Development of local coupled cluster response methods for vertical Ionization Potentials

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der Fakultät Chemie und Pharmazie
der Universität Regensburg

vorgelegt von

Gero Wälz

aus Ingolstadt

im Jahr 2017

Promotionsgesuch eingereicht am:	02.02.2017
Tag des Kolloquiums:	07.04.2017
Diese Arbeit wurde angeleitet von:	Prof. Dr. Martin Schütz

Prüfungsausschuss	
Vorsitzender:	Prof. Dr. Arno Pfitzner
Erstgutachter:	Prof. Dr. Martin Schütz
Zweitgutachter:	PD Dr. Denis Usvyat
Drittprüfer:	Prof. Dr. Ferdinand Evers
Ersatzprüfer:	Prof. Dr. Alkwin Slenczka

Danksagung

Ein herzliches Dankeschön geht an alle, die mich in den letzten Jahren unterstützt haben und ohne die die Fertigstellung einer solchen Arbeit schwer vorstellbar ist. Im Besonderen sind das:

Prof. Dr. Martin Schütz, der mich fachlich unterstützt hat und mir immer wieder das Vertrauen geschenkt hat dieser Aufgabe gewachsen zu sein.

Dr. Denis Usvyat für seine ständige Bereitschaft jede Frage stets gut gelaunt zu beantworten.

Dr. Tatiana Korona, die mir den Start dieser Arbeit mit ihrem kanonischen IP-CC2 Programm sehr erleichtert hat.

Den ehemaligen Mitarbeitern des Arbeitskreises, *Dr. Daniel Kats* und *Dr. Keyarash Sadeghian*, die für Fragen weiterhin zur Verfügung standen.

Dr. Stefan Loibl, für die schöne Zeit im Büro und in der Freizeit.

Meinen Kollegen *Katrin Ledermüller, Oliver Masur, Martin Christlmaier* und *Alexander Krach* für ihre Bereitschaft aufeinander einzugehen und mit ihrem eigenen Wissen weiterzuhelfen.

Meinen Kollegen *Thomas Merz* und *Matthias Hinreiner* für die vielen gemeinsam verbrachten Mittagsstunden und Kaffeerunden, sowie die exzellente Versorgung mit Kuchen.

Meinem Freund und Kollegen *David David*, im Besonderen für die etlichen Glaubensgespräche.

Meinen Eltern *Marita* und *Michael* gebührt der Dank nicht nur für die letzten Jahre, sondern für ihre Liebe und Unterstützung mein Leben lang.

Die Ergebnisse dieser Arbeit sind größtenteils bereits veröffentlicht worden:

Chapter 2 und Section 3.2

G. Wälz, D. Usvyat, T. Korona and M. Schütz

"A hierarchy of local coupled cluster singles and doubles response methods for ionization potentials"

The Journal of Chemical Physics

144, 084117 (2016), doi: 10.1063/1.4942234

Contents

1. Introduction

The term *theoretical chemistry* sounds at the first glance contradictory. Commonly, when people think of chemistry, they tend to imagine something happening in a laboratory where some kind of synthesis is performed. And, it is true, chemistry today is still a lot of synthesizing. But chemistry contains a lot more than that. Chemistry is also examination of substances with a variety of methods, e.g. with different forms of spectroscopy. Or, coming back to a theoretical chemistry, simulation of chemical processes and prediction of properties. A simulation can, for example, reveal that a planned synthesis does not lead to the desired product. Or it can be used to verify experimental results. The combination of experimental work and theoretical prediction is a powerful tool in the hands of today's chemists.

However, and this is a crucial point, application of theoretical methods is often not a straightforward task due to the underlying approximations, which might or might not be adequate for the substance or reaction in question. As a result, a specific method might be very accurate in one case, but fails spectacularly in another. Therefore, one has to use theoretical methods with care and be aware of their advantages and flaws.

One of the most widely used methods is the density functional theory (DFT) [1, 2]. It can be considered a semi-empirical method and its power lies in a good balance between accuracy and low computational cost, allowing for short time calculations compared to most ab initio methods. Therefore, large systems can be handled quite easily, while the accuracy of the results remains decent. On the other hand, ab initio methods do not depend on an empirical parameterization and are generally more reliable. Thus, many groups try to develop ab initio techniques as well-controlled approximations to the exact

solution, such that calculations of large molecules can be completed within a reasonable amount of time.

The basic ab initio approach for the ground state is the Hartree-Fock (HF) method [3]. It is often used as a starting point for more accurate (so-called post-HF) methods. There are several classes of post-HF methods. The most common are the many-body perturbation theory MP2 [4], configuration interaction (CI) methods [3] like CIS, CISD, etc., and the coupled cluster (CC) methods. The *Full* CI can be considered a reference method as it provides the exact solution of the Schrödinger equation in case of the complete one-electron basis. Unfortunately this method implies determination of the amplitudes for all possible excited determinants, leading to factorial scaling of the computational cost with system size. Therefore, Full CI applications are limited to very small molecules.

1.1. Coupled cluster methods

In this section the focus is on the CC approach which will be essential throughout this work. The coupled cluster ansatz offers a very convenient factorization of the Full CI wavefunction, maintaining the size-extensivity of the solution regardless of the excitation truncation level. Below, a short overview covers the basic formalism of this method.

The CC wave function is given by

$$|\mathrm{CC}\rangle = \exp(\mathbf{T})|0\rangle. \tag{1.1}$$

It is build on a reference wave function $|0\rangle$, which is usually the Hartree-Fock determinant. The cluster operator is given by

$$\mathbf{T} = \sum_i \mathbf{T}_i = \sum_i \sum_{\mu_i} t_{\mu_i} \tau_{\mu_i}, \tag{1.2}$$

which sums all kinds of possible excitations: single ($i = 1$), double ($i = 2$), triple ($i = 3$) excitations and so on. Cluster operators for a certain excita-

tion ($\mathbf{T}_i$) are implicitly defined in Eq. (1.2) as a product of corresponding amplitudes t_{μ_i} and excitation operators τ_{μ_i}.

Plugging the CC wave function into the time-independent Schrödinger equation yields

$$\mathbf{H}\exp(\mathbf{T})|0\rangle = E_0\exp(\mathbf{T})|0\rangle. \tag{1.3}$$

By projecting Eq. (1.3) onto the reference one obtains the energy:

$$\langle 0|\mathbf{H}\exp(\mathbf{T})|0\rangle = E_0, \tag{1.4}$$

while a projection of Eq. (1.3) onto `bra` configuration state functions (CSFs)

$$\langle\mu_i| = \langle 0|\tau^\dagger_{\mu_i} \tag{1.5}$$

provides the *unlinked* amplitude equations,

$$\langle\mu_i|\mathbf{H}\exp(\mathbf{T})|0\rangle = 0. \tag{1.6}$$

The amplitudes can be obtained by solving equations (1.6). An alternative formulation of the amplitude equations is obtained by multiplying with $\exp(-\mathbf{T})$ from left before projection:

$$\langle\mu_i|\exp(-\mathbf{T})\mathbf{H}\exp(\mathbf{T})|0\rangle = 0. \tag{1.7}$$

These equations define the *linked* CC formalism, which allows for a useful decomposition into a commutator series via the Baker-Campbell-Hausdorff (BCH) expansion,

$$\exp(-\mathbf{T})\mathbf{H}\exp(\mathbf{T}) = \mathbf{H} + [\mathbf{H},\mathbf{T}] + \frac{1}{2!}[[\mathbf{H},\mathbf{T}],\mathbf{T}] + \ldots \tag{1.8}$$

As a result, disconnected terms do not occur in the amplitude equations. Furthermore, the series (Eq. (1.8)) terminates after the fourth power of $\mathbf{T}$ for any number of particles or truncation in $\mathbf{T}$.

However, the so-called similarity transformed Hamiltonian, Eq. (1.8), is

not Hermitian. Although hermiticity is a desirable property, especially in the context of excited state calculations, it is not necessarily required. Indeed, the eigenvalues of a Hermitian operator are real and its eigenvectors orthogonal to each other, thus delivering physically meaningful results. A non-hermitian approach can generally yield complex eigenvalues, but in practice this is rarely the case. It is possible to formulate Hermitian coupled cluster ansätze. To this end the factor $\exp(-\mathbf{T})$ is replaced or augmented by an exponential function containing deexcitation operators $\mathbf{T}^\dagger$. This however leads to Eq. (1.8) not terminating after the fourth power of $\mathbf{T}$. Still, through years Hermitian CC approaches drew certain interest [5, 6]. They have been recently employed in the context of linear response methods: TD-VCC [7, 8] and TD-UCC-H [8]. Note that any exponential function containing $\mathbf{T}^\dagger$ does not automatically lead to a Hermitian method.

The projected amplitude equations are further simplified by projection onto *contravariant* CSFs, instead of the *covariant* ones introduced in Eq. (1.5). Contravariant CSFs, marked with a tilde, fulfill the condition

$$\langle \tilde{\mu}_I | \mu_J \rangle = \delta_{IJ}. \tag{1.9}$$

I and J stand collectively for orbital indices. By projection onto the contravariant CSF the amount of individual terms, appearing in the working equations, is then reduced considerably.

If all excitation levels are included in the summation, Eq. (1.2), the CC model becomes equivalent to Full CI and, thus, delivers the exact solution within the given basis. However, computationally this is even more demanding than Full CI, making it prohibitively expensive even for relatively small molecules. Computationally less demanding CC methods can be formulated by restricting this summation to certain excitation levels. Furthermore, additional truncations are possible due to perturbation theory arguments for higher excitation levels. Doing so enables various methods with increasing computational cost and accuracy: from CCS (which is equivalent to HF for the ground state), to CC2 [9], CCSD [10], CC3 [11], CCSD(T) [12, 13], CCSDT [14, 15],

CCSDTQ and so forth. The excitation space of CCSD, for example, consists exclusively of singles and doubles ($i = 1, 2$). The coupled cluster models CC2 (approximation to CCSD), CC3 or CCSD(T) (approximations to CCSDT) involve additional reductions in the doubles and triples amplitude equations, respectively, by ignoring contributions that are of higher than a certain order within the Møller-Plesset partitioning, Eq. (2.18).

Besides its adaptivity to approximations the coupled cluster model has further advantages to offer. In summary, these advantages are described by the statement that coupled cluster is suitable for a *"theoretical model chemistry"*, which implies that it is possible to study *"a wide range of problems at a uniform level of approximation. [...] The effectiveness of any model may be evaluated by comparing some of its details with real chemistry in areas where experimental data are available"* as stated by Pople et al. [16, p. 1]. The postulated requirements for such a theoretical model chemistry were slightly adjusted and augmented by Bartlett [17]. One of these requirements is the size-extensivity, which is fulfilled by coupled cluster but for example not by truncated CI methods. Size-extensivity implies the proper scaling of a model with the size of the molecule. Some other criteria are the applicability to excited states and open shell systems, which is possible for CC methods, as well as efficiency and cost effectiveness.

A lot of effort was put into the development of (efficient) coupled cluster implementations during the last decades. The first implementations aimed towards the establishment of coupled cluster methods as an alternative to existing methods back then. Since CC methods proved to be an excellent methodological choice (so far CCSD(T) is considered the gold standard) the current focus targets the efficiency of established methods to expand their applicability. A lot of success was achieved with local CC2 (LCC2) allowing for efficient ground and excited state calculations. For the latter, the current implementations make it possible to compute excitation energies [18, 19], orbital unrelaxed and relaxed first-order properties [20–23] and analytic gradients with respect to nuclear displacements [24].

The work presented here contributes one aspect to the ongoing effort of

efficient implementations. It allows calculation of ionization potentials for extended molecules (around 100 atoms and more) within a reasonable amount of time (see Tab. 2.2). By considering energy differences between ionization potentials it is also possible to obtain excitation energies of open-shell molecules. This topic is addressed in Sec. 3.2.

1.2. Approximations

A general problem of coupled cluster models are their high computational cost. Their scaling with respect to the molecular size $\mathcal{N}$ ranges from $\mathcal{N}^4$ (CCS), $\mathcal{N}^5$ (CC2), $\mathcal{N}^6$ (CCSD), $\mathcal{N}^7$ (CC3) to $\mathcal{N}^8$ (CCSDT). To achieve efficient methods for moderate and large molecules, approximations are introduced. One of the basic principles for approximations is that the error should be smaller than the error of the approximated method itself. The approximations used in this work are presented in this section.

