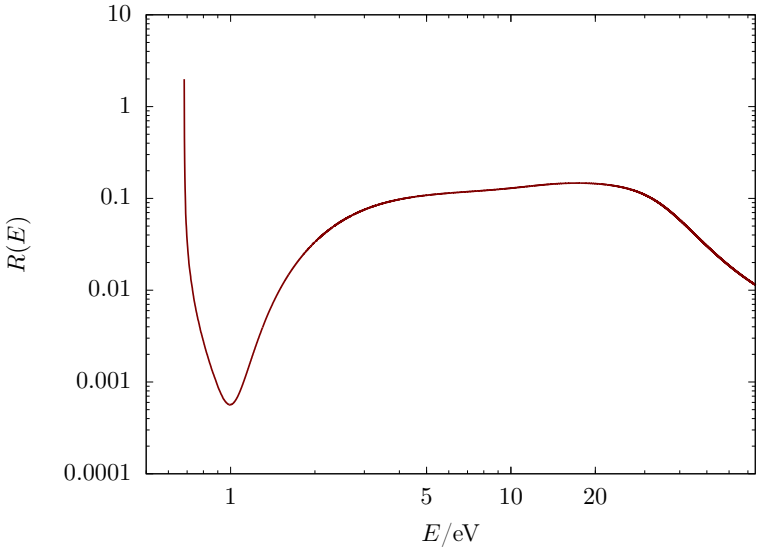


Figure 9: Branching ratio $R_{\mathrm{CBrClF}^{2+}/\mathrm{CHClF}^{2+}}$ for the for the comparison of CBrClF^{2+} versus CHClF^{2+} formation. The ratio is obtained as a function of the excitation energy. For internal energies above 1 eV the formation of CHClF^{2+} dominates. The steep increase of at low energies as well as the deacrease at high energies are possibly a numerical artifacts.

Table 5: Fragment composition, partial charges from Mulliken analysis and KER following from BOMD trajectories. The values are calculated from the last frame of the trajectories after Coulomb explosions of CHBrClF in different electronic charge and spin states.

Total charge	Spin state	E_{kin}/eV	Fragments
+3	doublet	8.9	$Br^{1.4+}$, $CHClF^{1.6+}$
+3	quartet	9.1	$Br^{1.4+}$, $CHClF^{1.6+}$
+2	singlet	3.5	Br^{+}, $CHClF^{+}$
+2	triplet	4.5	Br^{+}, $CHClF^{+}$

Table 6: The table shows the atomic distances of neighboring atoms, calculated at snapshots of the Coulomb explosion trajectories of $CHBrClF^{2+}$ and $CHBrClF^{3+\cdot}$, where the system has C_S symmetry.

Charge, Spin	C–H/Å	C–F/Å	C–Cl/Å	Cl–Br/Å	F–Br
2+, singlet	1.12	1.22	1.70	2.63	4.40
2+, triplet	1.14	1.36	1.65	3.47	4.19
3+, doublet	1.22	1.30	1.74	4.67	4.70
3+, quartet	1.22	1.27	1.66	4.29	4.20

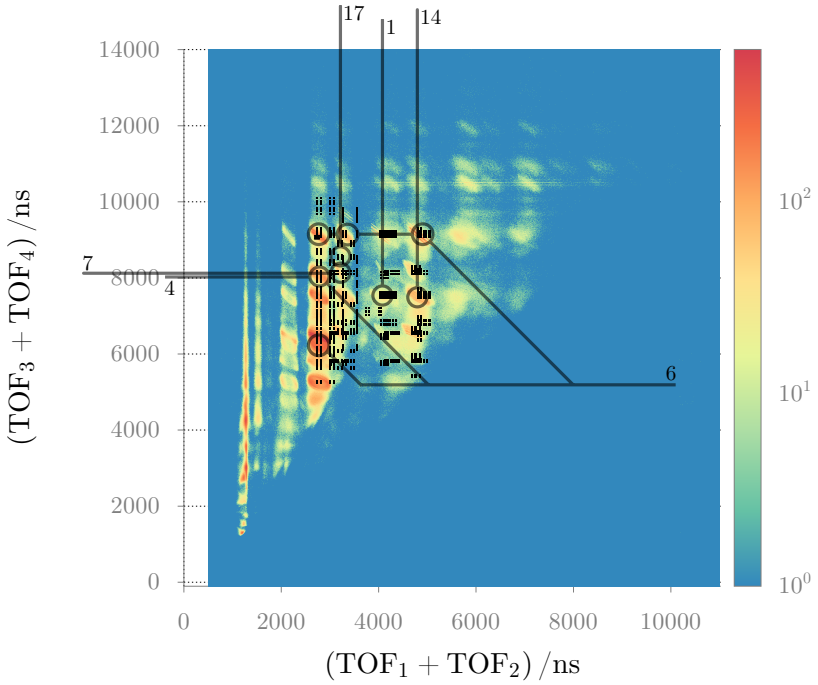


Figure 10: Pi4Co spectrum of CHBrClF overlaid with predicted appearance times. The histogram is colored according to the ion count and the predicted appearance times are shown as black spots. The circles indicate the time-range in which specific reaction channels should appear. The numbering refers to the channel enumeration in Fig. 11.

3.2 Quintuple ionization

3.2.1 Pi4Co analysis and branching ratios

The protocol established for PiPiCo analysis can also be used to analyze the Pi4Co (fourfold coincidence) spectrum of $CHBrClF^{5+\cdot}$. The experimental spectrum was again obtained by laser ionization and is shown in Fig. 10. As the fragmentation into five atoms was analyzed previously in Ref. [18], we focus here on the fragmentation into four particles.

Firstly, we determine the number of possible fragmentation channels under certain conditions: 10 different possibilities exist to create a diatomic molecule out of 5 distinct atoms. Then, we consider only those channels that leave Br or Cl in a charge state of 2+. The only channel, where we have only one possibility of distributing the 2+ charge contains the fragment BrCl. In total, we have thus 19 possibilities for a break-up into 4 particles. All of these channels together with the break-up into 5 particles are shown in the energy diagram in Fig. 11. The channels are ordered according to their appearance energy with respect to

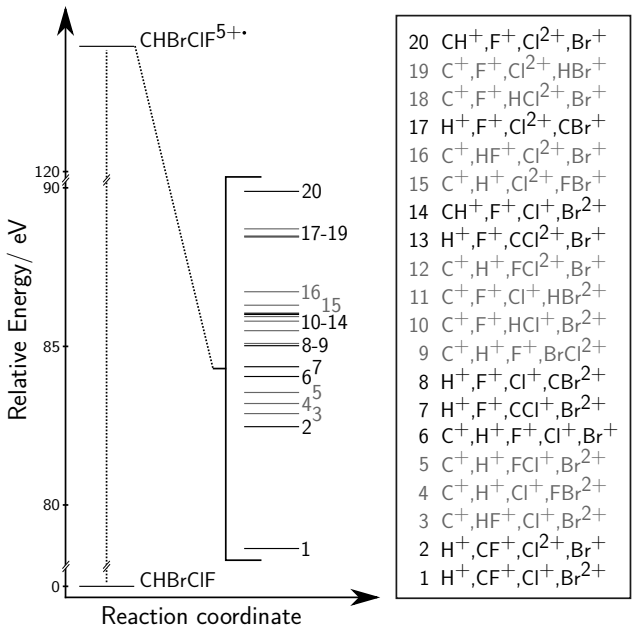


Figure 11: Appearance energies of fragmentation channels of $CHBrClF^{5+}$. Only channels leading to four or five fragments with Cl and Br in charge states + and 2+ are shown. The channels shown in dark color correspond to direct fragmentations, i.e., formation of CX^+ fragments.

the neutral parent molecule. The lowest energy fragmentation channel leads to the fragments CF^+, H^+, Cl^+ and Br^{2+}. The second channel lies approx. 5 eV higher. Fragmentation into five fragments is shown as channel number 6.

In Tab. 7 we show event counts from Pi4Co laser experiments: We searched for the break-up channels 1, 7, 14 and 17 and a sample of other channels. Only the direct fragmentation channels, which form CX^+ fragments, were significantly populated. Except for channel 17, we did not search for channels with Cl in the charge state 2+. Although channel 1 is energetically the most favorable, its relative strength is only a quarter of that of high lying channel 14. As in the branching ratios of $CHBrClF^{3+\cdot}$, the relative strength of the direct fragmentation channels is explained by statistical rate theory:

The analysis shows that the experimental ratio (shown in Tab. 7) correspond to an internal energy of 30 eV w.r.t. the energy of channel 1, which is close to the estimated excess energy of 45 eV (energy difference between fivefold vertically ionized CHBrClF and the lowest energy dissociation products). Theoretical branching ratios are also shown in Tab. 7. In Fig. 12, the branching ratios show that the energetically favored channel is usually stronger populated at low internal energies, whereas at higher energies population of the energetically unfavored channels increases significantly. Then, the CH^+ forming reactions are

Table 7: Analysis of experimental Pi4Co data for channels 1, 7, 14 and 17. Data of the laser experiment was used, setting a momentum gate according to the fragment masses with highest natural isotopic abundance. Other isotopologues were not considered. An energy threshold was used to remove statistical coincidences of unrelated events. The most abundant channels with a total charge of 4+ and a diatomic fragment is also the most abundant channels with a total charge of 5+. Channel 6 has been thoroughly discussed in Ref. [18] and was thus not included in the analysis.

Channel	absolute count	rel. strength	rel. strength computed @ 45 eV
1	9100	0.28	0.34
7	1300	0.04	0.13
14	33000	1.00	1.00
17	< 100	0	0.07
CF^+, H^+, Cl^+, Br^+ (1)	52000	0.35	0.38
CH^+, Br^+, Cl^+, F^+ (14)	150000	1.00	1.00

in general much faster. Predominance of the CH forming channel is explained by the number of translational and electronic states: The ratio of the mass dependent prefactor of the translational sums of states is given by

$$R_{\text{trans}} = \left(\frac{m_{\text{CX}_1} m_{\text{H}} m_{\text{X}_2} m_{\text{CX}_3}}{m_{\text{CH}} m_{\text{X}_1} m_{\text{X}_2} m_{\text{CX}_3}} \right)^{\frac{3}{2}} \tag{6}$$

For all three channels CX, X $\in$ {F,Cl,Br}, this ratio is below 0.05. Furthermore, no electronic states are available for the proton and thus the number of electronic states is significantly lower. Similar arguments explain also the predominance of signals in the Pi4Co spectrum that are attributed to dissociation into five particles, where one may not have been detected: The corresponding ratio of the translational sum of states prefactors of channel 1 w.r.t. channel 9 is $R_{\text{trans}} \approx 0.14$. Calculating the ratio using all DOFs gives $R(45\text{ eV}) \approx 10^{-31}$, which is much lower than expected. However, in this case one compares dissociations into products with different particle numbers. Hence for high energies, the translational sums of states ratio vanishes. We elaborate this further in Sec. 3.2.2.

3.2.2 Approximations to $R(E)$

We will now estimate the branching ratio in terms of the fragments' masses by exploiting the mass dependence of all DOF. In a system consisting of diatomic molecules and atoms, the main contribution to the SOS at high energies comes from the translational DOS, which is proportional to the product of the masses.

$$\rho_{\text{sys}}^{\text{trans}}(E) \propto \left(\prod_{i=1}^{N} m_i \right)^{\frac{3}{2}} E^{\frac{3N}{2}-1} \tag{7}$$

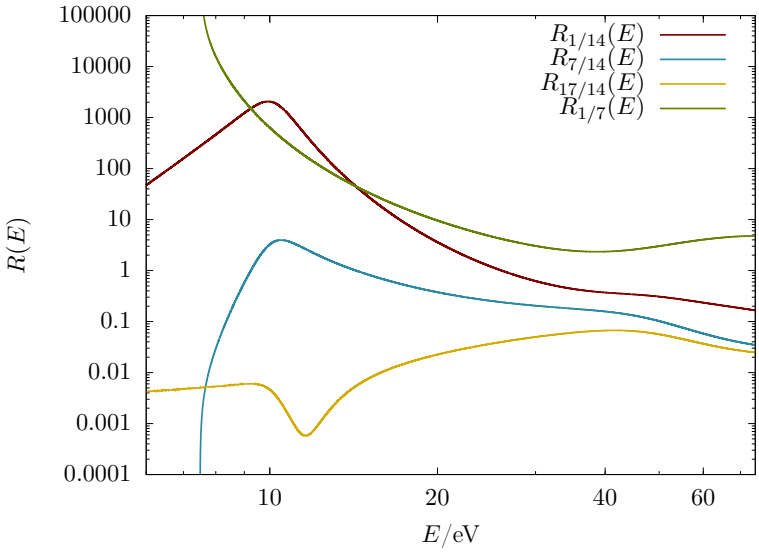


Figure 12: Branching ratios for direct fragmentations of $CHBrClF^{5+}$. At low energies usually only the energetically more favorable channel is populated whereas for higher energies population of the energetically unfavorable channel increases significantly.

The rotational DOS is proportional to $1/B$, B being the rotational constant, and thus proportional to the reduced mass of the diatomic fragment, whereas the vibrational DOS is proportional to ω.

$$\rho^{\mathrm{rot}}_{MA}(E) \propto \frac{1}{B} \propto \mu \tag{8}$$

$$\rho^{\mathrm{vib}}_{MA}(E) \propto \frac{1}{\omega} \propto \sqrt{\mu} \tag{9}$$

If we assume that only the lowest electronic state of all fragments contributes to the DOS, then we can simply integrate the translational DOS in order to obtain the SOS of the full system.

$$W^{\ddagger}_{\mathrm{sys}}(E) \propto \frac{\mu_{MA}}{\omega_{MA}} \left(\prod_i^N g_{i,0} \right) \left(\prod_i^N m_i \right)^{\frac{3}{2}} \frac{2E^{\frac{3N}{2}}}{3N} \tag{10}$$

$$\propto \mu_{MA}^{\frac{3}{2}} \left(\prod_i^N g_{i,0} \right) \left(\prod_i^N m_i \right)^{\frac{3}{2}} \frac{2E^{\frac{3N}{2}}}{3N} \tag{11}$$

Hence we get for the branching ratio

$$R(E) \propto \tilde{R}_1(E) = \left(\frac{\mu_{MA}}{\mu_{MB}}\right)\left(\frac{\omega_{MA}}{\omega_{MB}}\right)^{-1} \frac{\prod_i^N g_{Ai,0}}{\prod_i^N g_{Bi,0}} \left(\frac{\prod_i^N m_{Ai}}{\prod_i^N m_{Bi}}\right)^{\frac{3}{2}} \frac{N_B (E-E_{0A})^{\frac{3N_A}{2}}}{N_A (E-E_{0B})^{\frac{3N_B}{2}}} \tag{12}$$

$$\propto \tilde{R}_2(E) = \left(\frac{\mu_{MA}}{\mu_{MB}}\right)^{\frac{3}{2}} \frac{\prod_i^N g_{Ai,0}}{\prod_i^N g_{Bi,0}} \left(\frac{\prod_i^N m_{Ai}}{\prod_i^N m_{Bi}}\right)^{\frac{3}{2}} \frac{N_B}{N_A} \frac{(E-E_{0A})^{\frac{3N_A}{2}}}{(E-E_{0B})^{\frac{3N_B}{2}}} \tag{13}$$

The symbols $\tilde{R}_1(E)$ and $\tilde{R}_2(E)$ were introduced to emphasize the approximative character of both branching ratio estimates. In the limit of high energies, $E \gg E_{0A}, E \gg E_{0B}$, the qualitative branching ratio does not depend on the activation energies E_{0A} and E_{0B}. Then, the difference of the particle numbers $N_A - N_B$ solely determines the limiting value of the energy dependent term.

$$\lim_{E\to\infty} \frac{N_B (E-E_{0A})^{\frac{3N_A}{2}}}{N_A (E-E_{0B})^{\frac{3N_B}{2}}} \approx \lim_{E\to\infty} E^{\frac{3}{2}(N_A-N_B)} \tag{14}$$

$$= \begin{cases} \infty & , N_A > N_B \\ 1 & , N_A = N_B \\ 0 & , N_A < N_B \end{cases} \tag{15}$$

This result is understood in terms of the number of translational states available: The more particles a channel produces, the higher is the SOS at high energies. In the case of equal particle numbers, the qualitative branching ratio limit is thus determined by

$$\lim_{E\to\infty} R(E) \propto \left(\frac{\mu_{MA}}{\mu_{MB}}\right)^{\frac{3}{2}} \frac{\prod_i^N g_{Ai,0}}{\prod_i^N g_{Bi,0}} \left(\frac{\prod_i^N m_{Ai}}{\prod_i^N m_{Bi}}\right)^{\frac{3}{2}} \tag{16}$$

Despite the inherent simplifications, we think that this simple estimate will be of value in the future design and analysis of specific CEI experiments.

In Fig. 13 we compare the approximations in (12) and (13) with the direct count for all DOF. In general, all curves show the same overall trend. Necessarily, all features of the full calculation must be smoothed out. Hence we do not see small variations of $R(E)$ that are due to other effects than translation.

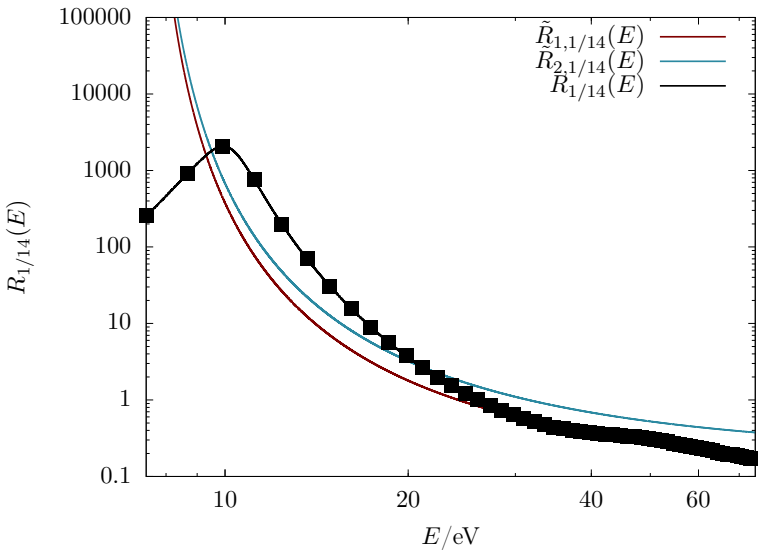


Figure 13: Comparison of branching ratios $R(E)$ for the reactions $CHBrClF^{5+\cdot}$ $\rightarrow$ CF^+, H^+, Br^{2+}, Cl^+ and $CHBrClF^{5+\cdot}$ $\rightarrow$ CH^+, F^+, Br^{2+}, Cl^+. The solid lines correspond to the approximations given in (12) (Approximation 1, red line) and (13) (Approximation 2, blue line). The black dots were calculated with a direct count algorithm. We used an energy resolution of 10 cm^{-1}, a maximum energy $E_{max} = 6 \cdot 10^5 cm^{-1} (\approx 75$ eV) and $\Delta E = 59842 cm^{-1}$.

3.2.3 Temperature-dependent fragmentation patterns

In the previous section we showed by statistical means that certain reaction paths only open up at higher internal energies but then may dominate the branching ratios. We show in the following that similar results can be obtained dynamically and are connected to structural rearrangements as they can occur at higher vibrational excitations.

A BOMD simulation of the $CHBrClF^{5+\cdot}$ Coulomb explosion at 0 K internal temperature shows fragmentation into CF^+, H^+, Br^{2+}, Cl^+ and F^+. Only at higher temperatures, other channels open up. In the structure of neutral CHBrClF, the C-F distance (1.348 Å) is already close to the bond length of CF^+ (1.275 Å). Only the C–H distance (1.083 Å) is closer to the CH^+ bond length (1.137 Å). However, on fivefold ionization the proton gets assigned a charge. Thus it is strongly repelled by the remaining four charges on the molecular fragment. The formation of a CH^+ bond is then not possible. Hence if one performs a BOMD simulation at $T = 0$ K, then the smallest structural rearrangement in order to form a molecular fragment is the one through channel 1 and hence we see the corresponding fragments in all BOMD simulations at 0 K. According to Fig. 11, the appearance energy of CH^+ fragment under fivefold ionization should be roughly 10 eV above the lowest energy channel. By increasing the temperature in the BOMD simulation (performed by assigning random momenta to each atom in the system such that the internal kinetic energy corresponds to the temperature while the total linear momentum remains zero) it is possible to open also other fragmentation channels shown in Fig. 11. Nevertheless the initial linear momenta have to be such that a distorted molecular structure is reached before the Coulomb repulsion leads to explosion. In our simulations the distributed kinetic energy was of the order of a few electron volt. As a rough estimate one can use the C–F stretching mode: The vibrational mode of $C–F^+$ has a wavenumber of ca. $1786\,\mathrm{cm}^{-1}$. Stretching the bond by about ≈ 0.5 Å amounts to a change in internal energy of 3.45 eV or 39986 K. Since the momenta are distributed over all atoms, we would need here to set the excess internal energy higher, e.g. to 55000 K.

At this point it has to be noted that our purely DFT based study of BOMD trajectories neglects electronically excited states, which are likely created by the high energy photons or intense laser fields used for the ionization of the gas

Table 8: Fragment composition, partial charges from Mulliken analysis and KER following from BOMD trajectories. The values are calculated from the last frame of the trajectories after Coulomb explosions of CHBrClF in different electronic charge and spin states. For comparison: The KER of 5 point charges with masses of the atoms C, H, Br, Cl and F in the arrangement of the neutral molecule is 74.15 eV.

Total charge	Spin state	E_{kin}/eV	Fragments
+5	doublet	43.1	H^+, $Cl^{1.3+}$, $Br^{1.6+}$, $CF^{1.1+}$
+5	quartet	39.9	H^+, Br^{2+}, $CClF^{2+}$

phase molecules in the experiment. For these states, multi-reference methods are typically necessary in order to describe the relevant BO hyper surfaces and wave packet propagation on multi-state surfaces might be important. Due to the plurality of possible fragmentation pathways, we did not pursue this in the current stage of our work.

3.3 Simulation of a COLTRIMS experiment

In the final stage of this work we show that classical and BOMD trajectories can be used to simulate a COLTRIMS experiment.

The charge and fragment arrangement from the last frame of the pentacation's Coulomb explosion trajectory on the doublet electronic spin PES were used in order to simulate a COLTRIMS detector signal. The fictitious detector signal was calculated by propagation of point charges within an external electric field. Since the point charges in the last frame of the BOMD trajectory were rarely integer, they were rounded to the nearest integer value. Bauer and Grimme provide in Ref. [30] a more sophisticated recipe for the calculation of fragment charges. However, our main interest lies in a qualitatively correct trajectory rather than in a reproduction of statistical signal intensities. Our calculations showed only a negligible dependence on the assigned charges.

The TOF data and event positions were then used in order to estimate the linear momenta at the beginning of the Coulomb explosion (corresponds to the beginning of the BOMD trajectory), $t = 0$ (Tab. 10, see supplementary material for the working equations). This procedure corresponds to the momentum reconstruction step as it was performed in the analysis of the experiments in Ref. [18, 19].

Tab. 9 shows the asymptotic linear momenta of the various fragments as taken from the last frame of the BOMD trajectory and reconstructed from the fictitious detector signals after propagation in the electric field. Although the

Table 9: Linear momenta after fragmentation of pentacationic CHBrClF. The linear momenta are calculated from the last frame of a BOMD trajectory, i.e. from the velocity (in atomic units) that each particle possesses in the last frame and from the atomic mass in units of the electron mass, and from the fictitious detector signals, i.e. from the TOF and the event positions of every fragment. However, as both sets are identical, we show only one. The momenta are rotated such that $\vec{p}_{\mathrm{Cl}}$ is collinear with the unit vector $\vec{e}_z$. The KER is calculated by $T_i = \frac{|\vec{p}_i|^2}{2m_i}$.

Fragment	$p_x/\frac{\hbar}{a_0}$	$p_y/\frac{\hbar}{a_0}$	$p_z/\frac{\hbar}{a_0}$	KER/eV
H^+	−15.3	50.2	−4.8	20.6
Cl^+	0.0	0.0	215.2	9.7
Br^{2+}	174.8	0.0	−173.2	5.7
CF^+	−159.5	−50.2	−37.2	7.1

former corresponds to the trajectory frame at $t = 29920\frac{\hbar}{E_\mathrm{h}}$ (≈ 720 fs) whereas

the latter has been reconstructed at $t = 0\frac{\hbar}{E_\mathrm{h}}$, they are in excellent agreement with each other. Hence the use of purely classical momentum reconstruction equations despite the quantum nature of the first few femtoseconds during the Coulomb explosion seems justified. Fig. 14 shows a graphical representation of the same findings: The neutral CHBrClF molecule is shown in its ground state structure. Additionally the reconstructed momentum vectors are represented as arrows in the color of the corresponding fragment whereas the normal plane of the asymptotic linear momentum vectors, taken from the last frame of the doublet electronic spin BOMD trajectory of the pentacation, is represented as small plate in the color of the fragment. The dark green linear momenta correspond to the CF^+ fragment. Furthermore we investigated the effect of the type of simulation (BOMD or classical MD [CMD]) on the total KER and the $\cos(\theta)$ which is used in order to identify the enantiomer, (Tab. 10). BOMD and classical MD show very similar values for $\cos(\theta_\mathrm{rec})$ calculated from the reconstructed linear momenta. The differences in the KERs is explained by the neglect of bonding potentials in the CMD simulation.

Table 10: Comparison of an BOMD with an CMD simulation for the breakup of CHBrClF into CF^+, H^+, Cl^+ and Br^{2+}. The cosine was calculated using the linear momenta of Br, Cl and H: $\cos(\theta) = \frac{\vec{p}_\mathrm{Br}}{|\vec{p}_\mathrm{Br}|} \cdot \left(\frac{\vec{p}_\mathrm{Cl} \times \vec{p}_\mathrm{H}}{|\vec{p}_\mathrm{Cl} \times \vec{p}_\mathrm{H}|}\right)$

Method	$\cos(\theta_\mathrm{rec})$	$\mathrm{KER}_\mathrm{rec}$/eV
BOMD	−0.677186	42.5
CMD	−0.666205	54.2

4 Conclusions

In this study, we give a comprehensive assessement of multiply charged CHBrClF. First, we searched for stable multiply charged ions of CHBrClF. Several of those located were achiral. As formation of these ions during COLTRIMS experiments could lead to the loss of information about the absolute configuration of the chiral precursor molecule, we pursued investigation of those structures further.

We could find planar structures for the singly, doubly and triply ionized CHBrClF. In the case of $CHBrClF^{+\cdot}$ the tetrahedral structure is the energetically most favorable and all other local minima found for the monocation lie energetically above the vertically ionized structure. All found $CHBrClF^{2+}$ isomers, except 2C, are lower in energy than the tetrahedral structure, but show dissociation energies smaller than the excess energy of the vertically ionized molecule by about 1 to 2 eV. The isomers of $CHBrClF^{3+\cdot}$ have dissociation paths which are exothermic by more than 5 eV.

The dissociation of $CHBrClF^{2+}$ and $CHBrClF^{3+\cdot}$ led only to two fragments in all investigated BOMD trajectories. Such events cannot be used to determine

the absolute configuration as one needs at least three non-coplanar linear momentum vectors. We could show that the Coulomb explosion of $CHBrClF^{3+\cdot}$ generates enough internal energy in the molecular fragment in order to overcome further dissociation barriers.

We conclude that inversion of the absolute configuration of CHBrClF during the COLTRIMS experiment is unlikely. Not only are species of higher symmetry meta- or unstable, but for the higher ionization states, which are more likely to give enough fragments, we could not find any stable planar conformers. Sequential ionization and fragmentation, as already mentioned in Ref. [18], could still lead to an inversion of absolute configuration or to wrong assignments due to fragment rotations.

In this study we established the use of a theoretical swiss army knife consisting of quantum chemical methods, structure optimization, molecular dynamics simulations and statistical rate theory. With this at hand, we are able to list all possible signals in CEI experiments. Despite numerical limitations, we can preselect important channels by structure and energy, estimate their relative strengths qualitatively and predict their usability for the determination of absolute configuration. Furthermore, we have the tools available to scan the huge amounts of experimental data for events that previously went unnoticed. We also provided formulas for the quick evaluation of branching ratios that require only knowledge of the fragments' masses that can be of aid in the design of future experiments.

Our study shows that the gas phase polycations of CHBrClF form a whole library of metastable achiral as well as chiral structures that are worthwhile to investigate with respect to their properties in multiion and chirality research.

5 Computational Section

Equilibrium structures as well as their harmonic vibrational frequencies were computed with the quantum chemistry program Turbomole 6.5. [31] The structures of the CHBrClF ions were optimized at the density functional level of theory (DFT) with the B3LYP functional and cc-pVTZ-F12 basis set for all atoms except Br, where we used cc-pVTZ-PP-F12 [32] and a fully-relativistic small-core Stuttgart-Cologne pseudo potential with 10 core electrons [33]. The energies in the self-consistent-field calculations were converged to 10^{-9} E_h. The root mean square of the density matrix was converged to 10^{-7}. The standard reference grid for DFT was used. Stationary points were confirmed by harmonic vibrational frequency calculations. Different starting structures with differing bonding situations were used in order to locate local minima on the PES. The total charges of the molecular systems ranged from +1 up to +5. For the structures obtained with the density functional calculations we also calculated single point energies on the coupled cluster level of theory (CCSD(T)-F12b) with Molpro. [34, 35] Here, we used basis sets, pseudo potential and auxiliary basis sets as described in Ref. [32].

Calculations regarding transition structures with the synchronous Transit-

guided Quasi-Newton method (QST2) [36, 37] and the Berny algorithm [38] and calculations of internal reaction coordinate (IRC) [39] were performed with Gaussian 09 [40] on the density functional level of theory with the B3LYP functional and a Def2SVP basis set [41]. A candidate for a transition structure was first interpolated from two starting structures with QST2. The obtained structure was reoptimized using the Berny algorithm. In order to check if we found the desired transition structure, forward and backward IRC calculations were done with a step size of 0.1 a_0 and a maximum number of points of 20 in each direction.

BOMD simulations of Coulomb explosions were done using Turbomole (versions 6.3.1 and 7.1) [42, 43]. Micro canonical trajectories were generated from $t = 0\ \frac{\hbar}{E_\mathrm{h}}$ till $t = 30000\ \frac{\hbar}{E_\mathrm{h}}$ using a time step of $80\ \frac{\hbar}{E_\mathrm{h}}$. Shorter time steps showed no difference in the trajectories, while time steps larger than $100\ \frac{\hbar}{E_\mathrm{h}}$ introduced artifacts. The relative drift of the total energy was of the order 10^{-5} whereas the absolute drift was of the order $10^{-2}\ E_\mathrm{h}$. All single point calculations were performed using DFT (B3LYP) and a Def2SVP basis on all atoms. Additionally, a small-core pseudo potential of the Stuttgart *multi-electron-Dirac-Fock-fit*-type [33] replaced the 10 innermost core electrons of Br. The BOMD trajectories were fed into a Mathematica 11 [44] simulation of the CEI experiment described in Ref. [18] by Pitzer et al. Furthermore, the Mathematica program was used to obtain a classical MD trajectory for comparison to the BOMD results. Local computer programs were written for the calculation of branching ratios and the listing of possible PiXCo signals.