1.2.1. Local ansatz

One of the key concepts behind the local CC methods for extended molecular systems developed by Schütz and co-workers is the local ansatz proposed by Pulay [25]. It utilizes the short range nature of electron correlation which can be accessed after transforming the canonical, delocalized orbitals into local ones. This is done separately for the occupied and virtual space. The former is localized according to a certain criterion, e.g. proposed by Pipek and Mezey [26] or Boys [27]. Most commonly used is the criterion of Pipek and Mezey, which minimizes the number of atoms the occupied orbitals are located on. Its computational complexity is as economical as Boys' method, but, in contrast to Boys, it is able to separate σ from π orbitals in planar molecules. According to the chosen criterion the unitary transformation matrix $\mathbf{W}$ is specified, which transforms the canonical occupied orbitals $\Phi_{\tilde{i}}^{\mathrm{CAN}}$

into localized molecular orbitals (LMOs) Φ_i^{LMO}:

$$|\phi_i^{\mathrm{LMO}}\rangle = |\phi_{\bar{i}}^{\mathrm{CAN}}\rangle W_{\bar{i}i}. \tag{1.10}$$

Here and in the following indices $i, j, k, \ldots$ denote occupied, $a, b, c, \ldots$ virtual, and $p, q, r, \ldots$ arbitrary local orbitals. Indices for canonical orbitals are decorated with a bar on top of the characters.

The virtual space, on the other hand, is spanned by projected atomic orbitals (PAOs). PAOs are obtained from projection of atomic orbitals (AOs) onto the virtual space. They are mutually nonorthogonal but orthogonal to the occupied LMOs.

The local orbitals are then used to reduce the number of allowed electron excitations. Two local approximations are usually utilized, one restricting the occupied the other the virtual space. Restrictions on the virtual space are introduced by creating smaller subspaces, so-called *domains*, which are build for each LMO (orbital domains), as well as pairs (pair domains) or triples (triple domains) of LMOs separately. An orbital domain $[i]$ contains only PAOs in spatial vicinity of LMO i. They are build, e.g., according to the procedure of Boughton and Pulay [28]. In case of double excitations from two LMOs i and j the corresponding pair domain $[ij]$ is obtained from the union of orbital domains $[i]$ and $[j]$.

Furthermore, restricted pair and triples lists are used to reduce the occupied space, allowing only certain LMO combinations and discarding the others. Such lists can be created, e.g., based on a distance criterion. Depending on the distance between two LMOs of a pair, pairs can be classified into strong, close, weak, distant and very distant pairs. This classification can then be used to treat pairs with different levels of theory, i.e., strong pairs with the highest level of theory and the other pair classes with progressively less accurate methods. The LCC2 implementation in `MOLPRO` [29] distinguishes only between strong and weak pairs, the latter being treated at MP2 level. Asymptotically, the pair- and domain-restrictions lower the overall scaling of a method up to linear, which makes the local approach particularly effective for large molecules.

However, distance criteria are only sufficient for the ground state. Creating suitable virtual orbital subspaces and pair lists for excited states is a more delicate task due to effects like charge-transfer, excitation to Rydberg states or large π systems. These examples share a non-local behaviour where pure distance criteria will fail. An alternative to the distance based criterion is presented in section 2.4.4.

1.2.2. Density fitting

Another important approximation is the density fitting (DF) approximation. It focuses on the two electron integrals

$$\begin{aligned}(pq|rs) &= \int \mathrm{d}\mathbf{r}_1 \int \mathrm{d}\mathbf{r}_2 \phi_p(\mathbf{r}_1)\phi_q(\mathbf{r}_1) r_{12}^{-1} \phi_r(\mathbf{r}_2)\phi_s(\mathbf{r}_2) \\ &= \int \mathrm{d}\mathbf{r}_1 \int \mathrm{d}\mathbf{r}_2 \rho_{pq}(\mathbf{r}_1) r_{12}^{-1} \rho_{rs}(\mathbf{r}_2), \end{aligned} \tag{1.11}$$

expressed in terms of one-particle densities ρ of orbital products in the second equality. These densities are substituted by approximated densities $\tilde{\rho}$ expanded in a fitting basis $\{\Xi_P(\mathbf{r})\}$:

$$\rho_{pq}(\mathbf{r}) \approx \tilde{\rho}_{pq}(\mathbf{r}) = \sum_P c_P^{pq} \Xi_P(\mathbf{r}). \tag{1.12}$$

Capital letters P, Q denote the fitting functions in the following. Following the approach of Dunlap [30], minimization of the fitting error with the Coulomb metric $J_{PQ} = (P|Q)$ leads to the linear equation system:

$$\sum_Q J_{PQ} c_Q^{pq} = (P|pq). \tag{1.13}$$

with the fitting coefficients c_P^{pq} as a solution. Finally, the four-indexed integrals can be expressed in terms of two three-indexed quantities:

$$(pq|rs) \approx \sum_P (pq|P) c_P^{rs}, \tag{1.14}$$

According to Eq. (1.13) the fitting coefficients are

$$c_P^{pq} = \sum_Q (J^{-1})_{PQ}(Q|pq). \tag{1.15}$$

In contrast to the local approximation, density fitting does not reduce the scaling of methods like DF-MP2 or DF-CC2, but the multiplicative prefactor is substantially lowered. On a sidenote, a step further in the direction of factorization of two-electron integrals is the tensor hypercontraction illustrated by Hohenstein et al. [31, 32], who decompose the four-indexed integrals into two-indexed rather than three-indexed quantities. In contrast to density fitting, such a technique lowers the nominal scaling of CC2 from $\mathcal{O}(\mathcal{N}^5)$ to $\mathcal{O}(\mathcal{N}^4)$.

1.2.3. Laplace transformation

In case of quantum chemical schemes involving orbital energy denominators, the Laplace transformation (LT) is another handy tool. It allows for a convenient partitioning of the eigenvalue problem. The doubles-doubles block of the Jacobi matrix in canonical CC2 is diagonal and the partitioned eigenvalue problem then reads,

$$A^{\text{eff}}_{\mu_1\nu_1}(\omega_{\bar{m}})R^{\bar{m}}_{\nu_1} = A_{\mu_1\nu_1}R^{\bar{m}}_{\nu_1} + A_{\mu_1\xi_2}\frac{A_{\xi_2\nu_1}R^{\bar{m}}_{\nu_1}}{\omega_{\bar{m}} - \Delta\epsilon_{\xi_2}}, \tag{1.16}$$

forming an effective singles eigenvalue problem. In local CC2, however, this is not possible, since the Fock matrix is not diagonal, which requires either solving the doubles equations or processing doubles and singles vector together in the Davidson procedure. Introducing the Laplace transformation though

$$\frac{1}{x} = \int_0^\infty \exp(-xt)dt \tag{1.17}$$

makes it possible to avoid these complications. Applying Eq. (1.17) to Eq. (1.16) and replacing the integration by a numerical quadrature yields

$$A^{\text{eff}}_{\mu_1\nu_1}(\omega_{\bar{m}})R^{\bar{m}}_{\nu_1} = A_{\mu_1\nu_1}R^{\bar{m}}_{\nu_1} - A_{\mu_1\xi_2}\int_0^\infty dt\ e^{-\Delta\epsilon_{\xi_2}t}e^{\omega_{\bar{m}}t}A_{\xi_2\nu_1}R^{\bar{m}}_{\nu_1}$$
$$\approx A_{\mu_1\nu_1}R^{\bar{m}}_{\nu_1} - A_{\mu_1\xi_2}\sum_{q=1}^{n_q} w_q e^{-\Delta\epsilon_{\xi_2}t_q}e^{\omega_{\bar{m}}t_q}A_{\xi_2\nu_1}R^{\bar{m}}_{\nu_1}. \qquad (1.18)$$

This quadrature allows one to convert the doubles residual from the local basis to canonical, multiply it with Laplace exponential factors, and convert the result back to the local basis as one transformation. This effectively reduces the Davidson space to be spanned by singles vectors only. For LT-DF-LCC2 [19] it was found that three integration points n_q, with corresponding quadrature points t_q and weights w_q, are sufficient. A Laplace transformed CC2 method for ionization potentials was also developed in this work, however, it was found that CC2 is not as accurate for ionization potentials as it is for excitation energies. Further development implied adding higher-order terms to the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ block of the Jacobian, which consequently made the Laplace transformation in this context hardly useful. Therefore, the Laplace transformation is not further outlined in this thesis.

1.3. Diagrammatic techniques

A simple way to develop working equations is the usage of diagrams. Diagrams are image representations of matrix elements, which provide a link to programmable expressions. The task of finding all necessary expressions is then replaced by finding all unique diagrams. Two example diagrams are shown in Fig. (1.1). The first diagram is connected to the bare excitation operator of the `bra` side drawn as a small dashed horizontal line. Diagrams closed like this refer to matrix elements obtained by projections on covariant CSFs. Open diagrams like the second one, on the other hand, refer to projections on

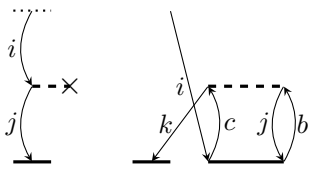

Figure 1.1.: Exemplary diagrams

contravariant CSFs.

Bold dashed horizontal lines, or interaction lines, represent fragments of the Hamilton operator like the Fock operator **F** (first diagram) or the fluctuation potential **W** (second diagram). They have as many vertices as electrons they act on. Since the Fock operator is an one-electron operator the second vertex is crossed out. Non-dashed horizontal lines, representing the cluster operators **T**, have as many vertices as electrons they are acting on, as well. Vertices usually have an incoming and outgoing vertical line, which reflect the action of the operator on an electron. However, operators that remove (add) one electron from (to) a molecular system, as it is the case for cluster operators of ionized (electron attached) states, have one excess incoming (outgoing) line.

Vertical lines are orbital substitutions relative to the Fermi vacuum and labeled with orbital indices. They can point either downwards representing occupied orbitals (hole lines) or upwards for virtual orbitals (particle lines). If vertical lines are connecting two vertices they are called *internal*. *External* lines have at least one open end. Exception from this rule are lines connected to the bare excitation operator: these lines are also considered external. With the following rules, diagrams can be translated into mathematical expressions:

- Each closed loop containing exclusively internal lines contributes a factor of two.
- The sign is given by $(-1)^{h+l}$ in which h is the number of hole lines and l the number of (closed plus non-closed) loops.
- Summation over internal lines.
- Horizontal lines contribute either an integral (bold dashed lines) or an

amplitude (non-dashed lines). The two-electron integrals take the form $(out_1\ in_1|out_2\ in_2)$, *out* labeled according to outgoing and *in* according to incoming vertical lines of a certain vertex 1 or 2. The Fock matrix is obtained similarly: $\langle out|F|in\rangle$. Labels of amplitudes are arranged by virtue of occupied (*occ*) and virtual (*virt*) indices rather than outgoing and incoming lines: $t^{occ_1\ occ_2\ \ldots}_{virt_1 virt_2 \ldots}$. The bare excitation operators contribute nothing.

- Vertical lines between bare excitation operators and interaction lines are dressed via similarity transformation with $\mathbf{T}_1$ amplitudes.
- In local basis every particle line not connected to an integral gives rise to the PAO overlap matrix $\mathbf{S}_{\mathrm{PAO}}$.

For example, evaluation of the two diagrams in Fig. (1.1) with the given rules yields

$$-\sum_j \hat{f}_{ji} u^j \qquad \text{and} \qquad -2\sum_{bcjk} (kc|jb) t^{ij}_{cb} u^k. \tag{1.19}$$

Note that the left expression results from a diagram with a covariant CSF, the right one from a diagram with a contravariant CSF.

2. Hierarchy of local coupled cluster ionization potentials

The content of this chapter has been published already in Ref. [33]. Authors involved in this work were Dr. Denis Usvyat, Dr. Tatiana Korona and my supervisor Prof. Dr. Martin Schütz. This chapter is taken entirely from the above mentioned publication. Some minor changes and additional notes have been added.

2.1. Introduction

In section 1.1 the coupled cluster theory was briefly presented for treating molecules in their ground state. This chapter covers the formalism for excited states, in particular for ionized states. One very versatile approach is response theory [34]. From response functions frequency-dependent molecular properties are obtained, from which expressions for the determination of excitation energies, transition strengths, polarizability, hyperpolarizabilities, etc. can be derived. This theory is used in the following to develop a formalism for ionized states and their ionization energies.