6 Experimental Section

Information on the COLTRIMS technique can be found in various reviews. [16, 45] We used a circular polarized laser beam with a wavelength of 800 nm and a pulse length of 40 fs. The laser power was 2.5 W, the repetition rate 100 kHz and the focal length 6 mm. The spectrometer had a length of 21 cm with an acceleration voltage of $116\ \frac{\mathrm{V}}{\mathrm{cm}}$. The MCP detector had a diameter of 80 mm. The probe was a racemic mixture of CHBrClF. Synchrotron data are obtained from the data set described in Ref. [19].

7 Acknowledgement

This work was supported by the State Initiative for the Development of Scientific and Economic Excellence LOEWE in the LOEWE-Focus ELCH. Sabrina Marquardt thanks the *Studienstiftung des Deutschen Volkes* and Sebastian Marquardt the *Fonds der Chemischen Industrie* for financial support. We thank the Center for Scientific Computing (CSC) Frankfurt for providing computer time.

8 Keywords

Multiply charged ions, Absolute configuration, COLTRIMS, chirality

References

[1] L. Pauling, *J. Chem. Phys.* **1933**, *1*(1), 56–59.

[2] M. Guilhaus, A. G. Brenton, J. H. Beynon, M. Rabrenovic, P. V. R. Schleyer, *J. Phys. B: Atomic and Molecular Physics* **1984**, *17*(17), L605.

[3] A. Belkacem, E. P. Kanter, R. E. Mitchell, Z. Vager, B. J. Zabransky, *Phys. Rev. Lett.* **1989**, *63*, 2555–2558.

[4] K. Vékey, *Mass Spectrom. Rev.* **1995**, *14*(3), 195–225.

[5] D. Schröder, H. Schwarz, *J. Phys. Chem. A* **1999**, *103*(37), 7385–7394.

[6] D. Mathur, *Phys. Rep.* **1993**, *225*(4), 193 – 272.

[7] T. Yatsuhashi, K. Toyota, N. Mitsubayashi, et al., *ChemPhysChem* **2016**, *17*(19), 2977–2981.

[8] K. Lammertsma, P. v. R. Schleyer, H. Schwarz, *Angew. Chem.* **1989**, *101*(10), 1313–1335.

[9] F. Maquin, D. Stahl, A. Sawaryn, et al., *J . Chem. Soc. Chem. Commun.* **1984**, 504.

[10] Y. Apeloig, M. Karni, B. Ciommer, G. Frenking, H. Schwarz, *Int. J. Mass. Spectrom. Ion Proc.* **1984**, *55*, 319.

[11] C. Guenat, F. Maquin, D. Stahl, W. Koch, H. Schwarz, *Int. J. Mass Spectrom. Ion Processes* **1985**, *63*(2), 265 – 272.

[12] J. Roithová, J. Žabka, Z. Herman, et al., *J. Phys. Chem. A* **2006**, *110*(20), 6447–6453, pMID: 16706400.

[13] D. Duflot, J.-M. Robbe, J.-P. Flament, *Int. J. Mass Spectrom. Ion Processes* **1997**, *171*(1), 215 – 230.

[14] R. P. Grant, S. R. Andrews, D. E. Parry, F. M. Harris, *Rapid Commun. Mass Spectrom.* **1998**, *12*(7), 382–388.

[15] R. Grant, F. Harris, D. Parry, *Int. J. Mass Spectrom.* **1999**, *192*(1-3), 111–123.

[16] R. Dörner, V. Mergel, O. Jagutzki, et al., *Phys. Rep.* **2000**, *330*(2–3), 95–192.

[17] J. Ullrich, R. Moshammer, A. Dorn, et al., *Rep. Prog. Phys.* **2003**, *66*(9), 1463–1545.

[18] M. Pitzer, M. Kunitski, A. S. Johnson, et al., *Science* **2013**, *341*(6150), 1096–1100.

[19] M. Pitzer, G. Kastirke, M. Kunitski, et al., *ChemPhysChem* **2016**, *17*, 2465–2472.

[20] Y.-L. He, L. Wang, *Struct. Chem.* **2009**, *20*(3), 461–479.

[21] J. M. Anglada, J. M. Bofill, S. Olivella, A. Solé, *J. Phys. Chem. A* **1998**, *102*(19), 3398–3406.

[22] S. Leach, M. Schwell, D. Talbi, et al., *Chem. Phys.* **2003**, *286*(1), 15 – 43.

[23] E. Rudberg, T. Brinck, *Chem. Phys.* **2004**, *302*, 217–228.

[24] H.-C. Jankowiak, J. Stuber, R. Berger, *J. Chem. Phys.* **2007**, *127*, 234101.

[25] C. Wang, D. Ding, M. Okunishi, et al., *Chem. Phys. Lett.* **2010**, *496*(1-3), 32 – 35.

[26] C. Wang, B. Wang, M. Okunishi, et al., *Chem. Phys.* **2014**, *430*, 40 – 46.

[27] G. Frenking, W. Koch, *J. Molec. Struct. THEOCHEM* **1984**, *110*, 49.

[28] W. Koch, G. Frenking, J. Gauss, D. Cremer, *J. Am. Chem. Soc.* **1986**, *108*, 5808.

[29] A. Kramida, Yu. Ralchenko, J. Reader, and NIST ASD Team, NIST Atomic Spectra Database (ver. 5.3), [Online]. Available: `http://physics.nist.gov/asd` [2017, July 14]. National Institute of Standards and Technology, Gaithersburg, MD., **2015**.

[30] C. A. Bauer, S. Grimme, *J. Phys. Chem. A* **2016**, *120*(21), 3755–3766.

[31] TURBOMOLE V6.5 2013, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

[32] J. G. Hill, K. A. Peterson, *J. Chem. Phys.* **2014**, *141*(9), 094106.

[33] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, M. Dolg, *J. Chem. Phys.* **2003**, *119*(21), 11113–11123.

[34] H.-J. Werner, P. J. Knowles, G. Knizia, et al., MOLPRO, version 2012.1, a package of ab initio programs, **2012**.

[35] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.

[36] C. Peng, H. Bernhard Schlegel, *Isr. J. Chem.* **1993**, *33*(4), 449–454.

[37] C. Peng, P. Y. Ayala, H. B. Schlegel, M. J. Frisch, *J. Comput. Chem.* **1996**, *17*(1), 49–56.

[38] X. Li, M. J. Frisch, *J. Chem. Theory Comput.* **2006**, *2*(3), 835–839, pMID: 26626690.

[39] K. Fukui, *Acc. Chem. Res.* **1981**, *14*(12), 363–368.

[40] M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al., Gaussian 09 Revision E.01, gaussian Inc. Wallingford CT 2009.

[41] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[42] TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

[43] TURBOMOLE V7.1 2016, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

[44] Wolfram Research, Inc., Mathematica 11.0, **2016**.

[45] M. Pitzer, *Journal of Physics B: Atomic, Molecular and Optical Physics* **2017**, *50*(15), 153001.

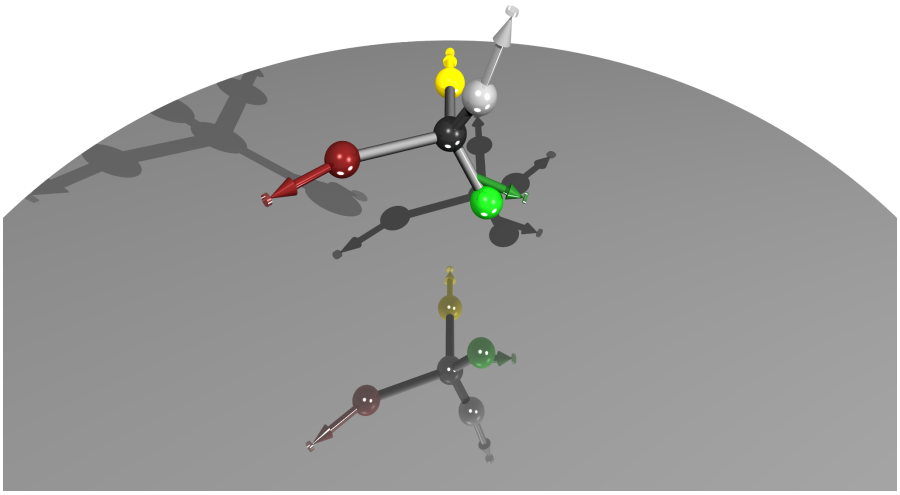


Figure 14: Graphical evaluation of reconstructed linear momenta. An image of the molecule CHBrClF in its neutral ground state structure is shown with the atoms having the colors C: gray, H: white, Br: brown, Cl: yellow, F: green and CF united atom: dark green. The image of the molecule is superimposed with the reconstructed Coulomb explosion momenta, here represented as arrows. The small disks represent the centered normal planes of the asymptotic momenta that are obtained directly from the last frame of the BOMD trajectory of pentacation's Coulomb explosion. The angles between the reconstructed momentum vectors and their corresponding references are: $\angle_{\mathrm{CF}} = 0.90°, \angle_{\mathrm{H}} = 0.82°, \angle_{\mathrm{Cl}} = 0°, \angle_{\mathrm{Br}} = 0.05°$. The momentum vector of H is scaled by a factor of 5.

5. Conclusion and Outlook

Major effects of relativity on the valence electrons originate from the core electrons, which is why modern pseudo potentials (PP) are often employed to mimic these effects in non-relativistic calculations. But the pseudo orbitals (PO) obtained through these calculations are neither sufficient for the calculation of core properties, as, e.g., the hyperfine coupling tensor or the electron density at the nucleus, nor can they serve as adequate starting orbitals for more refined, relativistic all-electron calculations. Hence this thesis targeted the development of a method to reconstruct all-electron molecular orbitals (MO) from their PO counterparts. Although I did not present a rigorous and in principle exact approach of the reconstruction, I devised a scheme that allows the implementation of an in principle automatic orbital reconstruction, similar to parametrized model Hamiltonians used in modern quantum chemistry software packages. I showed the applicability and versatility of my approach on numerous examples.

I started with the discussion of various HF-type model systems and described the analytic connection of total state energies and HF orbital energies in this context. From this starting point, a transformation from POs to MOs was formulated using the apparent shape consistency of POs and applied to various atomic and molecular systems. Only overlap integrals and a Gram–Schmidt orthogonalization are needed to construct the transformation and hence the method scales approximately cubically with the number of basis functions. The chosen reconstruction method likely represents a simple, but efficient method to reintroduce the nodal structure into POs. The reconstructed wave functions were used to calculate a wealth of core- and valence-like molecular properties. The results for these properties were in very good agreement with AE reference calculations, underlining the predictive power of the approach developed in this work. It was shown that numerous parameters influence the quality of reconstructed orbitals. Perhaps the most important parameter is the number virtual orbitals used to span the valence part of the PO: Although expectation values of core-like properties do not change much, the electric dipole moment (EDM) profits tremendously from using low-energy virtual orbitals in the reconstruction. This is especially important when applying the method as a zeroth order approximation to more sophisticated, relativistic methods, when satisfactory

starting orbitals are required to reduce the number of subsequent self-consistent field (SCF) iterations. Another important parameter is the basis in which the relativistic AO are expanded. The basis set must be chosen carefully, or else the orbitals obtained through reconstruction are ill-defined. However, the effort involved in constructing a proper basis set is limited, as one can precalculated a library of atomic calculations using large even-tempered basis sets. A protocol for the design of a corresponding basis set has been given and can in principle be used to obtain a reconstruction scheme for the entire periodic table including super-heavy elements.

The reconstruction method was successfully applied to obtain the relativistic scaling of hyperfine coupling constants of alkaline earth mono-fluorides. Reference and reconstructed values are in good agreement to each other. However, the break-down of an analytically derived relativistic enhancement factor for atomic numbers beyond 118 was discovered in the process. Despite the knowledge of this failure for almost 10 years, no final solution has been proposed to resolve the shortcoming. I could show that an empirically derived relativistic enhancement factor dating back to Fermi and Segrè eliminates the discontinuity at $Z = 118$, indicating applicability of the formula for atomic numbers till $Z \approx 125$.

Also, electronic structure parameters of nuclear spin-dependent parity violating interactions could be calculated in good agreement with reference results for the laser-coolable molecule $CH_3OSr^{+\bullet}$ and symmetry-independence of the method was shown using the examples of tetrahedral $PbCl_4$ and asymmetric NWHClF. I used derivatives of NWHFCl, namely NWHFBr and NWHFI, to demonstrate the power of my method in the accurate prediction of substituent effects on core-like properties.

Future investigations should take up the thread of my work to provide a more rigorous proof for the shape-consistency of energy-consistent PO. This connection would render energy- and shape-consistent PP theories as special cases of one unified theory, which likely explains some of the approximations imposed on this work.

In the field of chirality research, I contributed a molecular dynamics algorithm to study the (classical) dynamics of Coulomb explosions. Born–Oppenheimer molecular dynamics (BOMD) simulations were used to study the effects of the electronic structure on the dynamics of the Coulomb explosion. For identical fragment compositions, classical MD and BOMD trajectories showed only minor differences. Coulomb explosions imaging (CEI) algorithms were implemented to reconstruct the handedness of chiral molecules from time-of-flight data and fragment detection positions and was successfully applied to CMD and BOMD trajectories. Furthermore, the program was used to cross-check the experimental

analysis by providing data equivalent to a real experiment. Thus we could ensure that no random errors tamper with the analysis of CEI experiments.

In the future, the molecular dynamics (MD) simulations (BOMD or CMD) should be carefully adapted to properly account for a correct charge distribution during the Coulomb explosion. In the case of the CMD algorithm, first attempts have been made but remain preliminary. In BOMD simulations, fragments can exchange charges in the early state of the Coulomb explosion. The limiting charges are, however, fractional and thus physically not possible. As we did not expect significant changes due to the charges, I provided no solution to this problem, but indicated proposals from other works.

Planar constitutional isomers of CHBrClF were investigated with respect to their gas phase formation probability using BOMD. Additionally, it was shown that the formation of five atomic fragment ions from CHBrClF competes with the formation of one molecular fragment plus three atomic fragment ions. Although the energy gain by the formation of, e.g., a CF^+ and a Br^{2+} fragment is larger than that of forming five singly charged atomic fragments, corresponding signals in photoion coincidence measurements suggested predominance of the former channel over the latter. Branching ratios of competing reaction channels obtained through statistical means showed that kinetic effects can explain this predominance: A statistical model was built from numerical partition functions including kinetic, rotational, vibrational and electronic degrees of freedom and used in an RRKM-like picture to calculate the ratio of reaction rates. Analytical partition functions of quantum theoretical model systems were used to derive an analytic formula for branching ratios. Only atomic masses enter this model, rendering it especially useful in experimental applications. However, up to now the model treats only the lowest electronic state as a degree of freedom, biasing especially the branching ratio in the high-energy regime. Including electronic excitations could enhance the applicability of the model over a larger energy range.

A library of molecular CHBrClF fragment ions was used to explain many features of several photoion coincidence measurements. However, high photon energies and high photon densities open a multitude of reaction channels in the experiment. A full description of all possible (and experimentally available) reaction paths demands resilience, both on computer and human resources. Similar studies to the branching ratios of $CHBrClF^{5+\bullet}$ could be performed for $CHBrClF^{4+\bullet}$. However, it seems more promising to identify a model system to directly determine the absolute configuration of molecules involving large molecular fragments. Ideally, it should be easily obtainable with arbitrary enantiomeric excess. However, treating such a molecule efficiently within the models presented

herein may require additional information from other statistical theories: The implementation of the statistical adiabatic channel model in a computer program by colleagues paves the way towards estimates of dissociation reaction rates. Reducing the cost of this model is paramount to identify the most likely dissociation paths of large molecules at experimental conditions. Thus one could avoid expensive BOMD calculations in favor of cheaper CMD simulations, perhaps employing a united atom model for molecular fragments.

6. Zusammenfassung

Die vorliegende Arbeit beschreibt die Berechnung kernnaher molekularer Eigenschaften unter der Verwendung von Pseudopotentialen (PP). Dazu wurde eine Methode zur Rekonstruktion der Allelektronenmolekülorbitale (MO) aus energiekonsistent generierten Pseudoorbitalen (PO) entwickelt. Das erlaubt die effiziente Berechnung skalarrelativistischer MO aus einer Kombination von Allelektronenatomorbitalen (AO) und PO des Moleküls. Verglichen mit AO, haben die PO eine unterschiedliche Knotenstruktur und Amplitude in Kernnähe. Sie können daher zur Berechnung physikalisch sinnvoller Erwartungswerte von Eigenschaften, deren Hauptbeitrag durch die Orbitalbeiträge am Kernort bestimmt wird, nicht eingesetzt werden. Die entwickelte, einfach implementierbare Methode führt die Knotenstruktur wieder in die PO ein und erlaubt daher, mit nichtrelativistischen, effizient berechneten PO gute Abschätzungen für diese Erwartungswerte zu erhalten.

Aus den Gleichungen der Methode der zustandsgemittelten, selbstkonsistenten Felder mit Multikonfigurationsansatz (MCSCF), welche zur Erzeugung energiekonsistenter Pseudopotentiale genutzt wird, wurde in dieser Arbeit ein theoretischer Zusammenhang zwischen dem so erzeugten elektronischen Spektrum des Mehrelektronenhamiltonoperators und dem elektronischen Spektrum des effektiven Einelektronenfockoperators anhand mehrerer Modellsysteme aufgezeigt. Diese Beziehung wurde genutzt, um die *Formkonsistenz* energiekonsistenter PO zu begründen. Formkonsistenz beschreibt die Eigenschaft, dass ein PO außerhalb eines bestimmten (Kern-)Radius r_{core} dem entsprechenden MO exakt folgt. Im hier dargestellten Fall gilt dieser Zusammenhang in sehr guter Näherung. Die Formkonsistenz der PO bedingt, dass im Falle eines Atoms ein PO immer einen dominanten Atomorbitalpartner besitzt. Die Überlappung zwischen einem PO und dem zugehörigen AO ist nahe eins, da sie hauptsächlich durch den Bereich außerhalb von r_{core} bestimmt wird. Entwickelt man daher bspw. ein 5s PO in besetzten AO, so trägt vorwiegend das 5s AO bei. Das PO selbst enthält auch Anteile hochenergetischer, unbesetzter Orbitale, die zur Glättung der Funktion nahe des Atomkerns notwendig sind. Folglich wurde gezeigt, dass die Vernachlässigung der entsprechenden Terme die Knotenstruktur wieder einführt. Für MO ist dieser Ansatz jedoch nicht zielführend, da unbesetzte AOs

zur Beschreibung der Bindungspolarisierung notwendig sind. Durch Wahl eines Mischansatzes, d.h. Beimischung niederenergetischer, unbesetzter AO zum PO, konnte die AO-Polarisierung nahezu vollständig dargestellt werden, bei gleichzeitiger Rekonstruktion der Orbitale im Kernbereich. Entscheidend ist bei diesem Ansatz, dass er nur Integrale erhält, die in üblichen Quantenchemieprogrammen verfügbar sind und deren Berechnungszeit quadratisch mit der Anzahl der Basisfunktionen (N^2) steigt. Der teuerste Schritt im Algorithmus ist eine Orthogonalisierung mit der Gram–Schmidt-Methode, deren Kosten ungefähr mit N^3 steigen. Obwohl dies in dieser Arbeit nicht vorgenommen wurde, ist zu erwarten, dass diese Kosten durch eine Vorsortierung der berücksichtigten Vektoren noch reduziert werden können.

Die Rekonstruktionsmethode wurde an mehreren Beispielen getestet. Für atomare, geschlossenschalige Systeme war die Rekonstruktion immer (numerisch) exakt möglich. Anhand des Moleküls $BaF^{\bullet}$ wurde der Einfluss unterschiedlicher Parameter auf die Rekonstruktion untersucht. Es wurde gezeigt, dass eine Basissatzabhängigkeit der Methode vorliegt. Es gilt sicherzustellen, dass die gewählte Basis zur Berechnung der AO prinzipiell die Pseudoatomorbitale (PAO) exakt darstellen kann. Wird dies nicht berücksichtigt, können wichtige Anteile der Valenzwellenfunktion im Verlauf der Rekonstruktion verloren gehen. Weiterhin wurde analysiert, dass die Anzahl der unbesetzten Orbitale, die zur Darstellung der AO-Polarisierung notwendig ist, beschränkt ist. In den demonstrierten Fällen genügte es, die Orbitale, die zum PP durch die SA-MCSCF Fittingprozedur beitrugen, einzubeziehen.

Durch Rekonstruktion der Erdalkalimonofluoridorbitale wurde gezeigt, dass auch periodische Trends wiedergewonnen werden können. Im diskutierten Fall wurde die Skalierung der Hyperfeinstrukturparameter mit steigender Kernladungszahl diskutiert. Diese sollte im Fall wasserstoffartiger Atome linear verlaufen, sofern man die Hyperfeinstrukturparameter durch einen atomzahlabhängigen, relativistischen Verstärkungsfaktor dividiert. Für lineare Moleküle ergaben sich davon leichte Abweichungen. Die Skalierung der Eigenschaft mit rekonstruierten Wellenfunktionen ist jedoch sehr ähnlich zu der mit Allelektronenreferenzwellenfunktionen berechneten. In diesem Kontext wurde das Fehlverhalten eines typischen relativistischen Verstärkungsfaktors für Kernladungszahlen größer 100 diskutiert. Weiterhin wurde die Anwendbarkeit der Methode auch für paritätsverletzende Eigenschaften gezeigt, deren Hauptbeitrag von der Valenzelektronenwellenfunktion direkt am Kernort erzeugt wird. Am Beispiel des Moleküle $CH_3OSr^{+\bullet}$, $PbCl_4$ und NWHClF wurde gezeigt, dass die Methode unabhängig von der geometrischen Anordnung der Atome auch erfolgreich an größeren Molekülen angewandt werden kann. Abschließend wurde gezeigt, dass sowohl die Rekonstruktion mehrerer schwerer Zentren, als auch

die Vorhersage von Substituenteneffekten mit Hilfe rekonstruierter Wellenfunktionen möglich ist.

Auf dem Gebiet der Erforschung molekularer Chiralität war der Autor an zwei veröffentlichten Arbeiten und einer Arbeit, deren Veröffentlichung in Vorbereitung ist, beteiligt. Mit Hilfe eines Moleküldynamikprogramms zur Simulation von Coulombexplosionen konnte dabei besonders die Phase direkt nach der plötzlichen Ionisierung des Moleküls CHBrClF untersucht werden. Auch die Gleichungen zur Rekonstruktion der Absolutkonfiguration des untersuchten Moleküls aus Flugzeit und Detektionsort der Molekülfragmente wurden implementiert. Weiterhin erzeugte die Simulation einem Experiment ähnliche Daten, wodurch erstmals die Auswertungsroutinen des Experiments mit Daten simulierter, nicht-racemischer Mischungen überprüft werden konnte.

Für die zweite Veröffentlichung wurde das Moleküldynamikprogramm um harmonische und anharmonische Atombindungspotentiale erweitert. Außerdem wurde ein Ladungsaustauschmodell implementiert, mit Hilfe dessen die Ladungsumverteilung während der Coulombexplosion dargestellt werden kann.

In der noch unveröffentlichten Arbeit wurde mit Hilfe von Born–Oppenheimer Moleküldynamiksimulationen und statistischen Modellen die relative Häufigkeit einzelner Signale bei der simultanen Messung einzelner Photoionen nach der Coulombexplosion untersucht und erklärt. Abschließend wurde gezeigt, dass qualitative Aussagen zu Reaktionsverzweigungsverhältnissen über einen Vergleich der Zustandssummen involvierter Freiheitsgrade möglich ist. Mit Hilfe analytischer Gleichungen für die Zustandssummen einfacher Modelle konnte eine einfache Formel angegeben werden, die die Abschätzung des Reaktionsverzweigungsverhältnisses erlaubt.

7. Danksagung

Diese Arbeit wäre ohne die Unterstützung, die ich von vielen Seiten erfahren habe, nicht möglich gewesen. Ich möchte daher die Gelegenheit nutzen, mich bei all jenen zu bedanken.

Mein Dank gilt dem *Fonds der Chemischen Industrie*, dem *Frankfurt Institute for Advanced Studies*, der *Marburg University Research Academy*, der *Gesellschaft Deutscher Chemiker* und dem *Ursula-Kuhlmann-Fonds* für finanzielle und materielle Unterstützung während der Dissertation.

Lieber Robert, vielen Dank dafür, dass Du meine Begeisterung für die Wissenschaften über mittlerweile 7 Jahren gefördert hast und mir immer Freiraum für eigene Entfaltung und Gestaltung gelassen hast.

Lieber Ralf, vielen Dank, dass Du mir immer als Ratgeber zur Seite standest und mir zuhörtest, wenn ich in manch komplexer Situation jemanden zum Zuhören benötigte.

Lieber Reuti, vielen Dank für die tolle Arbeit, die Du jeden Tag leistest. Deine Hilfe ist von unschätzbarem Wert. Und auch wenn ich mich gegen das „SUSE von der Stange“ entschieden habe, hast Du mir bei allen Projekten mit Rat und Tat zur Seite gestanden.

Liebe Annette, Du bist die gute Seele der Gruppe. Vielen Dank für die Unterstützung bei allen „nicht-technischen“ Fragen und Problemen.

Liebe Mira, lieber Konstantin, lieber Steffen, ich habe Euch alle drei während Vertiefungsarbeiten betreut, ihr habt Euch alle für die TC und sogar für unsere Gruppe entschieden und habt Euch für einen Platz bei mir im Büro entschieden. Das geht runter wie Öl! Vielen Dank, dass Ihr tolle Kollegen seid, mit denen ich jede Idee zu jeder Zeit diskutieren konnte.

Liebe Anna, vielen Dank für deinen unbändigen und ansteckenden Optimismus. Ich wünsche Dir, dass Du ihn niemals verlierst.

Dear Diego, thank you for not running away. I really enjoyed our training sessions. Thank you for your support.

Дорогой Евгений, больше спасибо за лекциям по русскому языке. Твои подарки будут держать меня занятым в течение длительного времени.

Lieber Hauke, vielen Dank dass Du dein umfassendes Wissen zu Turbomole immer bereitwillig geteilt und mir so viele Stunden Recherche erspart hast. Vielen Dank für die Unterstützung bei der Lösung vieler kleiner und größerer Probleme im Theoretikeralltag.

Liebe Mama, lieber Papa, vielen Dank, dass Ihr immer für mich da gewesen seid, als es haarig wurde und mich in allen Lebenslagen, vor allem jüngst beim Papawerden, hervorragend beraten habt. Euch verdanke ich alles.

Lieber Chrissi, vielen Dank, dass ich immer auf dich zählen kann. Du bist mein Lieblingsbruder und mein bester Kumpel. Danke, dass es dich gibt.

Liebe Sabrina, vielen Dank für die vielen Stunden und Tage, die Du mit mir außerhalb der Universität meine Arbeit diskutiert hast. Vielen Dank, dass Du mein Leben jeden Tag aufs Neue bereicherst und gemeinsam mit mir alle Höhen und Tiefen erlebst. Vielen Dank, dass aus *Twogether Threegether* werden konnte. Ich liebe Dich.

Liebe Leonie, willkommen in dieser Welt. Dein Beitrag zur Dissertation und meiner Arbeit der letzten Jahre mag klein sein, aber er ist nicht unbedeutend. Du hast alles daran gesetzt, gesund geboren zu werden und es geschafft. Dafür und für jeden Moment, den ich mit Dir verbringen darf, möchte ich Dir danken. Ich wünsche Dir unbändige Energie, Kraft und Ausdauer, um die Welt zum Guten zu verändern und Freunde, die dich dabei unterstützen. Ich liebe Dich.

Bibliography

[1] D. Figgen, T. Saue, P. Schwerdtfeger, *J. Chem. Phys.* **2010**, *132*(23), 234310, DOI:10.1063/1.3439692.

[2] W. Heisenberg, *Z. Phys.* **1925**, *33*(1), 879–893, DOI:10.1007/bf01328377.

[3] M. Born, W. Heisenberg, P. Jordan, *Z. Phys.* **1926**, *35*(8-9), 557–615, DOI:10.1007/bf01379806.

[4] E. Schrödinger, *Ann. Phys.* **1926**, *384*(4), 361–376, DOI:10.1002/andp.19263840404.

[5] E. Schrödinger, *Ann. Phys.* **1926**, *384*(6), 489–527, DOI:10.1002/andp.19263840602.

[6] P. A. M. Dirac, *Proc. R. Soc. Lond. A Math. Phys. Sci.* **1928**, *117*(778), 610–624, DOI:10.1098/rspa.1928.0023.

[7] D. R. Hartree, *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*(01), 89, DOI:10.1017/s0305004100011919.

[8] D. R. Hartree, *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*(01), 111, DOI:10.1017/s0305004100011920.

[9] D. R. Hartree, *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*(03), 426, DOI:10.1017/s0305004100015954.

[10] V. Fock, *Z. Phys.* **1930**, *61*(1-2), 126–148, DOI:10.1007/bf01340294.

[11] J. Slater, *Phys. Rev.* **1951**, *81*(3), 385–390, DOI:10.1103/PhysRev.81.385.

[12] H. Hellmann, *J. Chem. Phys.* **1935**, *3*(1), 61–61, DOI:10.1063/1.1749559.

[13] H. Hellmann, *Acta Phys.Chim. U.R.S.S.* **1935**, *1*(6), 913–940.

[14] P. Gombás, *Z. Phys.* **1935**, *95*(9–10), 687–691, DOI:10.1007/BF01331305.

[15] P. Gombás, *Z. Phys.* **1935**, *94*(7–8), 473–488, DOI:10.1007/BF01330613.

[16] J. C. Phillips, L. Kleinman, *Phys. Rev.* **1959**, *116*, 287–294, DOI:10.1103/PhysRev.116.287.

[17] J. D. Weeks, S. A. Rice, *J. Chem. Phys.* **1968**, *49*(6), 2741–2755, DOI:10.1063/1.1670479.

[18] M. H. Cohen, V. Heine, *Phys. Rev.* **1961**, *122*, 1821–1826, DOI:10.1103/PhysRev.122.1821.

[19] Y. S. Lee, W. C. Ermler, K. S. Pitzer, *J. Chem. Phys.* **1977**, *67*(12), 5861–5876, DOI:http://dx.doi.org/10.1063/1.434793.

[20] L. R. Kahn, P. J. Hay, R. D. Cowan, *J. Chem. Phys.* **1978**, *68*(5), 2386–2397, DOI:10.1063/1.436009.

[21] M. Dolg, H. Stoll, H. Preuss, *J. Chem. Phys.* **1989**, *90*(3), 1730–1734.

[22] M. Dolg, H. Stoll, H. Preuss, R. M. Pitzer, *J. Phys. Chem.* **1993**, *97*(22), 5852–5859, DOI:10.1021/j100124a012.

[23] P. Pyykkö, H. Stoll, in *Chemical Modelling: Applications and Theory Volume 1*, Vol. 1, (Edited by A. Hinchliffe), The Royal Society of Chemistry, **2000**, 239–305, DOI:10.1039/9781847553317-00239.

[24] P. Schwerdtfeger, *ChemPhysChem* **2011**, *12*(17), 3143–3155, DOI:10.1002/cphc.201100387.

[25] X. Cao, M. Dolg, *WIREs Comput. Mol. Sci.* **2011**, *1*(2), 200–210, DOI:10.1002/wcms.28.