Another approach to excited (including ionized and electron attached) states is the equation-of-motion (EOM) CC framework. Methods for the calculation of ionization potentials (IPs) and properties of ionized states have been presented before in the EOM-CC framework [35–45], yet to the author's knowledge so far only for non-local canonical or natural orbitals, and without a DF based factorization of the Jacobian transforms. Note that ionization potentials can

also be simulated with existing programs for excitation energies by adding a very diffuse orbital to the space of virtual orbitals [39]. However, calculations employing this technique do not benefit from the lower scaling of a separate IP program.

2.2. Response theory for ionized states

As the time-dependent perturbation a formal non-physical operator is introduced, which destroys and creates a particle,

$$\mathbf{V}(t) = \sum_{k=-n}^{n} \exp(-i\omega_k t)\mathbf{V}(\omega_k), \tag{2.1}$$

$$\mathbf{V}(\omega_k) = \sum_{Y} \epsilon_Y(\omega_k)\mathbf{Y}, \quad \mathbf{Y} = \sum_{p} Y_p(a_{p\beta} + a_{p\beta}^{\dagger}), \tag{2.2}$$

with the perturbation strengths $\epsilon_Y(\omega_k)$ and the annihilation ($a_{p\beta}$) and creation ($a_{p\beta}^{\dagger}$) operators in second quantization. By treating ionization and electron attachment processes together, $\mathbf{Y}$ remains Hermitian. Together with the symmetry properties $\omega_{-k} = -\omega_k$, and $\epsilon_Y^*(\omega_k) = \epsilon_Y(\omega_{-k})$ the time-dependent perturbation $\mathbf{V}(t)$ is Hermitian. Since $\mathbf{V}(t)$ is unphysical the "integrals" Y_p of operator $\mathbf{Y}$ are set to one, for simplicity. With operator $\mathbf{X} \equiv \mathbf{Y}$, the resulting

exact linear response function for such a perturbation can then be written as

$$\begin{aligned}
\langle\langle X;Y\rangle\rangle_\omega &= \sum_{\bar{I}} \left(\frac{\langle 0|X|\bar{I}\rangle\langle\bar{I}|Y|0\rangle}{\omega - (E_{\bar{I}} - E_0)} - \frac{\langle 0|Y|\bar{I}\rangle\langle\bar{I}|X|0\rangle}{\omega + (E_{\bar{I}} - E_0)} \right) \\
&\quad + \sum_{\bar{A}} \left(\frac{\langle 0|X|\bar{A}\rangle\langle\bar{A}|Y|0\rangle}{\omega - (E_{\bar{A}} - E_0)} - \frac{\langle 0|Y|\bar{A}\rangle\langle\bar{A}|X|0\rangle}{\omega + (E_{\bar{A}} - E_0)} \right) \\
&= \sum_{pq} \left\{ \sum_{\bar{I}} \left(\frac{\langle 0|a^\dagger_{p\beta}|\bar{I}\rangle\langle\bar{I}|a_{q\beta}|0\rangle}{\omega - \omega_{\bar{I}}} - \frac{\langle 0|a^\dagger_{p\beta}|\bar{I}\rangle\langle\bar{I}|a_{q\beta}|0\rangle}{\omega + \omega_{\bar{I}}} \right) \right. \\
&\quad \left. + \sum_{\bar{A}} \left(\frac{\langle 0|a_{p\beta}|\bar{A}\rangle\langle\bar{A}|a^\dagger_{q\beta}|0\rangle}{\omega - \omega_{\bar{A}}} - \frac{\langle 0|a_{p\beta}|\bar{A}\rangle\langle\bar{A}|a^\dagger_{q\beta}|0\rangle}{\omega + \omega_{\bar{A}}} \right) \right\} \\
&= \sum_{ij} \sum_{\bar{I}} \left(\frac{\langle 0|a^\dagger_{i\beta}|\bar{I}\rangle\langle\bar{I}|a_{j\beta}|0\rangle}{\omega - \omega_{\bar{I}}} - \frac{\langle 0|a^\dagger_{i\beta}|\bar{I}\rangle\langle\bar{I}|a_{j\beta}|0\rangle}{\omega + \omega_{\bar{I}}} \right) \\
&\quad + \sum_{ab} \sum_{\bar{A}} \left(\frac{\langle 0|a_{a\beta}|\bar{A}\rangle\langle\bar{A}|a^\dagger_{b\beta}|0\rangle}{\omega - \omega_{\bar{A}}} - \frac{\langle 0|a_{a\beta}|\bar{A}\rangle\langle\bar{A}|a^\dagger_{b\beta}|0\rangle}{\omega + \omega_{\bar{A}}} \right), \qquad (2.3)
\end{aligned}$$

where $|\bar{I}\rangle$ are the ionized eigenstates living in the Fock subspace $F(M, N-1)$ with related energies $E_{\bar{I}}$, and $|\bar{A}\rangle$ the electron attached states with related energies $E_{\bar{A}}$ living in $F(M, N+1)$. M and N denote the number of spin orbitals and the number of electrons of the system, respectively. The linear response function $\langle\langle X;Y\rangle\rangle_\omega$ has poles for ionization energies $\omega_{\bar{I}}$ and electron affinities $\omega_{\bar{A}}$. In this chapter only the first part of $\langle\langle X;Y\rangle\rangle_\omega$, containing the poles for $\omega_{\bar{I}}$, is discussed. It contains only state functions for ground and ionized states. Ionized states are generated from ground state functions by an operator with one excess annihilator (cf. Eq. (2.8)). The time-dependent CC wavefunction after isolation of the phase can be written as

$$|\widetilde{\mathrm{CC}}\rangle = \exp(\mathbf{T}^{(0)} + \mathbf{T}^{(1)}(t) + \dots)|0\rangle, \qquad (2.4)$$

with

$$\begin{aligned}\mathbf{T}^{(0)} &= \mathbf{T}_1^{(0)} + \mathbf{T}_2^{(0)} + \ldots, \\ \mathbf{T}^{(1)}(t) &= \mathbf{T}_{\frac{1}{2}}^{(1)}(t) + \mathbf{T}_{\frac{3}{2}}^{(1)}(t) + \ldots,\end{aligned} \tag{2.5}$$

and

$$\begin{aligned}\mathbf{T}_1^{(0)} &= t_{\mu_1}^{(0)} \tau_{\mu_1} = t_a^i \tau_i^a, \\ \mathbf{T}_2^{(0)} &= t_{\mu_2}^{(0)} \tau_{\mu_2} = \frac{1}{2} t_{ab}^{ij} \tau_{ij}^{ab}, \\ \mathbf{T}_{\frac{1}{2}}^{(1)}(t) &= t_{\mu_{\frac{1}{2}}}^{(1)}(t) \tau_{\mu_{\frac{1}{2}}} = t^i(t) \tau_i, \\ \mathbf{T}_{\frac{3}{2}}^{(1)}(t) &= t_{\mu_{\frac{3}{2}}}^{(1)}(t) \tau_{\mu_{\frac{3}{2}}} = t_a^{ij}(t) \tau_{ij}^a.\end{aligned} \tag{2.6}$$

Note that Einstein convention is used above and in the following, i.e., repeated indices are implicitly summed up. Summations are written explicitly only if it is helpful for clarity.

In Eqs. (2.5) and (2.6) the particle rank m of related operators is introduced, i.e., the number of elementary operators of an operator string divided by two, as subscript indices in the individual $\mathbf{T}_m$ operators. For CC2 and CCSD the particle rank is truncated at $m = 2$ in the cluster operator. Furthermore, by virtue of the $2n + 1$ rule it is sufficient to consider amplitudes up to first order with respect to $\mathbf{V}(t)$. Orders with respect to the perturbation in time are given by superscripted numbers in parenthesis. Note that zeroth-order amplitudes with half-integer particle rank, as well as first order amplitudes with integer particle rank, are all zero.

The operators τ_m in Eq. (2.6) are all spin-adapted:

$$\tau_i^a = a_{a\alpha}^\dagger a_{i\alpha} + a_{a\beta}^\dagger a_{i\beta}, \qquad \tau_{ij}^{ab} = \tau_i^a \tau_j^b, \tag{2.7}$$

$$\tau_i = a_{i\beta}, \qquad \tau_{ij}^a = \tau_i^a \tau_j, \qquad \tau_{ijk}^{ab} = \tau_{ij}^{ab} \tau_k. \tag{2.8}$$

The operators in Eq. (2.7) with integer particle rank are spin- and particle-conserving or singlet-coupled excitation operators, generating a singlet state

when being applied to the closed shell reference determinant $|0\rangle$. On the other hand, the operators in Eq. (2.8) with half-integer rank produce a doublet state with $S = M_S = \frac{1}{2}$ and reduce the number of particles by one. As a sidenote, the LMO pair list for zeroth-order amplitudes t_{ab}^{ij} is triangular, while it is not for the first-order amplitudes t_a^{ij}.

The contravariant `bra` functions forming a biorthonormal set with the `ket` functions take the form

$$\begin{aligned}
\langle\tilde{\mu}_1| &= \langle\tilde{\Phi}_i^a| = \frac{1}{2}\langle 0|(\tau_i^a)^\dagger, \\
\langle\tilde{\mu}_2| &= \langle\tilde{\Phi}_{ij}^{ab}| = \frac{1}{6}\langle 0|(2\tau_{ij}^{ab} + \tau_{ji}^{ab})^\dagger, \\
\langle\tilde{\mu}_{\frac{1}{2}}| &= \langle\tilde{\Phi}_i| = \langle 0|(\tau_i)^\dagger, \\
\langle\tilde{\mu}_{\frac{3}{2}}| &= \langle\tilde{\Phi}_{ij}^a| = \frac{1}{3}\langle 0|(2\tau_{ij}^a + \tau_{ji}^a)^\dagger.
\end{aligned} \tag{2.9}$$

With these information the time-averaged second-order quasienergy Lagrangian can be set up:

$$\begin{aligned}
\{^{2n+1}L^{(2)}(t)\}_T = \sum_{k=-n}^{n} &\left[\left\langle 0 \middle| \left[\mathbf{V}(-\omega_k), \mathbf{T}_{\frac{1}{2}}^{(1)}(\omega_k)\right] \middle| \mathrm{CC}\right\rangle\right. \\
&+ \lambda_{\mu_m}^{(0)} \left\langle \tilde{\mu}'_m \middle| \left[\mathbf{V}(-\omega_k), \mathbf{T}_{\frac{1}{2}}^{(1)}(\omega_k) + \mathbf{T}_{\frac{3}{2}}^{(1)}(\omega_k)\right] \middle| \mathrm{CC}\right\rangle \\
&- \lambda_{\mu_m}^{(1)}(-\omega_k)\omega_k t_{\nu_l}^{(1)}(\omega_k) \left\langle \tilde{\mu}_m \middle| \tau_{\nu_l} \right| \mathrm{CC}\rangle \\
&\left. + \lambda_{\mu_m}^{(1)}(-\omega_k) \left\langle \tilde{\mu}'_m \middle| \mathbf{V}(\omega_k) + \left[\mathbf{H}^{(0)}, \mathbf{T}_{\frac{1}{2}}^{(1)}(\omega_k) + \mathbf{T}_{\frac{3}{2}}^{(1)}(\omega_k)\right] \middle| \mathrm{CC}\right\rangle\right]
\end{aligned} \tag{2.10}$$

where $|\mathrm{CC}\rangle = \exp(\mathbf{T}^{(0)})|0\rangle$ is the unperturbed CC wavefunction, $\langle\tilde{\mu}'_m| = \langle\tilde{\mu}_m|\exp(-\mathbf{T}^{(0)})$,

$$\begin{aligned}
\mathbf{T}_m^{(1)}(\omega_k) &= t_{\mu_m}^{(1)}(\omega_k)\tau_{\mu_m}, \quad \text{with} \\
t_{\mu_m}^{(1)}(t) &= \sum_{k=-n}^{n} t_{\mu_m}^{(1)}(\omega_k)\exp(-i\omega_k t),
\end{aligned} \tag{2.11}$$

$\lambda_{\mu_m}^{(1)}(\omega_k)$ is defined analogously, and $\mathbf{H}^{(0)}$ is the unperturbed Hamiltonian. In

Eq. (2.10) the particle-rank index m runs over $m = 1, 2$ for zeroth-order multipliers $\lambda^{(0)}_{\mu_m}$, and over $m = \frac{1}{2}, \frac{3}{2}$ for first-order multipliers $\lambda^{(1)}_{\mu_m}(\omega_k)$ and amplitudes $t^{(1)}_{\mu_m}(\omega_k)$. The Lagrangian contains no products of $\mathbf{T}^{(1)}_m(\omega_k)$ cluster operators, since related diagrams cannot be closed, neither by operators with integer particle rank nor in combination with one operator with half-integer particle rank. This implies that the second derivative of $\{^{2n+1}L^{(2)}(t)\}_T$ with respect to the first-order amplitudes is zero, which is a substantial simplification relative to CC response theory for electronically excited states. The linear response function with poles for the ionization potentials is obtained as the second derivative of $\{^{2n+1}L^{(2)}(t)\}_T$ with respect to the perturbation strengths ϵ:

$$\langle\langle X;Y\rangle\rangle'_\omega = \frac{d^2\{^{2n+1}L^{(2)}(t)\}_T}{d\epsilon_X(-\omega)d\epsilon_Y(\omega)} = \eta^X t^Y_{(\omega)} + \eta^Y t^X_{(-\omega)}, \tag{2.12}$$

with

$$\begin{aligned}\eta^Y_{\mu_l} &= \frac{\partial^2\{^{2n+1}L^{(2)}(t)\}_T}{\partial\epsilon_Y(-\omega)\partial t^{(1)}_{\mu_l}(\omega)} \\ &= \left\langle 0 \left| \left[\mathbf{Y}, \tau_{\mu_{\frac{1}{2}}}\right] \right| 0 \right\rangle \delta_{l\frac{1}{2}} + \lambda^{(0)}_{\mu_m} \left\langle \tilde{\mu}'_m \left| \left[\mathbf{Y}, \tau_{\frac{3}{2}}\right] \right| 0 \right\rangle \delta_{l\frac{3}{2}},\end{aligned} \tag{2.13}$$

and m being integer and l half-integer. Note that in the first two summands of Eq. (2.10) only the reference $|0\rangle$ of the `ket` $|\mathrm{CC}\rangle$ remains since $\mathbf{Y}$ cannot connect multiple `kets`.