[26] M. Dolg, X. Cao, *Chem. Rev.* **2012**, *112*(1), 403–480, DOI:10.1021/cr2001383.

[27] P. Pyykkö, J. P. Desclaux, *Acc. Chem. Res.* **1979**, *12*(8), 276–281, DOI:10.1021/ar50140a002.

[28] A. V. Titov, N. S. Mosyagin, A. N. Petrov, T. A. Isaev, *Int. J. Quantum Chem.* **2005**, *104*(2), 223–239, DOI:10.1002/qua.20418.

[29] G. Schreckenbach, T. Ziegler, *Int. J. Quantum Chem.* **1997**, *61*(6), 899–918, DOI:10.1002/(SICI)1097-461X(1997)61:6<899::AID-QUA3>3.0.CO;2-R.

[30] W. Kutzelnigg, W. Liu, *J. Chem. Phys.* **2005**, *123*(24), 241102, DOI:10.1063/1.2137315.

[31] Q. Sun, W. Liu, Y. Xiao, L. Cheng, *J. Chem. Phys.* **2009**, *131*(8), 081101, DOI:10.1063/1.3216471.

[32] Q. Sun, Y. Xiao, W. Liu, *J. Chem. Phys.* **2012**, *137*(17), 174105, DOI:10.1063/1.4764042.

[33] M. Barysz, A. J. Sadlej, *J. Chem. Phys.* **2002**, *116*(7), 2696–2704, DOI:10.1063/1.1436462.

[34] M. Iliaš, T. Saue, *J. Chem. Phys.* **2007**, *126*(6), 064102, DOI:10.1063/1.2436882.

[35] M. Barysz, Ł. Mentel, J. Leszczyński, *J. Chem. Phys.* **2009**, *130*(16), 164114, DOI:10.1063/1.3119714.

[36] W. Liu, D. Peng, *J. Chem. Phys.* **2006**, *125*(4), 044102, DOI:10.1063/1.2222365.

[37] A. Rutkowski, *J. Phys. B: At., Mol. Opt. Phys.* **1986**, *19*(2), 149, DOI:10.1088/0022-3700/19/2/005.

[38] A. Rutkowski, *J. Phys. B: At., Mol. Opt. Phys.* **1986**, *19*(21), 3431, DOI:10.1088/0022-3700/19/21/011.

[39] A. Rutkowski, *J. Phys. B: At., Mol. Opt. Phys.* **1986**, *19*(21), 3443, DOI:10.1088/0022-3700/19/21/012.

[40] A. Rutkowski, *Phys. Rev. A* **1996**, *53*, 145–151, DOI:10.1103/PhysRevA.53.145.

[41] W. Kutzelnigg, *Z. Phys. D* **1989**, *11*(1), 15–28, DOI:10.1007/BF01436580.

[42] W. Kutzelnigg, *Z. Phys. D* **1990**, *15*(1), 27–50, DOI:10.1007/BF01436910.

[43] S. Stopkowicz, J. Gauss, *J. Chem. Phys.* **2011**, *134*(6), 064114, DOI:10.1063/1.3522766.

[44] W. Schwalbach, S. Stopkowicz, L. Cheng, J. Gauss, *J. Chem. Phys.* **2011**, *135*(19), 194114, DOI:10.1063/1.3659316.

[45] C. Chang, M. Pélissier, P. Durand, *Phys. Scr.* **1986**, *34*(5), 394, DOI:10.1088/0031-8949/34/5/007.

[46] E. van Lenthe, E. J. Baerends, J. G. Snijders, *J. Chem. Phys.* **1993**, *99*(6), 4597–4610, DOI:10.1063/1.466059.

[47] E. van Lenthe, E. J. Baerends, J. G. Snijders, *J. Chem. Phys.* **1994**, *101*(11), 9783–9792, DOI:10.1063/1.467943.

[48] W. R. Daasch, L. E. McMurchie, E. R. Davidson, *Chem. Phys. Lett.* **1981**, *84*(1), 9, DOI:10.1016/0009-2614(81)85358-4.

[49] J. Cioslowski, P. Piskorz, *Chem. Phys. Lett.* **1996**, *255*(4–6), 315–319, DOI:10.1016/0009-2614(96)00404-6.

[50] E. A. Hinds, P. G. H. Sandars, *Phys. Rev. A* **1980**, *21*, 471–479, DOI:10.1103/PhysRevA.21.471.

[51] P. V. Coveney, P. G. H. Sandars, *J. Phys. B: At., Mol. Opt. Phys.* **1983**, *16*(20), 3727, DOI:10.1088/0022-3700/16/20/009.

[52] M. G. Kozlov, V. I. Fomichev, Y. Y. Dmitriev, L. N. Labzovsky, A. V. Titov, *J. Phys. B: At., Mol. Opt. Phys.* **1987**, *20*(19), 4939, DOI:10.1088/0022-3700/20/19/007.

[53] Y. Dmitriev, Y. Khait, M. Kozlov, L. Labzovsky, A. Mitrushenkov, A. Shtoff, A. Titov, *Phys. Lett. A* **1992**, *167*(3), 280–286, DOI:10.1016/0375-9601(92)90206-2.

[54] A. V. Titov, N. S. Mosyagin, V. F. Ezhov, *Phys. Rev. Lett.* **1996**, *77*, 5346–5349, DOI:10.1103/PhysRevLett.77.5346.

[55] M. G. Kozlov, A. V. Titov, N. S. Mosyagin, P. V. Souchko, *Phys. Rev. A* **1997**, *56*, R3326–R3329, DOI:10.1103/PhysRevA.56.R3326.

[56] A. N. Petrov, A. V. Titov, T. A. Isaev, N. S. Mosyagin, D. DeMille, *Phys. Rev. A* **2005**, *72*, 022505, DOI:10.1103/PhysRevA.72.022505.

[57] F. Schwabl, *Quantenmechanik (QM I)*, 7 Ed., Springer Berlin Heidelberg, **2007**, 978-3-54-073674-5.

[58] J. Crassous, C. Chardonnet, T. Saue, P. Schwerdtfeger, *Organic & Biomolecular Chemistry* **2005**, *3*(12), 2218, DOI:10.1039/b504212g.

[59] M. Avalos, R. Babiano, P. Cintas, J. L. Jiménez, J. C. Palacios, *Tetrahedron: Asymmetry* **2000**, *11*(14), 2845–2874, DOI:10.1016/s0957-4166(00)00265-2.

[60] F. Vester, T. L. V. Ulbricht, H. Krauch, *Die Naturwissenschaften* **1959**, *46*(2), 68–68, DOI:10.1007/bf00599091.

[61] G. L. J. A. Rikken, E. Raupach, *Nature* **2000**, *405*(6789), 932–935, DOI:10.1038/35016043.

[62] L. Pasteur, *Comptes Rendus hebdomadaires des Séances de l'Académie des Sciences* **1848**, *26*, 535–538.

[63] J. H. van't Hoff, *Archives néerlandaises des sciences exactes et naturelles* **1874**, *9*, 445–454.

[64] J.-A. L. Bel, *Bull. Soc. Chim. Paris* **1874**, *T22*, 337–347.

[65] J. M. Bijvoet, A. F. Peerdeman, A. J. V. Bommel, *Nature* **1951**, *168*(4268), 271–272, DOI:10.1038/168271a0.

[66] Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen, M. Fujita, *Nature* **2013**, *495*(7442), 461–466, DOI:10.1038/nature11990.

[67] Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen, M. Fujita, *Nature* **2013**, *501*(7466), 262–262, DOI:10.1038/nature12527.

[68] A. Buckingham, *Chem. Phys. Lett.* **2004**, *398*(1-3), 1–5, DOI:10.1016/j.cplett.2004.08.046.

[69] R. Berger, J. Courtieu, R. R. Gil, C. Griesinger, M. Köck, P. Lesot, B. Luy, D. Merlet, A. Navarro-Vázquez, M. Reggelin, U. M. Reinscheid, C. M. Thiele, M. Zweckstetter, *Angew. Chem. Int. Ed.* **2012**, *51*(33), 8388–8391, DOI:10.1002/anie.201107626.

[70] J. P. King, T. F. Sjolander, J. W. Blanchard, *J. Phys. Chem. Lett.* **2017**, *8*(4), 710–714, DOI:10.1021/acs.jpclett.6b02653.

[71] M. Pitzer, M. Kunitski, A. S. Johnson, T. Jahnke, H. Sann, F. Sturm, L. P. H. Schmidt, H. Schmidt-Böcking, R. Dörner, J. Stohner, J. Kiedrowski, M. Reggelin, S. Marquardt, A. Schießer, R. Berger, M. S. Schöffler, *Science* **2013**, *341*(6150), 1096–1100, DOI:10.1126/science.1240362.

[72] M. Pitzer, G. Kastirke, M. Kunitski, P. T. Jahnke, T. Bauer, C. Goihl, F. Trinter, C. Schober, K. Henrichs, J. Becht, S. Zeller, H. Gassert, M. Waitz, A. Kuhlins, H. Sann, F. Sturm, F. Wiegandt, R. Wallauer, L. P. H. Schmidt,

A. S. Johnson, M. Mazenauer, B. Spenger, S. Marquardt, S. Marquardt, H. Schmidt-Böcking, J. Stohner, R. Dörner, M. Schöffler, R. Berger, *ChemPhysChem* **2016**, *17*, 2465–2472, DOI:10.1002/cphc.201501118.

[73] S. Marquardt, S. Marquardt, M. Pitzer, K. Fehre, H. Schmidt-Böcking, R. Dörner, M. S. Schöffler, R. Berger, Theoretical study of gas phase fragmentation of multiply charged bromochlorofluoromethane (CHBrClF), **2018**.

[74] R. McWeeny, *Proc. Roy. Soc. Lond. A* **1959**, *253*(1273), 242–259, DOI:10.1098/rspa.1959.0191.

[75] R. McWeeny, *Rev. Mod. Phys.* **1960**, *32*(2), 335–369, DOI:10.1103/revmodphys.32.335.

[76] K. G. Dyall, *J. Chem. Inf. Comput. Sci.* **2001**, *41*(1), 30–37, DOI:10.1021/ci000048w.

[77] M. Dolg, U. Wedig, H. Stoll, H. Preuss, *J. Chem. Phys.* **1987**, *86*(2), 866–872, DOI:10.1063/1.452288.

[78] H. Hellmann, W. Kassatotschkin, *J. Chem. Phys.* **1936**, *4*(5), 324–325, DOI:10.1063/1.1749851.

[79] H. Hellmann, W. Kassatotschkin, *Acta Phys.Chim. U.R.S.S.* **1936**, *5*(1), 23–44.

[80] H. Yukawa, *Proc. Phys.-Math. Soc. Jpn.* **1935**, *17*, 48–57, DOI:10.11429/ppmsj1919.17.0_48.

[81] M. Kleiner, R. McWeeny, *Chemical Physics Letters* **1973**, *19*(4), 476–479, DOI:http://dx.doi.org/10.1016/0009-2614(73)85129-2.

[82] I. V. Abarenkov, V. Heine, *Philos. Mag.* **1965**, *12*(117), 529–537, DOI:10.1080/14786436508218898.

[83] L. Kahn, W. G. III, *Chem. Phys. Lett.* **1968**, *2*(8), 667 – 670, DOI:10.1016/0009-2614(63)80049-4.

[84] W. H. E. Schwarz, *Theor. Chim. Acta* **1968**, *11*(4), 307–324, DOI:10.1007/bf00568788.

[85] L. R. Kahn, P. Baybutt, D. G. Truhlar, *J. Chem. Phys.* **1976**, *65*(10), 3826–3853, DOI:10.1063/1.432900.

[86] M. E. Schwartz, J. D. Switalski, *J. Chem. Phys.* **1972**, *57*(10), 4125–4131, DOI:10.1063/1.1678039.

[87] J. D. Switalski, M. E. Schwartz, *J. Chem. Phys.* **1975**, *62*(4), 1521–1525, DOI:10.1063/1.430616.

[88] J. Flad, H. Stoll, H. Preuss, *J. Chem. Phys.* **1979**, *71*(7), 3042, DOI:10.1063/1.438710.

[89] H. Preuss, H. Stoll, U. Wedig, T. Krüger, *Int. J. Quantum Chem.* **1981**, *19*(1), 113–130, DOI:10.1002/qua.560190111.

[90] P. Fuentealba, H. Preuss, H. Stoll, L. V. Szentpály, *Chem. Phys. Lett.* **1982**, *89*(5), 418 – 422, DOI:10.1016/0009-2614(82)80012-2.

[91] L. E. McMurchie, E. R. Davidson, *J. Comput. Phys.* **1981**, *44*(2), 289–301, DOI:10.1016/0021-9991(81)90053-X.

[92] K. Dyall, K. Fægri, Jr., *Introduction to Relativistic Quantum Chemistry*, Oxford University Press, **2007**, 978-0-19-514086-6.

[93] B. Metz, Phd thesis, Fakultät Chemie der Universität Stuttgart, **2002**.

[94] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons Ltd., Chichester, **2012**, 978-0-47-196755-2.

[95] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*(3), 165 – 169, DOI:http://dx.doi.org/10.1016/0009-2614(89)85118-8.

[96] K. Gaul, *Techn. rep.*, Theoretische Chemie, Philipps–Universität Marburg, **2015**.

[97] K. Hirao, *J. Chem. Phys.* **1974**, *60*(8), 3215–3222, DOI:10.1063/1.1681510.

[98] R. McWeeny, *Methods of Molecular Quantum Mechanics*, 2 Ed., Academic Press Inc, **1992**, 978-0-12-486552-5.

[99] W. C. Topp, J. J. Hopfield, *Phys. Rev. B* **1973**, *7*(4), 1295–1303, DOI:10.1103/physrevb.7.1295.

[100] H. A. Kramers, in *Hand- und Jahrbuch der Chemischen Physik*, Vol. 1, (Edited by A. Eucken, K. L. Wolf), Akademische Verlagsgesellschaft M. B. H., Leipzig, **1938**.

[101] W. J. Hehre, R. F. Stewart, J. A. Pople, *J. Chem. Phys.* **1969**, *51*(6), 2657–2664, DOI:10.1063/1.1672392.

[102] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, T. L. Windus, *J. Chem. Inf. Model.* **2007**, *47*(3), 1045–1052, DOI:10.1021/ci600510j.

[103] G. H. Golub, C. F. Van Loan, *Matrix Computations*, J. Hopkins Uni. Press, **1996**, 0801854148.

[104] S. Lipschutz, M. L. Lipson, *Linear Algebra*, 4 Ed., Schaum's Outline Series, McGraw-Hill Companies, Inc., **2009**.

[105] D. B. Cook, *Theoretica Chimica Acta* **1972**, *27*(2), 161–163, DOI:10.1007/bf00528160.

[106] E. Baerends, D. Ellis, P. Ros, *Chemical Physics* **1973**, *2*(1), 41–51, DOI:10.1016/0301-0104(73)80059-x.

[107] G. B. Arfken, H. J. Weber, F. E. Harris, *Mathematical Methods for Physicists*, Elsevier LTD, Oxford, **2012**, 978-0-12-384654-9.

[108] W. Weltner, Jr., *Magnetic Atoms and Molecules*, Dover Publication, Inc., **1989**, 978-0-48-666140-7.

[109] I. B. Zel'dovich, *Sov. Phys. J. Exp. Theor. Phys.* **1958**, *6*, 1184.

[110] T. D. Lee, C. N. Yang, *Phys. Rev.* **1956**, *104*(1), 254–258, DOI:10.1103/physrev.104.254.

[111] C. G. Gray, G. Karl, V. A. Novikov, *Am. J. Phys* **2010**, *78*(9), 936–948, DOI:10.1119/1.3427412.

[112] T. A. Isaev, R. Berger, *Phys. Rev. A* **2012**, *86*, 062515, DOI:10.1103/PhysRevA.86.062515.

[113] T. Isaev, R. Berger, *J. Mol. Spectros.* **2014**, *300*, 26–30, DOI:10.1016/j.jms.2014.01.014.

[114] C. van Wüllen, *J. Chem. Phys.* **1998**, *109*(2), 392–399, DOI:10.1063/1.476576.

[115] W. Liu, C. van Wüllen, F. Wang, L. Li, *The Journal of Chemical Physics* **2002**, *116*(9), 3626–3634, DOI:10.1063/1.1446026.

[116] P. Fuentealba, L. von Szentpaly, H. Preuss, H. Stoll, *J. Phys. B* **1985**, *18*(7), 1287–1296, DOI:10.1088/0022-3700/18/7/010.

[117] M. Kaupp, P. v. R. Schleyer, H. Stoll, H. Preuss, *J. Chem. Phys.* **1991**, *94*(2), 1360–1366, DOI:10.1063/1.459993.

[118] I. S. Lim, H. Stoll, P. Schwerdtfeger, *J. Chem. Phys.* **2006**, *124*(3), 034107, DOI:10.1063/1.2148945.

[119] P. J. Hay, W. R. Wadt, *J. Chem. Phys.* **1985**, *82*(1), 299–310, DOI:10.1063/1.448975.

[120] W. J. Stevens, M. Krauss, H. Basch, P. G. Jasien, *Can. J. Chem.* **1992**, *70*(2), 612–630, DOI:10.1139/v92-085.

[121] D. P. Chong, *Can. J. Chem.* **1995**, *73*(1), 79–83, DOI:10.1139/v95-011.

[122] W. Johnson, G. Soff, *At. Data Nucl. Data Tables* **1985**, *33*(3), 405–446, DOI:10.1016/0092-640x(85)90010-5.

[123] E. Cohen, T. Cvitaš, J. Frey, B. Holmström, K. Kuchitsu, R. Marquardt, I. Mills, F. Pavese, J. S. M. Quack, H. Strauss, M. Takami, , A. Thor, *Quantities, Units and Symbols in Physical Chemistry, IUPAC Green Book*, 3 Ed., IUPAC & RSC Publishing, **2008**.

[124] S. Knecht, S. Fux, R. van Meer, L. Visscher, M. Reiher, T. Saue, *Theor. Chem. Acc.* **2011**, *129*(3-5), 631–650, DOI:10.1007/s00214-011-0911-2.

[125] R. Mastalerz, P.-O. Widmark, B. O. Roos, R. Lindh, M. Reiher, *J. Chem. Phys.* **2010**, *133*(14), 144111, DOI:10.1063/1.3491239.

[126] R. Berger, C. van Wüllen, *J. Chem. Phys.* **2005**, *122*(13), 134316, DOI:10.1063/1.1869467.

[127] R. Berger, N. Langermann, C. van Wüllen, *Phys. Rev. A* **2005**, *71*(4), DOI:10.1103/physreva.71.042105.

[128] S. Nahrwold, R. Berger, *J. Chem. Phys.* **2009**, *130*(21), 214101, DOI:http://dx.doi.org/10.1063/1.3103643.

[129] L. B. Knight, W. C. Easley, W. Weltner, M. Wilson, *J. Chem. Phys.* **1971**, *54*(1), 322–329, DOI:10.1063/1.1674610.

[130] C. Ryzlewicz, H.-U. Schütze-Pahlmann, J. Hoeft, T. Törring, *Chem. Phys.* **1982**, *71*(3), 389 – 399, DOI:10.1016/0301-0104(82)85045-3.

[131] N. Stone, *At. Data Nucl. Data Tables* **2005**, *90*(1), 75 – 176, DOI:10.1016/j.adt.2005.04.001.

[132] E. Fermi, E. Segrè, *Z. Phys.* **1933**, *82*(11-12), 729–749, DOI:10.1007/BF01334120.

[133] E. Fermi, E. Segrè, *Mem. Acad. d'Italia* **1933**, *4*, 131–158.

[134] O. P. Sushkov, V. V. Flambaum, I. B. Khriplovich, *Opt. Spectrosc. (USSR)* **1978**, *44*(1), 2.

[135] G. Racah, *Il Nuovo Cimento* **1931**, *8*(1), 178–190, DOI:10.1007/bf02959489.

[136] T. H. Dinh, V. A. Dzuba, V. V. Flambaum, *Phys. Rev. A* **2009**, *80*(4), DOI:10.1103/physreva.80.044502.

[137] T. H. Dinh, A. Dunning, V. A. Dzuba, V. V. Flambaum, *Phys. Rev. A* **2009**, *79*(5), DOI:10.1103/physreva.79.054102.

[138] K. Gaul, S. Marquardt, T. Isaev, R. Berger, *ArXiv e-prints* **2018**, 1805.05494.

[139] I. Kozyryev, L. Baum, K. Matsuda, J. M. Doyle, *ChemPhysChem* **2016**, *17*(22), 3641–3648, DOI:10.1002/cphc.201601051.

[140] T. A. Isaev, R. Berger, *Phys. Rev. Lett.* **2016**, *116*(6), DOI:10.1103/physrevlett.116.063006.

[141] S. Nahrwold, R. Berger, P. Schwerdtfeger, *J. Chem. Phys.* **2014**, *140*(2), 024305, DOI:10.1063/1.4852176.

[142] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, M. Dolg, *J. Chem. Phys.* **2003**, *119*(21), 11113–11123, DOI:10.1063/1.1622924.

[143] Z. Vager, R. Naaman, E. P. Kanter, *Science* **1989**, *244*(4903), 426–431, DOI:10.1126/science.244.4903.426.

[144] R. Dörner, V. Mergel, O. Jagutzki, L. Spielberger, J. Ullrich, R. Moshammer, H. Schmidt-Böcking, *Physics Reports* **2000**, *330*(2-3), 95–192, DOI:10.1016/s0370-1573(99)00109-x.

[145] C. Runge, *Math. Ann.* **1895**, *46*(2), 167–178, DOI:10.1007/BF01446807.

[146] W. Kutta, *Zeitschrift für Mathematik und Physik* **1901**, *46*, 435–453.

[147] H. Murgenau, G. M. Murphy, *Die Mathematik für Physik und Chemie*, Verlag Harri Deutsch, Frankfurt a. Main und Zürich, **1965**.

[148] E. W. Weisstein, Sphere Point Picking. From MathWorld—A Wolfram Web Resource, last visited on 2017-08-02.

[149] P. A. M. Dirac, *Proc. Roy. Soc. Lond. A* **1929**, *123*(792), 714–733.

[150] S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*(8), 1200–1211, DOI:10.1139/p80-159.

[151] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100, DOI:10.1103/PhysRevA.38.3098.

[152] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789, DOI:10.1103/PhysRevB.37.785.

[153] A. D. Becke, *J. Chem. Phys.* **1993**, *98*(7), 5648–5652, DOI:10.1063/1.464913.

[154] J. Eland, J. Sheahan, *Chem. Phys. Lett.* **1994**, *223*(5), 531 – 536, DOI:10.1016/0009-2614(94)00477-3.

[155] M. Dolg, Energy-consistent Pseudopotentials of the Stuttgart/Cologne Group, **2017**.

[156] G. Schaftenaar, J. Noordik, *J. Comput.-Aided Mol. Des.* **2000**, *14*(2), 123–134, DOI:10.1023/a:1008193805436.

[157] G. Schaftenaar, E. Vlieg, G. Vriend, *J. Comput.-Aided Mol. Des.* **2017**, *31*(9), 789–800, DOI:10.1007/s10822-017-0042-5.

[158] A. Nicklass, M. Dolg, H. Stoll, H. Preuss, *J. Chem. Phys.* **1995**, *102*(22), 8942–8952, DOI:10.1063/1.468948.

[159] K. A. Peterson, D. Figgen, M. Dolg, H. Stoll, *J. Chem. Phys.* **2007**, *126*(12), 124101, DOI:10.1063/1.2647019.

[160] W. Küchle, M. Dolg, H. Stoll, H. Preuss, *Molecular Physics* **1991**, *74*(6), 1245–1263, DOI:10.1080/00268979100102941.

[161] T. Hangele, M. Dolg, P. Schwerdtfeger, *J. Chem. Phys.* **2013**, *138*(17), 174113, DOI:10.1063/1.4803148.

[162] B. Metz, H. Stoll, M. Dolg, *J. Chem. Phys.* **2000**, *113*(7), 2563–2569, DOI:10.1063/1.1305880.

[163] D. Figgen, K. A. Peterson, M. Dolg, H. Stoll, *J. Chem. Phys.* **2009**, *130*(16), 164108, DOI:10.1063/1.3119665.

Appendix A.

Basis sets

In this appendix the basis sets are presented that were generated in the course of the work. All of them are of the *even-tempered* type, which means that all exponential coefficients are in a fixed relation to each other.

$$\beta = \frac{\alpha_i}{\alpha_{i+1}} \tag{A.1}$$

Hence the full basis is defined by three parameters per angular momentum: First the initial coefficient α_i, second the ratio β of the coefficients a_i and a_{i+1} and third the number of coefficients N. These values are presented below and formatted $\alpha_0 : \beta[M : N]$, i.e., the basis started with a coefficient of α_0, used a ratio of β and the coefficients M through N were taken for the specific angular momentum basis.

A.1. even-temp

s	500000000 :	2.6	[1 : 26]
p	500000000 :	2.6	[2 : 26]
d	500000000 :	2.6	[12 : 25]
f	500000000 :	2.6	[15 : 22]

A.2. even-temp_ext

s	500000000 :	2.6	[1 : 29]
p	500000000 :	2.6	[2 : 29]
d	500000000 :	2.6	[12 : 28]
f	500000000 :	2.6	[15 : 22]

A.3. even-temp_ext2

s	500000000 :	2.6	[1 : 29]
p	500000000 :	2.6	[2 : 29]
d	500000000 :	2.6	[12 : 28]
f	500000000 :	2.6	[15 : 27]

A.4. even-temp_v2

s	1000000 :	2.6	[1 : 27]
p	100000 :	2.6	[1 : 23]
d	5000 :	2.6	[1 : 18]
f	1000 :	2.6	[1 : 15]

A.5. even-temp_v7

s	3153128230 :	2.6	[1 : 33]
p	3153128230 :	2.6	[1 : 30]
d	5000 :	2.6	[1 : 20]
f	1000 :	2.6	[1 : 15]

Appendix B.

Reconstruction Parameters

The orbital reconstruction depends on several parameters. These parameters are necessary to define for reproducibility of the work. In this chapter, I provide tables of the parameters. Each table has the following layout:

	PP-Atom	AE-Atom
Method	Method PP	Method AE
Basis	`ECPXXMDF`	`even-temp`
PP	ECPXXMDF	-
Core orbitals	-	a'–b'
Valence orbitals	a–b	b'+1–c

The *method* defines the quantum chemical approach taken to converge the orbitals. For the PP-system, this is restricted Hartree–Fock (RHF), restricted open-shell Hartree–Fock (ROHF) or UHF SCF in a NR framework, whereas the AE system is usually treated in by 1c ZORA HF. Only in a few cases I deviated from these defaults.

The *basis* defines the basis sets used in the calculations. An "@" denotes a specific basis for a specific atom. All bases without the @-sign are employed at all atoms which are not mentioned specifically. In general, I used only two distinct types of basis sets: The Ahlrichs basis sets denoted by `def2-XXXX` and the basis sets provided with the Stuttgart PPs denoted by the name of the PP, sometimes appended by a string that denotes the type of split-valence used in the basis (E.g., `_VTZ` means *valence triple zeta*).

In the row *PP* I describe the employed PP by the name used in the database of Dolg. [155] There, Dolg also describes the nomenclature used to define the PP: Each PP is named by a keyword `ECPnXY`, where `n` is the number of electrons contained in the core. If a PP was constructed with a single-valence electron ion as reference, then `X` is `S`. If instead a neutral atom was chosen as reference, then `X` is `M`. The letter `Y` describes the employed Hamiltonian and can be

HF (Hartree–Fock), WB (Wood–Boring), DF (Dirac–Hartree–Fock), DFB (Dirac–Hartree–Fock–Coulomb–Breit and DFQ (Dirac–Hartree–Fock–Coulomb–Breit plus QED contribution). For the AE atom, no PP was employed.

The rows *Core orbitals* and *Valence orbitals* each provide orbital indices. The orbitals were obtained through SCF calculations using the quantum chemistry program package TURBOMOLE [95] and transformed into MOLDEN format [156, 157] by TURBOMOLE. In these files, the orbitals are stored with increasing energy, i.e., orbital 1 denotes the first orbital occuring in the MOLDEN-file and has the lowest energy (e.g., a 1s orbital). The numbers usually denote an orbital range that was included in the reconstruction, followed in consecutive lines by the orbital angular moment labels to which the range corresponds.[1]

In a PP-system, no core orbitals exist, whereas for the AE-system they define the space orthogonal to the space on which the PP valence orbitals are mapped. The PP valence orbitals usually comprise the occupied orbitals of a PP system. In contrast, the AE valence orbitals also contain virtual AO indices (see Ch. 3.2.8, Ch. 3.5 and Ch. 3.6.3.3 for the discussion of the influence of virtual orbitals on the reconstructed orbitals).

For one case ($BaF^{\bullet}$, see Ch. 3.6.3.3) I investigated the dependece of the quality of expectation values calculated from reconstructed orbitals on the number of virtual orbitals included in the reconstruction. As is shown there, core-like properties do not depend much on the number of virtual orbitals, whereas valence-like properties profit considerably from these orbitals. Thus I usually included about 40–60 virtuals from the AE atomic calculation in the reconstruction. A reasonable assumption for the selection of virtuals is to select those virtual orbitals that were ecplicitly correlated in the MCSCF calculation used to derive the PP. We tested the assumption for the atomic cases and achieved excellent agreement with the reference results.

[1] Note I do not label the PP orbitals as I usually reconstructed only occupied orbitals. Note further that the virtuals do not follow the Aufbau principle.

B.1. Atoms

B.1.1. Kr

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP10MDF	even-temp
PP	ECP10MDF [142]	
Core orbitals	-	1–10
		1s–2s
		2p
Valence orbitals	1–26	11–54
	3s–4s	3s–5s
	3p–4p	3p–5p
	3d	3d–4d

	PP-Atom	AE-Atom
Method	NR, HF	1c, ZORA, UHF
Basis	ECP28MWB_VTZ	even-temp
PP	ECP28MWB [158]	
Core orbitals	-	1–28
		1s–3s
		2p–3p
		3d
Valence orbitals	1–8	29–72
	4s	4s–6s
	4p	4p–6p
		4d–5d

According to Ref. [158], the 6p orbital was not contained in the space active for the generation of the PP. However, we include it anyway in the reconstruction, as the active 5d orbital lies energetically higher. Ref. [158] contains no information if the 4d orbitals were treated as active and hence we assumed them to be active.

	PP-Atom	AE-Atom
Method	NR, UHF	NR, UHF
Basis	ECP28MHF	even-temp
PP	ECP28MHF [158]	
Core orbitals	-	1–28 1s–3s 2p–3p 3d
Valence orbitals	1–8 4s 4p	29–72 4s–6s 4p–6p 4d–5d

According to Ref. [158], the 6p orbital was not contained in the space active for the generation of the PP. However, we include it anyway in the reconstruction, as the active 5d orbital lies energetically higher. Ref. [158] contains no information if the 4d orbitals were treated as active and hence we assumed them to be active.

B.1.2. Xe

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	def2-TZVP	even-temp
PP	def2-pp (extended ECP28MDF [142])	
Core orbitals	-	1–28 1s–3s 2p–3p 3d
Valence orbitals	1–26 4s–5s 4p–5p 3d–4d	29–72 4s–6s 4p–6p 4d–5d

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP46MWB_VTZ	even-temp
PP	ECP46MWB [158]	
Core orbitals	-	1–46
		1s–4s
		2p–4p
		3d–4d
Valence orbitals	1–8	47–90
	5s	5s–7s
	5p	5p–7p
		5d–6d

According to Ref. [158], the 7p orbital was not contained in the active space for the generation of the PP. However, we include it anyway in the reconstruction, as the active 6d orbital lies energetically higher. Ref. [158] contains no information if the 5d orbitals are active and hence we assumed them to be active.