$t^Y_{(\omega)}$ is obtained from the stationary conditions

$$(\mathbf{A} - \omega\mathbf{M})t^Y_{(\omega)} + \xi^Y = 0, \tag{2.14}$$

with the CC Jacobian $\mathbf{A}$,

$$\begin{aligned}A_{\mu_m\nu_l} - \omega M_{\mu_m\nu_l} &= \frac{\partial^2\{^{2n+1}L^{(2)}(t)\}_T}{\partial\lambda^{(1)}_{\mu_m}(-\omega)\partial t^{(1)}_{\mu_l}(\omega)}, \\ A_{\mu_m\nu_l} &= \left\langle \tilde{\mu}'_m \left| \left[\mathbf{H}^{(0)}, \tau_{\nu_l}\right] \right| \mathrm{CC} \right\rangle,\end{aligned} \tag{2.15}$$

metric

$$M_{\mu_m \nu_l} = \langle \tilde{\mu}_m | \tau_{\nu_l} | 0 \rangle, \tag{2.16}$$

and

$$\xi^Y_{\mu_m} = \frac{\{^{2n+1}L^{(2)}(t)\}_T}{\partial \lambda^{(1)}_{\mu_m}(-\omega) \partial \epsilon_Y(\omega)} = \langle \tilde{\mu}'_m \, | \mathbf{Y} | \, \mathrm{CC} \rangle = \langle \tilde{\mu}'_m \, | \mathbf{Y} | \, 0 \rangle \, , \tag{2.17}$$

with m and l both half-integer indices. From Eq. (2.17) it is clear that $\xi^Y_{\mu_m}$ is non-zero for $m = \frac{1}{2}$. Therefore, $t^Y_{(\omega)}$ has poles for the singular matrix $\mathbf{A} - \omega \mathbf{M}$ (cf. Eq.(2.14)), which, in turn, leads to poles for these ω in the linear response function $\langle\langle X; Y \rangle\rangle_\omega$ according to Eq. (2.12). The eigenvalues of the CC Jacobian $\mathbf{A}$ hence correspond to the ionization potentials of the molecular system as described by the CC model.

2.3. Approximate coupled cluster model CC2

As outlined earlier, CCSD scales as $\mathcal{N}^6$. The motivation for CC2 is to obtain a cheaper coupled cluster method which still contains correlation energy unlike CCS. Christiansen et al. [9] presented this method 1995. CC2 is designed such that the ground state energy is correct through second order, in contrast to CCSD which is correct through third order. Therefore, the total energy is of MP2 quality, however, CC2 allows for calculating excitation energies and transition moments. The CC2 model is based on the Møller-Plesset partitioning of the Hamiltonian,

$$\mathbf{H}^{(0)} = \mathbf{F}^{[0](0)} + \mathbf{W}^{[1](0)}, \tag{2.18}$$

with $\mathbf{F}^{[0](0)}$ representing the Fock operator, and $\mathbf{W}^{[1](0)}$ the fluctuation potential. The order with respect to the fluctuation potential, and with respect to $\mathbf{V}(t)$ is indicated by superscript numbers in brackets, and parenthesis, respectively. In the following these superscripts are dropped for the Hamiltonian and its fragments. Lower scaling and an energy correct through second order is achieved by approximation of the CCSD amplitude equations: the singles amplitude equations remain unchanged and the doubles amplitude equations are approximated to be correct through first order with respect to the fluctuation

potential,

$$0 = \left\langle \tilde{\mu}_1 \left| \hat{\mathbf{H}} + \left[\hat{\mathbf{H}}, \mathbf{T}_2^{(0)} \right] \right| 0 \right\rangle, \tag{2.19}$$

$$0 = \left\langle \tilde{\mu}_2 \left| \hat{\mathbf{H}} + \left[\mathbf{F}, \mathbf{T}_2^{(0)} \right] \right| 0 \right\rangle, \tag{2.20}$$

with the similarity-transformed Hamiltonian

$$\hat{\mathbf{H}} = \exp(-\mathbf{T}_1^{(0)}) \mathbf{H} \exp(\mathbf{T}_1^{(0)}). \tag{2.21}$$

The singles amplitudes have a special role: although singles amplitudes appear at second order with respect to the fluctuation potential for the first time, this is due to the fact that Hartree-Fock is used as a reference [11]. Otherwise, singles already appear at zeroth order. Since singles are important to approximate orbital relaxation effects they are assigned to be zeroth-order in $\mathbf{W}$. The similarity transformation of the Hamiltonian is a convenient way to keep the singles amplitudes in the equations. This transformation leads to dressed integrals,

$$\hat{(pq|rs)} = (\mu\nu|\rho\sigma) \lambda^p_{\mu p} \lambda^h_{\nu q} \lambda^p_{\rho r} \lambda^h_{\sigma s}, \tag{2.22}$$

decorated with a hat. The coefficient matrices λ^p and λ^h transform from AO basis (indexed by Greek letters $\mu, \nu, \dots$) to MO basis. They are calculated from the LMO and PAO coefficient matrices $\mathbf{L}$ and $\mathbf{P}$, respectively, the PAO overlap matrix $\mathbf{S}$ and zeroth-order singles amplitudes $t^{(0)}_{\mu_1}$:

$$\begin{aligned} \lambda^p_{\mu a} &= P_{\mu a} - L_{\mu i} t^i_a S_{a' a}, & \lambda^p_{\mu i} &= L_{\mu i} \\ \lambda^h_{\mu a} &= P_{\mu a}, & \lambda^h_{\mu i} &= L_{\mu i} + P_{\mu a} t^i_a. \end{aligned} \tag{2.23}$$

These coefficient matrices are the origin for the fifth diagrammatic rule in Sec. 1.3, stating that only vertical lines between the *bra* side and an interaction line are similarity transformed. Looking closely at the four transformation matrices reveals that only outgoing lines which point upwards ($\lambda^p_{\mu a}$) and incoming lines which point downwards ($\lambda^h_{\mu i}$) include the singles amplitudes. The

only possibility for such lines are vertical lines connecting the *bra* side with an interaction line. Vertical lines going out ($\lambda^p_{\mu i}$) or coming in ($\lambda^h_{\mu a}$) from below of an interaction line are not dressed.

The density fitting approximation can still be applied in a straightforward way,

$$(pq\hat{|}rs) \approx (pq\hat{|}P)\hat{c}^{rs}_P. \tag{2.24}$$

As a result, the three-indexed quantities are similarity transformed. The dressed Fock matrix is defined as

$$\hat{f}_{pq} = \hat{h}_{pq} + \sum_i \Big(2(ii\hat{|}pq) - (iq\hat{|}pi)\Big). \tag{2.25}$$

In this work, constraints according to the local ansatz presented in Sec. 2.4.4 are imposed on the zeroth-order doubles amplitudes $\mathbf{T}_2^{(0)}$:

$$\mathbf{T}_2^{(0)} = \frac{1}{2}\sum_{ij\in P_0}\sum_{ab\in[ij]} t^{ij}_{ab}\tau^{ab}_{ij}, \tag{2.26}$$

with P_0 refering to the (restricted) ground state pair list. The pair list and domains for the ground state are determined on basis of spatial locality arguments. Detailed information for the test calculations can be found in Section 2.5.

The CC2 Jacobian for ionized states is obtained by differentiation of the time-averaged second-order quasienergy Lagrangian for the CC2 model, yielding

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} \mathbf{A}_{\frac{1}{2}\frac{1}{2}} & \mathbf{A}_{\frac{1}{2}\frac{3}{2}} \\ \mathbf{A}_{\frac{3}{2}\frac{1}{2}} & \mathbf{A}_{\frac{3}{2}\frac{3}{2}} \end{pmatrix} \\ &= \begin{pmatrix} \left\langle \tilde{\mu}_{\frac{1}{2}} \middle| \left[\hat{\mathbf{H}}, \tau_{\nu_{\frac{1}{2}}}\right] \exp(\mathbf{T}_2^{(0)}) \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{1}{2}} \middle| \left[\hat{\mathbf{H}}, \tau_{\nu_{\frac{3}{2}}}\right] \middle| 0 \right\rangle \\ \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{W}}, \tau_{\nu_{\frac{1}{2}}}\right] \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{1}{2}} \middle| \left[\mathbf{F}, \tau_{\nu_{\frac{3}{2}}}\right] \middle| 0 \right\rangle \end{pmatrix} \end{aligned} \tag{2.27}$$

The ionization potentials $\omega_{\bar{I}}$ for the lowest ionized states $|\bar{I}\rangle$ are obtained by

solving the right eigenvalue problem

$$\mathbf{A}R(\bar{I}) = \omega_{\bar{I}}\mathbf{M}R(\bar{I}). \tag{2.28}$$

To this end a Davidson procedure generalized to nonsymmetric matrices is employed [46, 47] such that only matrix trial-vector products $\mathfrak{V}(I) = \mathbf{A}U(I)$, rather than the full matrix $\mathbf{A}$, are needed. Note that $\bar{I} \in \{1, \ldots, N_{\bar{I}} \leq N_{Dav}\}$ denotes a particular ionized state. $N_{\bar{I}}$ is equal to the number of states treated in a multistate calculation. On the other hand, $I \in \{1, \ldots, N_{Dav}\}$ corresponds to a certain basis vector of the Davidson subspace, which, in turn, belongs to a certain $\bar{I}$. At the start of every calculation and at a refresh of the subspace $N_{Dav} = N_{\bar{I}}$. With every iteration basis vectors are added to the Davidson subspace and N_{Dav} grows accordingly. No state-specific local approximations are introduced on the trial vectors $U(I)$ and eigenvectors $R(\bar{I})$ at the CC2 level. Only the truncation of the zeroth-order doubles amplitudes is utilized in the computation of the intermediate V_{ia}^P (cf. Eq. (2.31)). Moreover, the locality in the orbitals is exploited for prescreening in the evaluation of the diagrams given below.