	PP-Atom	AE-Atom
Method	NR, UHF	NR, UHF
Basis	ECP46MHF	even-temp
PP	ECP46MHF [158]	
Core orbitals	-	1–46
		1s–4s
		2p–4p
		3d–4d
Valence orbitals	1–8	47–90
	5s	5s–7s
	5p	5p–7p
		5d–6d

According to Ref. [158], the 7p orbital was not contained in the active space for the generation of the PP. However, we include it anyway in the reconstruction, as the active 6d orbital lies energetically higher. Ref. [158] contains no information if the 5d orbitals are active and hence we assumed them to be active.

B.1.3. Rn

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP60MDF_VTZ	even-temp
PP	ECP60MDF [159]	
Core orbitals	-	1–60
		1s–4s
		2p–4p
		3d–4d
		4f
Valence orbitals	1–26	61–104
	5s–6s	5s–7s
	5p–6p	5p–7p
	5d	5d–6d

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP78MWB	even-temp
PP	ECP78MWB [160]	
Core orbitals	-	1–78
		1s–4s
		2p–4p
		3d–4d
		4f
Valence orbitals	1–8	79–104
	6s	5s–7s
	6p	5p–7p
		5d–6d

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP78MHF	even-temp
PP	ECP78MHF [160]	
Core orbitals	-	1–78 1s–4s 2p–4p 3d–4d 4f
Valence orbitals	1–8 6s 6p	79–104 5s–7s 5p–7p 5d–6d

B.1.4. Og

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP92MDF_TZVP	even-temp
PP	ECP92MDFB [161]	
Core orbitals	-	1–92 1s–5s 2p–5p 3d–5d 4f–5f
Valence orbitals	1–26 6s–7s 6p–7p 7d	79–80,95–144 6s–9s 6p–9p 6d–7d

The valence orbital range for the AE atom is not continuous, because the 6s valence orbitals are energetically below the 5f core orbitals included in the PP.

	PP-Atom	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP92MDF_TZVP	even-temp
PP	ECP92MDFQ [161]	
Core orbitals	-	1–92
		1s–5s
		2p–5p
		3d–5d
		4f–5f
Valence orbitals	1–26	79–80,95–144
	6s–7s	6s–9s
	6p–7p	6p–9p
	7d	6d–7d

The valence orbital range for the AE atom is not continuous, because the 6s valence orbitals are energetically below the 5f core orbitals included in the PP.

B.2. Diatomic Molecules

B.2.1. CaF

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP10MDF@Ca,def2-TZVP	even-temp_v7
PP	ECP10MDF@Ca [118]	
Core orbitals	-	1–10
		1s–2s
		2p
Valence orbitals	1–19	11–68
		3s–8s
		3p–4p
		3d–4d
		4f

The ordering of the virtual orbitals does not necessarily follow the Aufbau principle. Furthermore, we preferred to include all virtual until a certain orbital is reached. Although no information is given in Ref. [118], the 4d orbital was likely part of the active space during the PP generation. Thus we included it in the reconstruction.

B.2.2. SrF

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP28MDF@Sr,def2-TZVP	even-temp_v7
PP	ECP28MDF@Sr [118]	
Core orbitals	-	1–28 1s–3s 2p–3p 3d
Valence orbitals	1–19	29–92 4s–9s 4p–6p 4d–5d 4f

g-functions ($l = 4$) were removed from the ECP28MDF.

B.2.3. BaF

	PP-Molecule	AE-Atom
Method	NR, UHF	NR, UHF
Basis	ECP46MHF@Ba,def2-TZVP	even-temp
PP	ECP46MHF@Ba [117]	
Core orbitals	-	1–46 1s–4s 2p–4p 3d–4d
Valence orbitals	1–19	47–90 5s–7s 5p–7p 5d–6d

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP54SDF@Ba,def2-TZVP	even-temp
PP	ECP54SDF@Ba [116]	
Core orbitals	-	1–54 1s–5s 2p–5p 3d–4d
Valence orbitals	1–11	55–98 6s–7s 6p–7p 5d–6d

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP46MWB@Ba,def2-TZVP	even-temp
PP	ECP46MWB@Ba [117]	
Core orbitals	-	1–46
		1s–4s
		2p–4p
		3d–4d
Valence orbitals	1–19	47–98
		5s–8s
		5p–8p
		5d–6d

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP46MDF@Ba,def2-TZVP	even-temp
PP	ECP46MDF@Ba [118]	
Core orbitals	-	1–46
		1s–4s
		2p–4p
		3d–4d
Valence orbitals	1–19	47–98
		5s–8s
		5p–8p
		5d–6d

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP46MDF@Ba,def2-TZVP	even-temp_v7
PP	ECP46MDF@Ba [118]	
Core orbitals	-	1–46
		1s–4s
		2p–4p
		3d–4d
Valence orbitals	1–19	47–110
		5s–10s
		5p–7p
		5d–6d
		4f

g-functions ($l = 4$) were removed from the ECP46MDF.

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	LANL2DZ@Ba,def2-TZVP	even-temp_v7
PP	LANL2DZ@Ba [119]	
Core orbitals	-	1–46 1s–4s 2p–4p 3d–4d
Valence orbitals	1–19	47–102 5s–10s 5p–7p 5d–6d 4f

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	SBKJCp@Ba,def2-TZVP	even-temp_v7
PP	SBKJC@Ba [120]	
Core orbitals	-	1–54 1s–5s 2p–5p 3d–4d
Valence orbitals	1–19	55–98 6s–10s 6p–7p 5d–6d 4f

B.2.4. RaF

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP78MDF@Ra,def2-TZVP	even-temp_v7
PP	ECP78MDF@Ra [118]	
Core orbitals	-	1–78 1s–5s 2p–5p 3d–5d 4f
Valence orbitals	1–19	79–136 6s–11s 6p–7p 5d–7d 5f

g-functions ($l = 4$) were removed from the basis set ECP78MDF.

B.2.5. UbnF

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP92MDFQ_TZVP@Ubn, def2-TZVP	even-temp_v7
PP	ECP92MDF@Ubn [161]	
Core orbitals	-	1–78,81–94 1s–5s 2p–5p 3d–5d 4f–5f
Valence orbitals	1–37	79–80,95–168 6s–12s 6p–8p 5d–8d 6f

g-functions ($l = 4$) were removed from the basis set ECP92MDF.

B.3. Polyatomic Molecules

B.3.1. $CH_3OSr^{\bullet}$

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP28MDF@Sr	even-temp_v7
PP	ECP28MDF@Sr [118]	
Core orbitals	-	1–28 1s–3s 2p–3p 3d
Valence orbitals	1–27	29–60,77–92,109–124,139–148,165–180 4s–8s 4p–8p 4d–8d

In Ref. [118] no information is given on the configurations used in the construction of the PP. It is likely that the configruations were chosen similar to Ba. In this case, the AE AOs basis would have to consist of the orbitals 4s–6s, 4p–6p, 4d–5d. However, the basis sets even-temp_7 and ECP28MDF) were diffuse enough to improve the description of the orbitals close to the carbon atoms significantly in the PP calculation (see Fig. B.1). But the diffuse basis functions contribute to the orbitals 4s–6s and especially 4p–6p, as well as 4d–5d. As the diffuse orbitals seem to contribute significantly to the orbital at the position of C and O, we need to include orbitals up

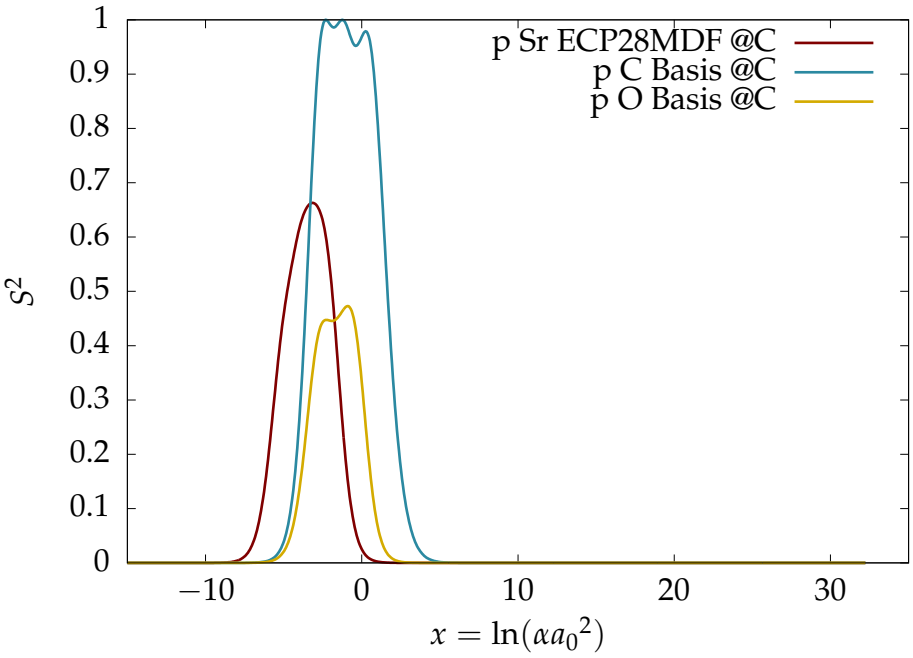


Figure B.1.: Basis set profile of diffuse basis sets in CH_3OSr. The basis set profiles were generated at the position of the carbon atom. Hence the carbon line indicates the basis set profile of the def2-TZVP basis employed for C, while the remaining lines show, which effective exponential coefficient the basis functions of other atoms generate at the carbon atom. As coordinates we used

to 8d in the AE AO basis. A different approach, e.g., to choose a more compact basis set for Sr and more diffuse basis sets for C and O in the PP calculation, sucessfully reconstructed the AE wave function at Sr as well as at C and O. However, as the calculated core-properties were relatively insensitive w.r.t. the chosen AE AO space, we show in this work only the values for the parameters in the table above.

B.3.2. $PbCl_4$

	PP-Molecule	AE-Atom
Method	NR, UHF	1c, ZORA, UHF
Basis	ECP60MDF_VTZ@Pb, def2-TZVP@Cl	even-temp_v7
PP	ECP60MDF@Pb [162]	
Core orbitals	-	1–46, 49–62 1s–4s 2p–4p 4d 4f
Valence orbitals	1–90	47–48, 63–108 5s–9s 5p–7p 5d–6d

By employing fractional occupation numbers in the AE atom, the atomic system was spin-averaged to avoid broken spin-symmetry of the reference orbitals. I.e., each of the 6p spin-orbitals was occupied by $\frac{1}{3}$ electron. As the 5s orbitals were lower in energy than the 4f orbitals, the core orbitals needed to be described by two distinct index ranges.

B.3.3. NWHFCl

	PP-Molecule	AE-Atom
Method	NR, RKS(LDA)	1c, ZORA, RKS(LDA)
Basis	def2-TZVP@W, def2-SVP	even-temp_v7
PP	ECP60MWB@W [163]	
Core orbitals	-	1–23, 28–34 1s–4s 2p–4p 4d 4f
Valence orbitals	1–24	24–27, 35–55 5s–10s 5p–7p 5d–6d

We used a restricted Kohn–Sham ansatz to generate the AOs and MOs, employing fractional occupation numbers for partially filled shells in order to enforce spin-symmetric orbitals. The 5s and 5p orbitals were lower in energy than the 4f, which is why two individually continuous index ranges describe the core orbitals.

B.3.4. NWHFBr

	PP-Molecule	AE-Atom
Method	NR, RKS(LDA)	1c, ZORA, RKS(LDA)
Basis	def2-TZVP@W, def2-SVP	even-temp_v7
PP	ECP10MDF@Br [142]	
Core orbitals	-	1–5 1s–2s 2p
Valence orbitals	1–28	6–19, 23–30 3s–5s 3p–5p 4d

We used a restricted Kohn–Sham ansatz to generate the AOs and MOs, employing fractional occupation numbers for partially filled shells in order to enforce spin-symmetric orbitals.

The tungsten orbitals were reconstructed identical to the NWHFCl system. Only the valence orbitals of the PP-molecule had to be changed in the second reconstruction step, of course.

B.3.5. NWHFI

	PP-Molecule	AE-Atom
Method	NR, RKS(LDA)	1c, ZORA, RKS(LDA)
Basis	def2-TZVP@W, def2-SVP	even-temp_v7
PP	ECP28MDF@I [142]	
Core orbitals	-	1–14 1s–3s 2p–3p 3d
Valence orbitals	1–24	15–28, 32–39 4s–10s 4p–6p 4d–5d

A modified version of ECP28MDF was used, as implemented in TURBOMOLE to avoid g-type projectors.

We used a restricted Kohn–Sham ansatz to generate the AOs and MOs, employing fractional occupation numbers for partially filled shells in order to enforce spin-symmetric orbitals. The 5s and 5p orbitals were lower in energy than the 4f, which is why two individually continuous index ranges describe the core orbitals.

The tungsten orbitals were reconstructed identical to the NWHFCl system. Only the valence orbitals of the PP-molecule had to be changed in the second reconstruction step, of course.

Appendix C.

Reconstruction Data

Table C.1.: Core and valence properties of NWHClF. Expectation values are shown which were calculated from reconstructed and all-electron orbitals. The orbital reconstruction was performed at the tungsten atom. The relative error is given w.r.t. the AE reference calculation.

Property	RC	AE	$\varepsilon_{\mathrm{rel}}/\%$
Kinetic energy$/E_{\mathrm{h}}$	21 117.681	21 120.805	1.48×10^{-2}
Mass velocity$/E_{\mathrm{h}}$	−87 494.319	−87 487.314	8.01×10^{-3}
$\langle -\frac{Z}{r} \rangle / E_{\mathrm{h}}$	−43 871.848	−43 890.212	4.18×10^{-2}
$\lvert\Psi(\mathrm{W})\rvert^2 / {a_0}^{-3}$	1 821 741.360	1 821 607.632	7.34×10^{-3}
$\lvert\Psi(\mathrm{Cl})\rvert^2 / {a_0}^{-3}$	2914.243	2914.757	1.76×10^{-2}
1-e$^-$-Darwin corr.$/E_{\mathrm{h}}$	11 280.916	11 280.089	7.33×10^{-3}
1-e$^-$-Darwin corr. (W)$/E_{\mathrm{h}}$	11 276.361	11 275.533	7.34×10^{-3}
1-e$^-$-Darwin corr. (Cl)$/E_{\mathrm{h}}$	4.144	4.145	1.76×10^{-2}
EDM$/e\,a_0$	497.336	497.311	4.93×10^{-3}

Table C.2.: Core and valence properties of NWHBrF. Expectation values are shown which were calculated from reconstructed and all-electron orbitals. The reconstruction was performed in subsequent steps: First, the orbitals were reconstructed at the bromine atom, and second at the tungsten atom. The relative error is given w.r.t. the AE reference calculation.