Diagrams for the CC2 Jacobian are shown in Fig. (2.1). From these the expressions for the right matrix trial-vector product $\mathfrak{V}(I) = \mathbf{A}U(I)$ are obtained. The final working equations, including the factorization of ERIs according to the DF approximation, are

$$\mathfrak{v}^i = -u^j \hat{f}_{ji} + u^k Z_{ik} + \hat{f}_{jb}\tilde{u}_b^{ji} - {}^{\mathrm{fc}}W_k^P(ki\hat{|}P), \tag{2.29}$$

$$\mathfrak{v}_a^{ij} = \hat{c}_{ai}^P \hat{B}_j^P + f_{ab}u_b^{ij} - S_{aa'}u_{a'}^{ik}f_{kj} - S_{aa'}u_{a'}^{kj}f_{ki}, \tag{2.30}$$

with the intermediates

$$\begin{aligned} Z_{ik} &= -(kc|P)V_{ic}^P, & V_{ia}^P &= \tilde{t}_{ab}^{ik}c_{kb}^P \\ {}^{\mathrm{fc}}W_i^P &= c_{kb}^P\tilde{u}_b^{ki}, & \hat{B}_i^P &= -u^k(ki\hat{|}P). \end{aligned} \tag{2.31}$$

Amplitudes and trial vectors decorated with a tilde correspond to contravariant

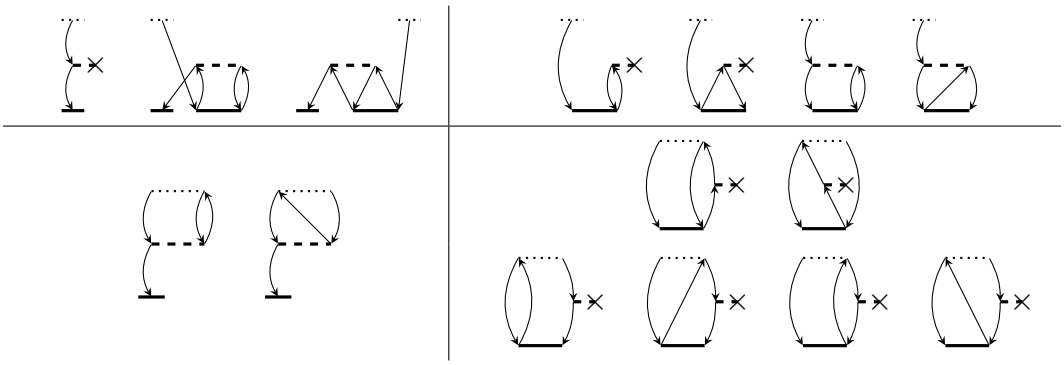

Figure 2.1.: Covariant diagrams of the CC2 Jacobian arranged in corresponding block structure and in order according to Eqs. (2.29) and (2.30). To each summand of these equations belong two consecutive diagrams, except the first one of $\mathbf{A}_{\frac{1}{2}\frac{1}{2}}$. Using contravariant `bra` CSFs cancels the second diagram of each couple of diagrams in $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$.

`bra` functions as defined in Eq. (2.9). They are calculated according to

$$\tilde{t}_{ab}^{ij} = 2t_{ab}^{ij} - t_{ab}^{ji}, \quad \text{and} \quad \tilde{u}_{a}^{ij} = 2u_{a}^{ij} - u_{a}^{ji}. \tag{2.32}$$

For a clearer formulation the explicit dependence of trial vectors U and products $\mathfrak{V}$ on the ionized states is dropped.

As is evident from Fig. (2.1) fifteen covariant diagrams contribute to the CC2 Jacobian. However, only eight summands remain in the final working equations (2.29) and (2.30). This is an effect of the contravariant `bra` CSFs and their corresponding contravariant amplitudes and trial vectors (Eq. (2.32)), as well as their ability to reduce the number of actual expressions in the working equations.

As a sidenote, the right matrix trial-vector produt $\mathfrak{V}$ for the ADC(2), also known as TD-UCC[2]-H [8], Jacobian $\mathbf{A}$ is very similar to Eqs. (2.29)-(2.31). Only the $\mathfrak{v}^i$ part differs, reading instead

$$\mathfrak{v}^i = -u^j f_{ji} + \frac{1}{2} u^k (Z_{ik} + Z_{ki}) - {}^{\mathrm{fc}}W_k^P(ki|P) \tag{2.33}$$

and all integrals and Fock matrix elements are undressed. Since ADC(2) relies

on a MP2 ground state calculation the singles amplitudes $t_{\mu_1}^{(0)}$ are zero and hence ADC(2) is Hermitian.

2.4. Additional higher order diagrams

The CC2 model for ionized states as specified in Sec. 2.3 does by itself not provide ionization potentials of satisfactory accuracy. It is, however, used to generate initial guesses for the right eigenvectors $R(\bar{I})$, as well as initial state-specific local approximations. Due to the generation of an electron hole through the ionization process, orbital relaxation effects are expected to be of greater importance for ionization potentials than for electronic excitation energies, where the CC2 model already provides acceptable accuracy for many applications.

In order to improve on the CC2 model, higher-order diagrams are added to the CC2 Jacobian, while still sticking to the $m = \frac{1}{2} \oplus \frac{3}{2}$ excitation space. From Eq. (2.27) it is clear that the $\mathbf{A}_{\frac{1}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{1}{2}\frac{3}{2}}$ submatrices are already complete in the sense that they contain all possible diagrams, whereas for the submatrices $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ this is not the case. The latter are only correct to first- and zeroth-order with respect to the fluctuation potential $\mathbf{W}$, respectively. Therefore, the order of these submatrices is increased step by step. For further reference the methods of this hierarchy are called IP-CCSD[k]$_{\mathrm{CC2}}$, with $k \in [0, 1, 2, \mathrm{f}]$. This acronym implies that $\mathbf{A}$ used in the eigenvalue problem (2.28) is an approximation to the CCSD Jacobian, correct to order k with respect to $\mathbf{W}$ on top of a CC2 ground state calculation. Consequently, the pure CC2 approach corresponds to the acronym IP-CCSD[0]$_{\mathrm{CC2}}$. IP-CCSD[f]$_{\mathrm{CC2}}$ as well as IP-CCSD$_{\mathrm{CC2}}$ denote the full CCSD Jacobian.

2.4.1. IP-CCSD[1]$_{CC2}$

In a first step the orders of $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ are both increased by one:

$$\mathbf{A}_{\frac{3}{2}\frac{1}{2}} = \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{W}} + \left[\hat{\mathbf{W}}, \mathbf{T}_2^{(0)} \right], \tau_{\nu_{\frac{1}{2}}} \right] \middle| 0 \right\rangle, \tag{2.34}$$

$$\mathbf{A}_{\frac{3}{2}\frac{3}{2}} = \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{F}} + \hat{\mathbf{W}}, \tau_{\nu_{\frac{3}{2}}} \right] \middle| 0 \right\rangle. \tag{2.35}$$

This implies the addition of two related third-order terms to the $m = \frac{3}{2}$ amplitude equations in the time-averaged second-order quasienergy CC2 Lagrangian, from which the Jacobian is obtained by differentiation. Note that the lowest orders with respect to $\mathbf{W}$ of $\mathbf{T}_{\frac{1}{2}}^{(1)}$ and $\mathbf{T}_{\frac{3}{2}}^{(1)}$ and related multipliers are 0 and 1, respectively. Due to the extra terms in the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ block of the Jacobian, this block is no longer diagonal in the canonical basis, which prevents the partitioning of the eigenvalue problem to an effective $m = \frac{1}{2}$ eigenvalue problem and, thus, the application of the Laplace transformation (cf. Section 1.2.3). For that reason, the Fock matrix in $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ is dressed, including further higher-order terms at no additional cost. Corresponding diagrams are presented in Figures (B.1) and (B.2), respectively, in the Appendix.

Extending IP-CCSD[0]$_{CC2}$ just according to Eq. (2.35), i.e., without adding the second-order contribution to $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ in Eq. (2.34) leads to a method for which the ionization potentials of eigenstates with predominantly $m = \frac{3}{2}$ character are treated formally at first order with respect to the fluctuation potential, rather than at zeroth order as in CC2 or ADC(2). However, test calculations have shown that a method with such a Jacobian is not really superior to IP-CCSD[0]$_{CC2}$: sometimes the ionization potentials are better than those of IP-CCSD[0]$_{CC2}$, but often worse. A similar behaviour was also found for excitation energies of the ADC(2)-x method [48], where, analogously, just the doubles-doubles block of the ADC(2) Jacobian is augmented by the first order term $\langle \tilde{\mu}_2 | [\mathbf{W}, \tau_{\nu_2}] | 0 \rangle$ [49, 50]. IP-CCSD[1]$_{CC2}$, on the other hand, is clearly superior to IP-CCSD[0]$_{CC2}$, as demonstrated in Section 2.5.

Due to the replacement of the undressed by the dressed Fock operator in $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$, the undressed Fock matrix elements in Eq. (2.30) have to be substituted

by their dressed counterparts. The other two higher-order terms in Eqs. (2.34) and (2.35), defining the IP-CCSD[1]$_{\mathrm{CC2}}$ method, add the following terms to the IP-CCSD[0]$_{\mathrm{CC2}}$ matrix trial-vector product given in Eq. (2.30):

$$\begin{aligned}{}^{[1]}\Delta\mathfrak{v}_a^{ij} = S_{aa'}\left(t_{a'b}^{ik}\hat{Y}_b^{kj} + t_{ba'}^{jk}\hat{Y}_b^{ki} - V_{ia'}^P\hat{B}_j^P\right) - \hat{Y}_{abc}t_{bc}^{ij} + (ai\hat{|}P)^{\mathrm{fc}}W_j^P \\ + S_{aa'}u_{a'}^{kl}(ki\hat{|}P)\hat{c}_{lj}^P - u_b^{kj}(ab\hat{|}P)\hat{c}_{ki}^P - u_b^{ik}(ab\hat{|}P)\hat{c}_{kj}^P, \end{aligned} \tag{2.36}$$

with the new intermediates

$$\hat{Y}_a^{ij} = (ij\hat{|}P)B_a^P, \quad \hat{Y}_{abc} = (ab\hat{|}P)B_c^P, \quad \text{and} \quad B_a^P = u^k c_{ka}^P. \tag{2.37}$$

In contrast to standard IP-CC2 or IP-ADC(2) diagrams, which (apart from the contraction of the ground state amplitudes in Eq. (2.31) outside the Davidson iterations) scale nominally at most as $\mathcal{O}(\mathcal{N}^4)$ with molecular size $\mathcal{N}$, these additional diagrams scale nominally as $\mathcal{O}(\mathcal{N}^5)$. In Section 2.4.4 local correlation techniques are introduced to reduce this scaling.

2.4.2. IP-CCSD[2]$_{\mathrm{CC2}}$

A further step towards the full CCSD Jacobian is the IP-CCSD[2]$_{\mathrm{CC2}}$ method, where an additional second-order term is added to $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$:

$$\mathbf{A}_{\frac{3}{2}\frac{3}{2}} = \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{F}} + \hat{\mathbf{W}} + \left[\hat{\mathbf{W}}, \mathbf{T}_2^{(0)}\right], \tau_{\nu_{\frac{3}{2}}}\right] \middle| 0 \right\rangle. \tag{2.38}$$

Corresponding diagrams are illustrated in Figure B.3 in the Appendix. The Jacobian of the IP-CCSD[2]$_{\mathrm{CC2}}$ method is identical to that of the EOMIP-CCSD(2) method presented by Stanton and Gauss [37], yet the latter is based on a MP2 rather than a CC2 ground state calculation. Consequently, EOMIP-CCSD(2) corresponds to the acronym IP-CCSD[2]$_{\mathrm{MP2}}$. The IP-CCSD[2]$_{\mathrm{CC2}}$

adds the following further terms to Eqs. (2.30) and (2.36):

$$
\begin{aligned}
{}^{[2]}\Delta\mathfrak{v}_a^{ij} = S_{aa'} \Big\{ & Z_{ik} u_{a'}^{kj} + Z_{jk} u_{a'}^{ik} + Z_{a'c} u_c^{ij} + V_{ia'}^P \, {}^{\mathrm{int}}W_j^P \\
& + \left(t_{a'b}^{il} u_c^{jk} + t_{a'b}^{lj} u_c^{ik} - \tilde{t}_{a'b}^{il} u_c^{kj} \right) K_{bc}^{kl} - t_{a'b}^{ij} Y_b + Y^{ijkl} u_{a'}^{kl} \Big\}, \qquad (2.39)
\end{aligned}
$$

with the new intermediates

$$
\begin{aligned}
Z_{ac} &= -V_{ka}^P (kc|P), & {}^{\mathrm{int}}W_i^P &= (kb|P) \tilde{u}_b^{ki}, & K_{ab}^{ij} &= (ia|P) c_{jb}^P, \\
Y_a &= (ka|P)^{\mathrm{fc}} W_k^P, & Y^{ijkl} &= t_{bc}^{ij} K_{bc}^{kl}. & & \qquad (2.40)
\end{aligned}
$$

2.4.3. IP-CCSD[f]$_{\mathrm{CC2}}$

Finally, the full IP-CCSD$_{\mathrm{CC2}}$ method includes a further fourth-order term in $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$, containing the dressed Fock matrix:

$$
\mathbf{A}_{\frac{3}{2}\frac{1}{2}} = \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{W}} + \left[\hat{\mathbf{F}} + \hat{\mathbf{W}}, \mathbf{T}_2^{(0)} \right], \tau_{\nu_{\frac{1}{2}}} \right] \middle| 0 \right\rangle . \qquad (2.41)
$$

This leads to the additional term in the matrix trial-vector product,

$$
{}^{[f]}\Delta\mathfrak{v}_a^{ij} = -S_{aa'} \tilde{t}_{a'b}^{ij} \hat{f}_{kb} u^k. \qquad (2.42)
$$

Corresponding diagrams are presented in Figure B.4 in the Appendix. The additional computational effort to go from IP-CCSD[2]$_{\mathrm{CC2}}$ to IP-CCSD$_{\mathrm{CC2}}$ is only minor. For a MP2 ground state $\hat{f}_{kb}$ is zero and, therefore, Eq. (2.42) is zero. Hence, IP-CCSD[2]$_{\mathrm{MP2}}$ is already equivalent to the complete method IP-CCSD[f]$_{\mathrm{MP2}}$.

Table 2.1 compiles the hierarchy of IP-CCSD$[k]_M$ methods with corresponding synonyms, if available. Furthermore, the correctness with respect to the fluctuation potential of the ionization energies with dominant $m = \frac{1}{2}$ and $m = \frac{3}{2}$, as well as the correctness of the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ submatrices is given: all methods apart from $k = 0$ describe ionization energies with dominant $m = \frac{1}{2}$ and $m = \frac{3}{2}$ character correct through second- and first-order, respectively. The

Table 2.1.: List of individual IP-CCSD$[k]_M$ methods and eventually existing synonyms. The correctness with respect to $\mathbf{W}$ of the ionization energies with dominant $m = \frac{1}{2}$ and $m = \frac{3}{2}$ character (according to the analysis of Ref. [51]), and the correctness of the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ submatrices is given. The related nominal scaling of the computational cost of the additional Jacobian times trial vector diagrams with respect to the number of occupied (n_o), virtual (n_v), and fitting (N_f) functions is also provided.

Methods IP-CCSD$[k]_M$			Order				
k	M	Synonyms	$\omega_{\frac{1}{2}}$	$\omega_{\frac{3}{2}}$	$\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$	$\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$	Scaling[a]
0	MP2[b]	IP-ADC(2)	2	0	1	0	$n_o^2 n_v N_f$
1	MP2		2	1	2	1	$n_o^2 n_v^3$
2	MP2	EOMIP-CCSD(2)	2	1	2	2	$n_o^3 n_v^2$
0	CC2	IP-CC2	2	0	1	0	$n_o^2 n_v N_f$
1	CC2		2	1	2	1	$n_o^2 n_v^3$
2	CC2		2	1	2	2	$n_o^3 n_v^2$
f	CC2		2	1	2	2	$n_o^2 n_v^2$
f	CCSD	EOMIP-CCSD	2	1	2	3	$n_o^2 n_v^3$

[a] The scaling of the ground state calculation is $n_o^2 n_v^2 N_f$ for all methods except EOMIP-CCSD, where it is $n_o^2 n_v^4$.

[b] For $k = 0$ and $M =$ MP2 the Jacobian is symmetrized according to Eq. (2.33).

highest nominal scaling of the computational cost of the Jacobian times trial vector diagrams, with which the method is augmented by increasing k, is also given in Table 2.1. Note however that the nominal scaling of the EOMIP-CCSD method is governed by the ladder diagrams of the ground state CCSD calculation and thus has an overall nominal scaling of $n_o^2 n_v^4$.

Within the excitation space $m = \frac{1}{2} \oplus \frac{3}{2}$, the method can only be further

improved by going to higher-order ground state methods, i.e., by inclusion of higher-order doubles and triples amplitudes $t_{\mu_2}^{[2](0)}, t_{\mu_3}^{[2](0)}$. The effect of this on the ionization energies is presumably small, since the difficulty lies in an adequate description of the open-shell excited state. To further improve the methods, one therefore has to go to $m = \frac{1}{2} \oplus \frac{3}{2} \oplus \frac{5}{2}$ excitation spaces and beyond. This topic is discussed in chapter 3.

2.4.4. Local approximations

In Section 1.2.1 the local ansatz was introduced, together with two approximations, one restricting the occupied space, the other the virtual space. As already mentioned before, a pure distance criterion, as used for the ground state, is not a proper choice. In this section local approximations affecting the excited states are discussed.

As shown in Table 2.1, the additional terms of the methods beyond the pure IP-CC2 model exhibit a higher-order scaling of the computational cost with molecular size, namely $\mathcal{O}(\mathcal{N}^5)$ instead of $\mathcal{O}(\mathcal{N}^4)$. Consequently, these terms are much more expensive than an initial IP-CCSD$[0]_{\mathrm{CC2}}$ calculation. Note that for the CC2 ground state calculation and its amplitudes local approximations according to Eq. (2.26) are invoked.

In order to reduce the computational cost of these additional terms, local approximations focusing on the $m = \frac{3}{2}$ trial vectors $u_a^{ij}(I)$ and related matrix trial vector products $\mathfrak{v}_a^{ij}(I)$ are introduced. Primarily, the pairs ij are restricted to the pair list $P_{\bar{I}}$. To this end, after the initial IP-CCSD$[0]_{\mathrm{CC2}}$ calculation, a subset of important pairs is determined by analyzing their individual contribution,

$$d^{ij}(\bar{I}) = r_a^{ij}(\bar{I}) S_{aa'} r_{a'}^{ij}(\bar{I}), \tag{2.43}$$

to the norm of the $m = \frac{3}{2}$ part of the present approximations to the right eigenvectors $r_a^{ij}(\bar{I})$. Note that in the previous equation, the repeated LMO indices i, j are excluded from the implicit summation. The individual d^{ij} are then normalized by their sum and sorted according to decreasing magnitude. The pair list $P_{\bar{I}}$ comprises the pairs ij with largest d^{ij} until the cumulative d^{ij}

reaches a certain specified threshold κ_e. For certain pairs ij in pair list $P_{\bar{I}}$ also symmetry related pairs are included.

The individual $P_{\bar{I}}$ are modified in the course of the Davidson diagonalization such that an appropriate state-specific local approximation is attained for each ionized state $|\bar{I}\rangle$. In the critical initial phase of the first few iterations of the Davidson procedure where the lowest ionized states have to be found, $P_{\bar{I}}$ is re-built in each iteration, i.e., a refresh of the Davidson procedure is enforced. Thereafter, the $P_{\bar{I}}$ are then further re-constructed only in each refresh triggered by the usual thresholds [19].

Applying state-specific approximations without any further thoughts is counterproductive due to the mixing of states I and J, when building the small effective Jacobian $A'(\omega)$ living in the Davidson subspace:

$$A'_{IJ}(\omega_{\bar{J}}) = \tilde{u}^i(I)\mathfrak{v}^i(J) + \tilde{u}^{ij}_a(I)\mathfrak{v}^{ij}_a(J). \tag{2.44}$$

A pair ij in pair list $P_{\bar{I}}$, but *not* in $P_{\bar{J}}$ would result in a zero contribution, denying the initial importance of this pair for state $\bar{I}$. This is of course not an acceptable behaviour. Since only one virtual orbital index occurs in the vectors of the $m = \frac{3}{2}$ excitation space, they can easily be stored on disk even without truncating ij according to $P_{\bar{I}}$. Therefore, pairs $ij \notin P_{\bar{I}}$ are not kept zero in $\mathfrak{v}^{ij}_a(I)$, but still calculated at the level of IP-CCSD[0]$_{\mathrm{CC2}}$, i.e., according to Eq. (2.30) without further terms. For pairs $ij \in P_{\bar{I}}$, on the other hand, $\mathfrak{v}^{ij}_a(I)$ is calculated according to Eq. (2.30) and augmented by Eqs. (2.36), (2.39), and (2.42), depending on the level of the method. This setting enables a multistate Davidson diagonalization with state-specific local approximations.

Apart from the truncations of ground state amplitudes t^{ij}_{ab} and trial vectors $u^{ij}_a(I)$ as discussed above, locality is also utilized in the three-index ERIs $(ab\hat{|}P)$ and the intermediate $\hat{Y}_{abc}$ of Eqs. (2.36) and (2.37): due to the exponential decay in the integral with respect to the distance between the two centers of the PAOs indexed by a and b, the number of non-negligible integrals of this type scales as $\mathcal{O}(\mathcal{N}^2)$ with system size $\mathcal{N}$. The same holds also for the intermediate $\hat{Y}_{abc}$. Hence, truncating PAO pairs a and b according to an overlap criterion

decreases the size of the $(ab\hat{|}P)$ integral distribution and the $\hat{Y}_{abc}$ object, as well as the computational cost of related terms in Eq. (2.36) substantially. To this end, the maximum value S_{AB} of the corresponding block of the PAO overlap matrix $S_{ab}, \forall a \in A, b \in B$ is assigned to every center pair A, B. A simple check of S_{AB} against a threshold given in the input (the AB-threshold) determines which center pairs A, B are kept in $(ab\hat{|}P)$ and $\hat{Y}_{abc}$. Furthermore, since the intermediate $\hat{Y}_{abc}$ is contracted with the ground state doubles amplitudes t_{bc}^{ij} in Eq. (2.36) the center pairs B, C are also restricted in $\hat{Y}_{abc}$. This is achieved by setting up a list of non-vanishing center pairs B, C with $B \in [ij] \wedge C \in [ij], \forall ij \in P_0$. This yields a diagonally dominant B, C list since there is no ij including both B and C in the pair domain $[ij]$ with B being very far from C. Test calculations for the determination of the AB-threshold are summarized in Tables A.5-A.8.

The much smaller set of the all-internal three-index ERIs $(ij\hat{|}P)$ can be reduced similarly as just described for $(ab\hat{|}P)$: $(ij\hat{|}P)$ decays exponentially with respect to the distance between the LMOs indexed i and j. However, due to the orthogonality of the LMOs, an overlap criterion is inappropriate. Instead, a product of Löwdin partial charges of the LMOs i and j is used, as suggested by Kats [52, Eq. (2)], with the corresponding ij-threshold. Different ij-thresholds are tested and presented in Table A.4.

2.5. Test calculations

The IP-CCSD$[k]_M$ methods introduced in this chapter have been implemented in the `MOLPRO` program package [29]. In this section vertical ionization potentials are presented, as well as test calculations to investigate the local error. For all calculations the cc-pVDZ AO basis set [53] was used, together with the related `MP2FIT` [54] fitting basis and the `JKFIT` [55] fitting basis related to at least the cc-pVTZ AO basis. The `JKFIT` basis is used for the construction of Fock and dressed Fock matrices. The occupied orbitals were localized by employing Pipek-Mezey localization [26] already mentioned in Section 1.2.1.

Besides the IP-CCSD$[k]_M$ methods, IPs are also obtained from the Δ approach: here, IPs are obtained as the difference of total energies of the N and $N-1$ electron system. The Δ approach is expected to provide higher accuracy at a given level of the correlation treatment than the corresponding linear response or EOM method, since it explicitly includes orbital relaxation for both, the N and the $N-1$ electron system, through optimized Hartree-Fock orbitals.

2.5.1. Accuracy of the IP-CCSD$[k]_M$ hierarchy

In this Section the accuracy of the IP-CCSD$[k]_M$ hierarchy is investigated for a test set of small to medium sized molecules by comparison to the ΔCCSD(T) reference. The individual molecules and states of the test set, along with the corresponding IPs, are listed in Table A.1 of the appendix. Most of the molecules have been used already for testing the local CC linear response approach for electronic excitation energies. Their geometries can be found in the supplementary material of Ref. [33].

Davidson diagonalizations were carried out for seven states simultaneously (three for water), yet results are presented only for the 3-6 lowest states depending on the availability of a related ΔCCSD(T) result. Figure 2.2 displays the mean absolute errors (MAE) of the IPs of the various methods relative to ΔCCSD(T). The ΔMP2, ΔCCSD, and the ΔCCSD(T) reference values are obtained as a difference between a closed-shell calculation and a corresponding open-shell calculation with one electron removed from the molecule. By utilizing point group symmetry in the open-shell calculations, it is possible to obtain the ionization potential of the lowest lying ionized state in each irreducible representation of the point group. Furthermore, also EOMIP-CCSD [36] results and IPs obtained as Hartree-Fock orbital energies according to Koopmans' theorem are included. The EOMIP-CCSD calculations were performed with the `CFOUR` program package [56]. All IP-CCSD$[k]_{\mathrm{CC2}}$ and IP-CCSD$[k]_{\mathrm{MP2}}$ calculations referring to Fig. 2.2 were carried out without any local approximations and are thus equivalent to their canonical counterpart.