Property	RC: 1.Br, 2.W	AE	$\varepsilon_{\mathrm{rel}}/\%$
Kinetic energy$/E_{\mathrm{h}}$	23 387.207	23 394.309	3.04×10^{-2}
Mass velocity$/E_{\mathrm{h}}$	−87 867.123	−87 830.184	4.21×10^{-2}
$\langle -\frac{Z}{r} \rangle / E_{\mathrm{h}}$	−49 727.428	−49 751.066	4.75×10^{-2}
$\lvert\Psi(\mathrm{W})\rvert^2 / {a_0}^{-3}$	1 821 739.408	1 821 783.709	2.43×10^{-3}
$\lvert\Psi(\mathrm{Br})\rvert^2 / {a_0}^{-3}$	46 505.029	46 512.631	1.63×10^{-2}
1-e$^-$-Darwin corr.$/E_{\mathrm{h}}$	11 412.910	11 413.206	2.60×10^{-3}
1-e$^-$-Darwin corr. (W)$/E_{\mathrm{h}}$	11 276.349	11 276.623	2.43×10^{-3}
1-e$^-$-Darwin corr. (Br)$/E_{\mathrm{h}}$	136.150	136.173	1.63×10^{-2}
EDM$/e\,a_0$	534.301	534.301	4.99×10^{-5}

Table C.3.: Core and valence properties of NWHFI. Expectation values are shown which were calculated from reconstructed and all-electron orbitals. The reconstruction was performed in subsequent steps: First, the orbitals were reconstructed at the iodine atom, and second at the tungsten atom. The relative error is given w.r.t. the AE reference calculation.

Property	RC: 1.I, 2.W	AE	$\varepsilon_{\mathrm{rel}}/\%$
Kinetic energy$/E_{\mathrm{h}}$	28 615.514	28 621.133	1.96×10^{-2}
Mass velocity$/E_{\mathrm{h}}$	−92 971.585	−92 949.242	2.40×10^{-2}
$\langle -\frac{Z}{r} \rangle / E_{\mathrm{h}}$	−61 376.156	−61 397.828	3.53×10^{-2}
$\lvert\Psi(\mathrm{W})\rvert^2 / {a_0}^{-3}$	1 821 739.646	1 821 783.462	2.41×10^{-3}
$\lvert\Psi(\mathrm{I})\rvert^2 / {a_0}^{-3}$	280 475.890	280 488.268	4.41×10^{-3}
1-e$^-$-Darwin corr.$/E_{\mathrm{h}}$	12 520.193	12 520.519	2.60×10^{-3}
1-e$^-$-Darwin corr. (W)$/E_{\mathrm{h}}$	11 276.350	11 276.622	2.41×10^{-3}
1-e$^-$-Darwin corr. (I)$/E_{\mathrm{h}}$	1243.432	1243.487	4.41×10^{-3}
EDM$/e\,a_0$	552.773	552.887	2.07×10^{-2}

Appendix D.

Exemplary Implementation of the Reconstruction

```
(* Define primitive basis function χ̄ *)
normGauss[a_,x0_]  :=
     Evaluate[
         Assuming[
             Element[a, Reals] &&
              a > 0,
             Integrate[Exp[-a (x-x0)^2] Exp[-a (x-x0)^2], {x, -Infinity, Infinity}]
          ]
      ];
primitiveGaussian[a_,x0_] := Sqrt[1/normGauss[a,x0]] Exp[-a (x-x0)^2]
primitiveOverlap[b_,xb_,a_,xa_]:=Evaluate[
   Assuming[
      Element[a,Reals]&&
      Element[b,Reals]&&
      a > 0&&
      b > 0,
      Integrate[primitiveGaussian[b,xb]primitiveGaussian[a,xa],{x,-Infinity,
    Infinity}]
   ]
]
primitiveXYZOverlap[ b_,{xb_,yb_,zb_},a_,{xa_,ya_,za_} ]:=
   primitiveOverlap[b,xb,a,xa]*
   primitiveOverlap[b,yb,a,ya]*
   primitiveOverlap[b,zb,a,za]

(* Flatten basis given for easier matrix handling*)
flatBasis[ basis_ ] :=
   Module[
   {newbas},
      newbas = {};
      Do[
         newbf = {};
         Do[
```

```
            AppendTo[ newbf,
               { basis[[ atom, 1 ]],
                 basis[[ atom, 2, bf, primitive, 1]],
                 basis[[ atom, 2, bf, primitive, 2]]}
            ]
         ,{primitive, Length[ basis[[ atom, 2, bf ]] ] }
         ];
         AppendTo[newbas, newbf ];
      ,{atom, Length[ basis ] }
      ,{bf, Length[ basis[[ atom, 2 ]] ] }
      ];
      Return[ newbas ]
   ]

(* Normalize contraction coefficients *)
normalizeContraction[ basis_ ]:=
   Do[
      contrNorm =
         Sum[
            basis[[ bf, l, 3]] * basis[[ bf, k, 3]] *
            primitiveXYZOverlap[
               basis[[ bf, l, 2]],
               basis[[ bf, l, 1]],
               basis[[ bf, k, 2]],
               basis[[ bf, k, 1]]
            ]
            ,{l,Length[ basis[[ bf ]] ]}
            ,{k,Length[ basis[[ bf ]] ]}
         ];
      basis[[ bf, ;;, 3 ]] /= Sqrt[contrNorm];
      ,{bf, Length[ basis ]}
   ]
SetAttributes[ normalizeContraction, HoldFirst ];

(* Calculate overlap matrix S *)
overlapMatrix[ basis1_, basis2_ ] :=
   Table[
      Sum[
         basis1[[ bf1, l, 3]] * basis2[[ bf2, k, 3]] *
            primitiveXYZOverlap[
               basis1[[ bf1, l, 2]],
               basis1[[ bf1, l, 1]],
               basis2[[ bf2, k, 2]],
               basis2[[ bf2, k, 1]]
            ]
         ,{l,Length[ basis1[[ bf1 ]] ]}
         ,{k,Length[ basis2[[ bf2 ]] ]}
      ]
      ,{bf1, Length[ basis1 ]}
```

```
      ,{bf2, Length[ basis2 ]}
   ]

(* Define basis A *)
basisA = {
   { (* Li *)
      { 0,0,0 },
      {
         {
            { 0.63628975000     ,-0.099967228357407  },
            { 0.14786005000     , 0.39951282743192  },
            { 0.04808867800     , 0.70011546549963  }
         }
      }
   },
   { (* H *)
      {0,0,2.8848847899035},
      {
         {
            { 0.168856, 0.44463470222969},
            { 0.623913, 0.53532764149294},
            { 3.425250, 0.15432889664648}
         }
      }
   }
}

(* Define basis B *)
basisB = {
   { (* Li *)
      { 0,0,0 },
      {
         {
            {16.119575000, 0.15432897000},
            { 2.936200700, 0.53532814000},
            {0.7946504900, 0.44463454000}
         },
         {
            { 0.63628975000     ,-0.09996722900     },
            { 0.14786005000     , 0.39951283000     },
            { 0.04808867800     , 0.70011547000     }
         }
      }
   },
   { (* H *)
      {0,0,2.8848847899035},
      {
         {
            { 0.168856, 0.44463470222969},
```

```
            { 0.623913, 0.53532764149294},
            { 3.425250, 0.15432889664648}
         }
      }
   }
}


(* Calculate the overlap of two primitive basis functions *)
Print[primitiveOverlap[ a, xa, b, xb]]
Print["\n"]

(* Calculate the overlap matrix S_BB *)
Print["SBB"]
basisB = flatBasis[basisB]
normalizeContraction[ basisB ]
SBB = overlapMatrix[basisB,basisB]
Print[SetPrecision[MatrixForm[SBB],8]]
Print["\n"]

(* Define M and orthogonalize to get M_⊥ *)
Print[ "Non-orthogonal MMat"]
mMatNotOrth = Transpose[
      {
         { 0.99188411559721    , 0.031694515622807    , 0.000000},
         {-0.27911025978142    , 1.029918488798400    , 0.000000},
         { 0.00000000000000    , 0.000000000000000    , 1.000000}
      }
   ];
Print[ SetPrecision[MatrixForm[mMatNotOrth//Transpose],8] ]
Print["\n"]

Print[ "Orthogonalize last row of mMatNotOrth to get M_perp" ]
mMatNotOrth =  Transpose[
                  Orthogonalize[
                     Transpose[mMatNotOrth], #1.SBB.#2 &,
                     Method->"ModifiedGramSchmidt"
                  ]
               ]
newvec = mMatNotOrth[[;;,3]]

newvecNorm = Sum[
      newvec[[i]]*newvec[[j]]
      SBB[[i,j]]
   ,{i,1,3}
   ,{j,1,3}
   ]
newvec /= Sqrt[newvecNorm];

```

```
(* M_⊥ *)
mMat = Transpose[
      {
         Transpose[mMatNotOrth][[1]],
         Transpose[mMatNotOrth][[2]],
         newvec
      }
   ]

Print[ SetPrecision[MatrixForm[mMat//Transpose],8] ]
Print["\n"]

(* M_⊥core and M_⊥val *)
Print[ "Seperate mcore and mval" ]
mcore = Transpose[Transpose[mMat][[1]]]
mval  = Transpose[Transpose[mMat][[{2,3}]]]

(* Overlap matrix S_AB *)
Print[" Overlap between bases A and B "]
basisA = flatBasis[basisA]
SAB = overlapMatrix[basisA, basisB]

Print[SetPrecision[MatrixForm[SAB],8]]
Print["\n"]

(* Reconstruction *)
Print[ "Mapping the coefficients..." ]
ppCoeffs = {0.52489119392610,0.66131688845210}
rcCoeffs = ppCoeffs.SAB.mval.Transpose[mval]
Print[SetPrecision[rcCoeffs,8]]
Print["\n"]
Print[ "... and normalizing" ]
norm = rcCoeffs.SBB.rcCoeffs;
rcCoeffs /= Sqrt[norm]
Print[SetPrecision[rcCoeffs,8]]
Print["\n"]
```

Acronyms

1c one-component. 50, A-13

2c two-component. 2, 19, 62, 67, 73, 74

4c four-component. 2, 55

ACO atomic core orbital. 1, 10, 12, 16, 40, 41, 72

AE all-electron. VII, 2–7, 9, 12, 16–20, 25–31, 33, 35–42, 45, 46, 48–50, 53–57, 61–63, 67–74, A-20

AO atomic orbital. II, VII, 4, 9, 18, 20, 35–42, 45, 49, 50, 54, 61, 62, 74, A-27

AUO atomic unoccupied orbital. 10

AVO atomic valence orbital. 10, 19

BF basis function. 9, 41, 42, 44, 45

BOMD Born–Oppenheimer molecular dynamics. 99, 108, 109, 145, 146

CAS computer algebra system. 79

CC coupled cluster. 20

CEI Coulomb explosion imaging. II, 76, 77, 79, 81, 99, 109

CI configuration interaction. 4, 20

CIS configuration interaction singles. I, 26, 30, 33

CMD classical molecular dynamics. 79, 80, 108, 145, 146

COLTRIMS COLd Target Recoil Ion Momentum Spectroscopy. VII, 7, 76–80

CSF configuration state function. 9

DC Dirac–Coulomb. 2, 56

DCB Dirac–Coulomb–Breit. 2

DFT density functional theory. 17, 99, 108

DHF Dirac–Hartree–Fock. 1, 2, 19, A-14

DHF-C/B Dirac–Hartree–Fock–Coulomb–Breit. 50, A-14

DPT direct perturbation theory. 2

EC energy-consistent. 2, 3, 6, 7, 9, 10, 12, 18, 25, 26, 31, 33–36, 38, 50, 54, 72, 74
ECP effective core potential. 1, 10, 12, 25
EDM electric dipole moment. 48, 51, 53, 56, 57, 60, 61

FC frozen core. VII, 10–12, 16, 40
Full-CI full configuration interaction. 20

GHF generalized Hartree–Fock. 67
GPK generalized Phillips–Kleinman. 1, 16, 18
GS Gram–Schmidt. 45, 47

HF Hartree–Fock. 2, 14, 16–18, 20, 26–31, 33, 34, 36, 42, 44, 45, 48, 53, 62, 67, 71, 73, 143, A-38
HFS hyperfine structure. II, 54, 63
HOMO highest-occupied molecular orbital. 68

IOTC infinite order two component. 2

LCAO-MO linear combination of atomic orbitals to molecular orbitals. 11, 12
LHS left hand side. 32

MC multi-configuration. 20
MCDHF multi-configuration Dirac–Hartree–Fock. 19
MCHF multi-configuration Hartree–Fock. 19
MCP multi-channel plate. 76
MCSCF multi-configuration self-consistent field. 7, 9, 21, 23, 26, 33, 34, 38, 74, A-14
MD molecular dynamics. 145
MO molecular orbital. 4, 6, 9, 12, 20, 21, 28, 32, 39–41, 46–49, 61, 62, 72, A-27
MP model potential. 12
MR multi reference. 20
MV mass-velocity. 53
MVD mass-velocity Darwin. 6
MVO molecular valence orbital. 55

NOCR nonvariational one-center restoration procedure. 4

NR non-relativistic. 1–4, 19, 44, 50, 51, 54–56, 68, 73, A-13

OACO orthogonalized atomic core orbital. 46, 48

OAVO orthogonalized atomic valence orbital. 46, 47

PAO pseudo atomic orbital. 2, 4, 10, 45, 49, 53, 54

PES potential energy hypersurface. 108

PI2CO Photoion-Photoion-Coincidence. 108, 109

PI4CO fourfold photoion coincidence. 108, 109

PIXCO *X*-fold photoion coincidence. 109

PK Phillips–Kleinman. I, 1, 13, 16, 26, 33

PMO pseudo molecular orbital. 10, 25, 39–42, 44, 46, 47, 49, 50, 53

PO pseudo orbital. I, II, V, VII, 3–5, 7, 9, 10, 13, 14, 18–20, 25, 26, 28, 32–40, 49, 50, 54, 61, 72

PP pseudo potential. I, II, V, VII, 1–7, 9, 10, 12–14, 16–20, 25–44, 46, 47, 50, 53–57, 60, 63, 67, 71–74, 99, A-27

PT perturbation theory. 20

PVO pseudo valence orbital. 19

RC reconstructed. 68, 70, 71

RHF restricted Hartree–Fock. A-13

RHS right hand side. 61

RK Runge–Kutta. VII, 79, 80

ROHF restricted open-shell Hartree–Fock. A-13

SA state-averaged. 9, 21–23, 27

SC shape-consistent. 4, 6, 12, 18, 26, 36, 49, 72, 74

SCF self-consistent field. 12, 22, 26, 32, 54, 62, 65, 67, 73, 74, A-14

SD Slater determinant. 9, 21, 23

SR scalar-relativistic. 1, 6, 19, 20, 50, 54, 55, 62, 67, 71, 73, 74

SS state-specific. 21

TOF time-of-flight. 77, 78

UHF unrestricted Hartree–Fock. 19, 36, 67, A-13

w.r.t. with respect to. 9, 13, 16, 17, 19, 22, 23, 26, 35, 39, 42, 53–55, 63, 71–73, 75, 78, 98, 108, 109, A-25

X2C exact two component. 2

ZORA zeroth order regular approximation. 2, 3, 19, 50, 52, 62, 65, 67, 74, A-13

Symbols

l angular momentum. 17

Z atomic charge. 17, 35, 51, 58, 63–67, 144

ϕ atomic orbital. 9–16, 18, 27, 29, 30, 33–37, 39–42, 46

a_0 Bohr radius. 64

Φ configuration state function. 9, 15, 16, 23

ε Error. 53

ε_{rel} Relative error $\left|\frac{x-x_0}{x_0}\right|$. 48, 63, 68, 71, 72, A-30

α Fine structure constant. In atomic units $\alpha \approx \frac{1}{137}$. 13, 63, 64, 66

μ magnetic moment. 51, 68

m mass. 58

χ Gaussian type basis function. 9, 39–42, 44, 46, 56, 73

ψ molecular orbital. 4, 5, 9, 10, 21, 39–41, 44, 49, 64

$\bar{\chi}$ primitive Gaussian type basis function. 9, 44, A-31

ε orbital energy. 14, 42

r radius. 50, 51, 57–59, 64

R Relativistic scaling factor. VII, 63–65

S_{e} electron spin quantum number. 51

I spin quantum number. 51, 52, 68

Ψ Electronic state. 9, 18, 21–23, 26, 27, 29, 30, 52

Ψ Electronic wave function. 10, 11, 53