As can be seen from Fig. 2.2, the simplest methods, namely ADC(2) and

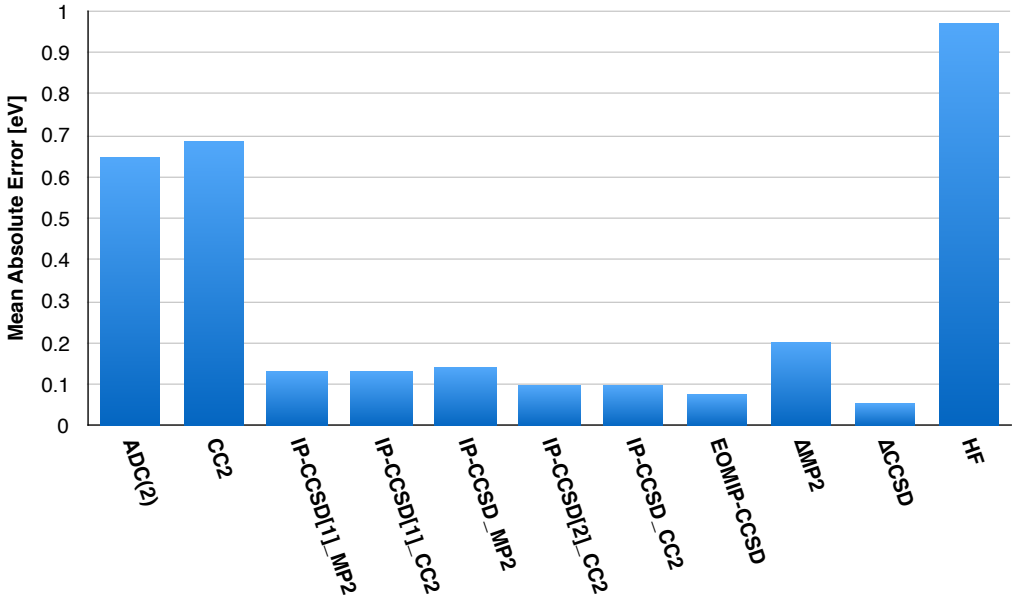

Figure 2.2.: Mean absolute errors of the vertical IPs (in [eV]) of the individual methods relative to the reference ΔCCSD(T). The MAE is calculated over all ionized states of all molecules of the test set. The AB- and ij-thresholds were both set to 10^{-8}, and the pair lists P_0 and $P_{\bar{I}}$ remained unrestricted.

CC2, exhibit large errors relative to ΔCCSD(T). The differences in the ionization potentials range from 0.3 eV to more than 1 eV. The MAEs are in the range 0.6-0.7 eV, which is only a slight improvement over pure Hartree-Fock orbital energies (Koopmans' theorem), where the MAE amounts to about 1 eV. The situation improves drastically on going from ADC(2) and CC2 to the IP-CCSD$[1]_{\mathrm{MP2}}$ or IP-CCSD$[1]_{\mathrm{CC2}}$ methods, where differences mostly smaller than 0.2 eV are observed. The MAE for these methods amounts to 0.13 eV.

IP-CCSD$[2]_{\mathrm{CC2}}$ exhibits a further improvement compared to IP-CCSD$[1]_{\mathrm{CC2}}$, and the MAE improves to 0.10 eV. The differences between IP-CCSD$[2]_{\mathrm{CC2}}$ and full IP-CCSD$_{\mathrm{CC2}}$ are very minor, both regarding accuracy and computational cost. EOMIP-CCSD, which employs CCSD rather than CC2 ground state amplitudes, is again somewhat more accurate than IP-CCSD$_{\mathrm{CC2}}$ (MAE is 0.07 eV), but also considerably more costly, since four-external ladder diagrams

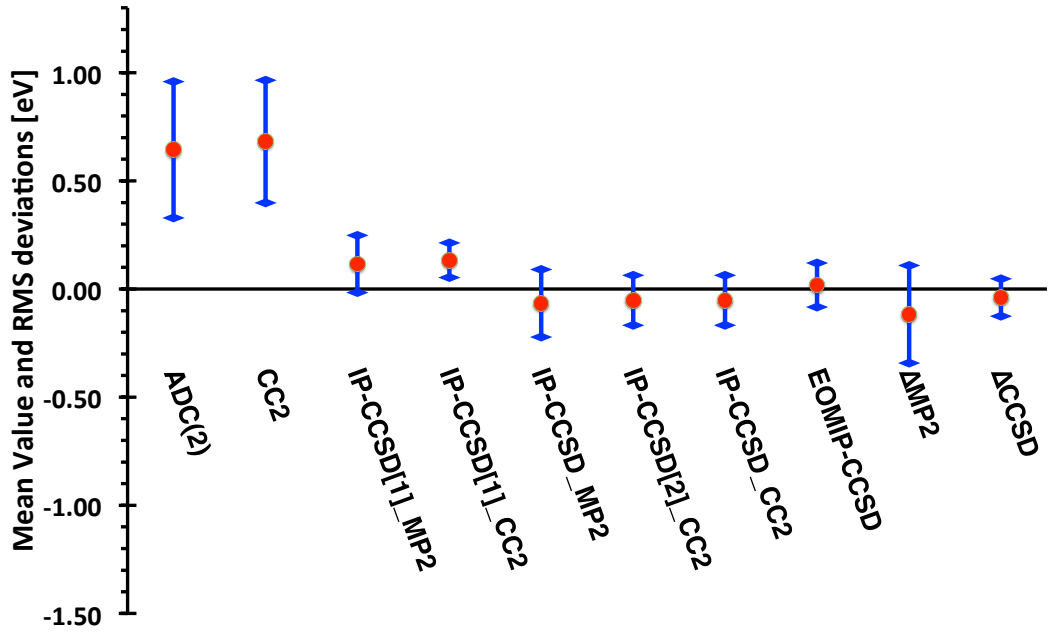

Figure 2.3.: Mean (in red) and RMS (in blue) errors of vertical IPs (in [eV]) of the individual methods relative to the reference ΔCCSD(T). The Mean and RMS errors are calculated over all ionized states of all molecules of the test set. The AB- and ij-thresholds were both set to 10^{-8} and the pair list P_0 and $P_{\bar{I}}$ remain unrestricted.

have to be evaluated.

Figure 2.3 shows the mean and RMS error of the IPs for the same methods as in Fig. 2.2, apart from the Hartree-Fock orbital energies, for which the mean±RMS error amounts to -0.95 ± 0.78 eV. It can be seen that IP-ADC(2), IP-CC2, as well as the $k = 1$ methods, IP-CCSD$[1]_{\mathrm{MP2}}$ and IP-CCSD$[1]_{\mathrm{CC2}}$, generally underestimate the ΔCCSD(T) reference value, while the $k \geq 2$ methods also overestimate it in many cases. Furthermore, the IP-CCSD$[k]_{\mathrm{CC2}}$ methods are generally somewhat more accurate than the corresponding IP-CCSD$[k]_{\mathrm{MP2}}$ methods, as is evident from the RMS errors. Therefore, inclusion of ground state singles is recommended.

2.5.2. Accuracy of local approximations

Figure 2.4 compiles, for each of the test molecules, the MAE of the IPs of local vs. non-local (=canonical) IP-CCSD$[k]_{\mathrm{MP2}}$ and IP-CCSD$[k]_{\mathrm{CC2}}$ over all computed states of the molecule. IPs of the individual molecules and states are listed in Table A.2. The orbital domains of the ground state were obtained according to the Boughton-Pulay (BP) procedure with a criterion of 0.98, and extended by the next nearest neighbour atoms forming a covalent bond with any of the atoms of the BP domain (`iext`=1 option in `MOLPRO`). The ground state pair list P_0 is truncated according to a distance criterion of 10 bohrs in all calculations. The excited state pair list $P_{\bar{I}}$ was determined as described in Section 2.4.4, employing a threshold $\kappa_e = 0.99$ (cf. Tab. A.3). The AB-threshold employed to restrict the center pair list A, B of the three-index ERIs $(ab\hat{|}P)$ and intermediate $\hat{Y}_{abc}$ was set to 10^{-2}. Test calculations indicate that such a value for the AB-threshold is a reasonable setting for basis sets without diffuse functions. Some related data is given in Tables A.4-A.8 in the appendix. Likewise, for the threshold restricting the pairs i, j in $(ij\hat{|}P)$ on the basis of Löwdin partial charges, a conservative value of 10^{-4} was used.

As can be seen from Fig. 2.4, the mean absolute local errors are all below 0.025 eV. For IP-ADC(2) and IP-CC2, the $P_{\bar{I}}$ remains untruncated, and the small local errors are solely caused by the truncation of P_0 and the ground state pair domains $[ij]$. Interestingly, the IP-CCSD$[1]_M$ method exhibits significantly larger local errors than the higher-order methods IP-CCSD$[k]_M$ with $k \geq 2$. This appears to be related to the tendency of the $k = 1$ methods to locate states with dominant or large $m = \frac{3}{2}$ character (cf. Section 3.2), which are not found by the $k \geq 2$ methods: for example, the largest errors of the $k = 1$ methods were seen for the D_2 and D_3 states of furan, i.e., 0.066 eV for M = CC2, and 0.071 eV for M = MP2, respectively. Both the D_2 state (for M = CC2) and the D_3 state (for M = MP2) are dominated by the same $\tau_{\mathrm{HOMO}}\tau^{\mathrm{LUMO}}_{\mathrm{HOMO}}$ ionization process with $m = \frac{3}{2}$ character. HOMO and LUMO are abbreviations for the highest occupied and lowest unoccupied molecular

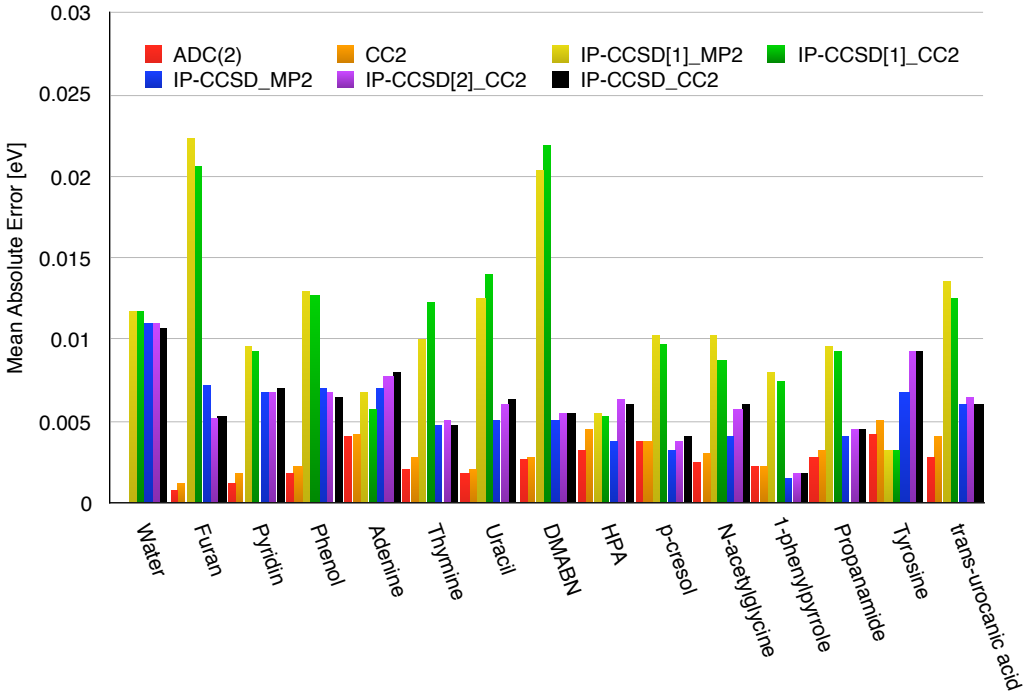

Figure 2.4.: Mean absolute local errors of the vertical IPs (in [eV]) of the individual methods, i.e., MAEs of the differences in the IPs between local and and corresponding non-local (canonical) methods (the averaging is done over all ionized states of all molecules of the test set). For the local calculations the AB- and ij-threshold were set to 10^{-2} and 10^{-4}, respectively, P_0 truncated according to a distance criterion of of $R_0 = 10$ bohrs, and $P_{\bar{I}}$ according to a threshold of $\kappa_e = 0.99$ (see text). For the non-local calculations the AB- and ij-thresholds were both set to 10^{-8} and the pair lists P_0 and $P_{\bar{I}}$ remained unrestricted.

orbital, respectively. The corresponding weight amounts to about 70%. The other methods, on the other hand, only find states with small or insignificant $m = \frac{3}{2}$ character, and the local errors are smaller, e.g., 0.005 eV for the D_2 state of furan when calculated with the IP-CCSD$[2]_{CC2}$ method. As it appears, states with dominant or substantial $m = \frac{3}{2}$ character are more sensitive to the local approximation as specified by the present settings. Nevertheless, generally, the local errors can be considered as acceptable on the scale of the

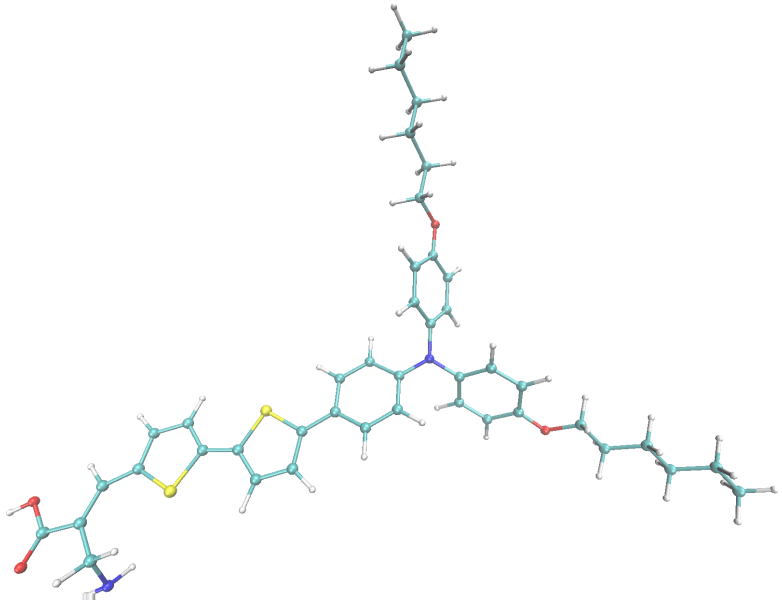

Figure 2.5.: D21L6

accuracy observed relative to ΔCCSD(T) (cf. Section 2.5.1), particularly so for the $k \geq 2$ methods.

2.5.3. Calculations on D21L6

As an example for a bigger molecule calculations on the 3-(5-(5-(4-(bis(4-(hexyloxy)phenyl)amino)phenyl)thiophene-2-yl)thiophene-2-yl)-2-cyanoacrylic acid (D21L6) dye are presented. The dye is shown in Fig. 2.5. This molecule is utilized as an organic sensitizer for solar-cell applications [57]. D21L6 comprises 98 atoms and 262 correlated electrons. The local approximations for the ground state amplitudes t_{ab}^{ij} reduce the pair list P_0 from 8646 to 2613 and the average size of the virtual space from 759 to 129 functions, i.e., the number of amplitudes from 4981 million to 48 million by two orders of magnitude. The size of the intermediate $\hat{Y}_{abc}$ reduces from 852 million elements (6.6 GB) to 160 million elements by a factor of 5.3 by virtue of the restrictions in the ab and bc ranges as discussed in Section 2.4.4. All these truncations do not depend

on the individual ionized states.

The seven lowest ionized states of D21L6 were computed with the local IP-CCSD$[k]_{\mathrm{CC2}}$ and IP-CCSD$[k]_{\mathrm{MP2}}$ methods. The results of these calculations are compiled in Table 2.2. As already seen for the previous test set of molecules, the $k = 0$ methods significantly underestimate the IPs when taking IP-CCSD[f]$_{\mathrm{CC2}}$ as the reference. Differences to IP-CCSD[f]$_{\mathrm{CC2}}$ range from 0.56-0.94 eV with MAEs of 0.79 eV for IP-CCSD[0]$_{\mathrm{CC2}}$ and 0.72 eV for IP-CCSD[0]$_{\mathrm{MP2}}$. For the $k = 1$ methods, and this in contrast to all other methods, the D_2 and D_4 states account for a large fraction of the $m = \frac{3}{2}$ character. The weights of the $m = \frac{3}{2}$ CSFs amounts to $25 - 40\%$. This behaviour persists with full $P_{\bar{I}}$ lists ($\kappa_e = 1$) and is therefore not caused by the local approximation, but a peculiarity of the $k = 1$ methods. Enhanced $m = \frac{3}{2}$ character when using the $k = 1$ methods was also observed for some of the molecules of the test set, primarily for higher states outside those entering the statistics in Section 2.5.1 and of course for the acridine radical in Section 3.2.

Limiting the statistics to states that can be clearly assigned, the MAEs of IP-CCSD[1]$_{\mathrm{CC2}}$ and IP-CCSD[1]$_{\mathrm{MP2}}$ relative to IP-CCSD[f]$_{\mathrm{CC2}}$ amount to 0.32 eV and 0.23 eV., respectively. Ionization energies without ground state singles, i.e. $M = \mathrm{MP2}$, are generally shifted to somewhat larger values, hence reducing the MAE of IP-CCSD[1]$_{\mathrm{MP2}}$. The $k = 2$ methods, on the other hand, provide results in good agreement with the reference, but also at a comparable cost as IP-CCSD[f]$_{\mathrm{CC2}}$ itself: the first Davidson diagonalization at the $k = 0$ level (used to generate a first guess of the local approximation) takes about 5 h, the subsequent second Davidson diagonalization about 15 h. At first glance it is somewhat counterintuitive that the simpler $k = 1$ methods actually take more computation time than IP-CCSD[f]$_{\mathrm{CC2}}$ itself. This can be attributed to the occurrence of enhanced $m = \frac{3}{2}$ character in the D_2 and D_4 states and a slower convergence of the Davidson diagonalization. In iteration 24 still four states are not yet converged for the $k = 1$ methods, among them D_2 and D_4, while for the other methods just a single state has not converged at that step. This slower convergence make the $k = 1$ methods computationally more expensive than the $k \geq 2$ methods.

Table 2.2.: Results for D21L6. The vertical ionization energies are given in eV. The timings for ground-state, the initial $k = 0$ diagonalization, and the second $k > 0$ diagonalization with additional terms included are all given in hours [h]. In parenthesis the number of Jacobian times trial vector products are given, which were computed during the iterative diagonalizations. Convergence criteria for the individual states during the diagonalizations are the norm of the residual (10^{-4}) and the energy change in the eigenvalue ($7.65 \cdot 10^{-6}$), i.e., the IP. The calculations were performed in parallel mode on seven AMD 6180 SE cores at 2.5 GHz.

	IP-CCSD$[k]_M$						
	$k = 0$		$k = 1$		$k = 2$		$k = \mathrm{f}$
	ADC(2)	$M = \mathrm{CC2}$	$M = \mathrm{MP2}$	$M = \mathrm{CC2}$	$M = \mathrm{MP2}$	$M = \mathrm{CC2}$	$M = \mathrm{CC2}$
	IPs [eV]						
D_0	5.414	5.317	6.082	5.968	6.354	6.257	6.258
D_1	6.516	6.400	6.935	6.817	7.285	7.186	7.188
D_2	7.363	7.316	7.811[a]	7.639[a]	8.165	8.145	8.145
D_3	7.533	7.451	7.888	7.847	8.305	8.239	8.241
D_4	7.891	7.862	8.179[b]	8.021[b]	8.703	8.700	8.701
D_5	8.237	8.146	8.436	8.340	8.860	8.793	8.794
D_6	8.271	8.195	8.476	8.415	8.997	8.923	8.924
HF + M [h]	1.8	5.2	1.9	5.1	1.9	5.0	5.1
$k = 0$ step [h]	3.9(143)	3.7(134)	4.2(143)	4.0(134)	4.0(143)	4.0(134)	4.0(134)
$k > 0$ step [h]	...	...	17.9(170)	22.8(172)	12.4(136)	15.1(137)	15.6(137)

[a] Ionized state with dominant $m = \frac{3}{2}$ character: $\tau_{\mathrm{HOMO}}\tau^{\mathrm{LUMO}}_{\mathrm{HOMO}}$, $\tau_{\mathrm{HOMO\text{-}1}}\tau^{\mathrm{LUMO}}_{\mathrm{HOMO}}$.
[b] Ionized state with dominant $m = \frac{3}{2}$ character: $\tau_{\mathrm{HOMO}}\tau^{\mathrm{LUMO}}_{\mathrm{HOMO}}$, $\tau_{\mathrm{HOMO}}\tau^{\mathrm{LUMO+1}}_{\mathrm{HOMO}}$.

In order to get a handle on the local error in this molecule an IP-CCSD[f]$_{\mathrm{CC2}}$ calculation with full pair list $P_{\bar{I}}$ ($\kappa_e = 1$) was carried out, which is the most critical local approximation in the local IP-CCSD$[k]_M$ hierarchy. Furthermore, the AB- and ij-thresholds were each reduced by two orders of magnitude relative to their default values. The effect on the IPs is quite small, the largest deviation amounts to 0.045 eV, and the MAE over the seven states to 0.011 eV. Therefore, one can conclude that the local approximations are uncritical also for such an extended molecule as D21L6.

3. Hierarchy extension to the five-half excitation space

3.1. Introduction

The highest order method IP-CCSD$_{\mathrm{CC2}}$ of the hierarchy presented in the previous chapter is already a very good tool to obtain accurate ionization energies even for large molecules. Various approximations are utilized such that it is possible to obtain, for example, the seven lowest ionization energies of D21L6 within one day. However, the hierarchy of methods in the $m = \frac{1}{2} \oplus \frac{3}{2}$ excitation space is only good for ionization processes with dominant $m = \frac{1}{2}$ character, even though the $k = 1$ methods also tend to locate states with dominant $m = \frac{3}{2}$ character. For all IP-CCSD$[k]_M$ methods the former processes are correct to second order, whereas the latter are only correct to first order (cf. Tab. 2.1). In order to further improve the accuracy of ionization energies in general and particularly those with dominant $m = \frac{3}{2}$ character the hierarchy presented so far is extended. One possible route is to stick to the excitation space $m = \frac{1}{2} \oplus \frac{3}{2}$ and introduce higher-order ground state methods, i.e., by inclusion of higher-order doubles and triples amplitudes $t_{\mu_2}^{[2](0)}, t_{\mu_3}^{[2](0)}$. The effect of this on the ionization energies is presumably small. Another, more promising way is to extend the excitation space to $m = \frac{1}{2} \oplus \frac{3}{2} \oplus \frac{5}{2}$. That extension introduces higher order contributions to the Jacobian such that $m = \frac{1}{2}$ and $m = \frac{3}{2}$ dominated ionizations are correct to third and second order, respectively. Since IPs can be used to calculate excitation energies of radicals, a hierarchy extension will generally improve ionization energies themselves, but in particular allows for

an improved description of certain types of radical excitations. This topic is addressed in the next Section.

3.2. Excitation energies of radicals

The content of this section has been published already in Ref. [33] and is reused here with some additional notes.

The application of the IP-CCSD$[k]_M$ methods presented in the previous chapter is not limited to the calculation of ionization energies. They can also be used to compute excitation energies of radicals as differences of proper ionization energies. This is illustrated in Figures 3.1 and 3.2 for different scenarios. Figure 3.1 [58] pictures different situations for a radical. In its ground state the highest occupied molecular orbital (HOMO) of the radical is singly occupied. The HOMO is then called a SOMO (singly occupied molecular orbital). An excitation of the radical allows for different scenarios:

- Type I corresponds to single excitations into the SOMO.
- Type II are single excitations from the SOMO into the virtual space.
- Type III collects excitations from doubly occupied molecular orbitals into the virtual space.

Besides solving the associated eigenvalue problem, excitation energies can also be obtained by energy differences between final and initial states. One way to arrive at any radical state is the proper ionization of the corresponding closed-shell molecule (obtained by adding an extra electron to the radical). Figure 3.2 [59] illustrates, how the determinants of the radical in Fig. 3.1 are obtained through such ionizations of closed-shell systems.

The initial and final state of an excitation SOMO $-$ n $\rightarrow$ SOMO (Type I), for example, are obtained by ionizing the closed-shell system in the $m = \frac{1}{2}$ excitation space, i.e., IPs related to τ_{HOMO} and $\tau_{\mathrm{HOMO-n}}$, with $\tau_{(\mathrm{HOMO-n})} = a_{(\mathrm{HOMO-n})\beta}$. The difference between these two IPs yields the corresponding

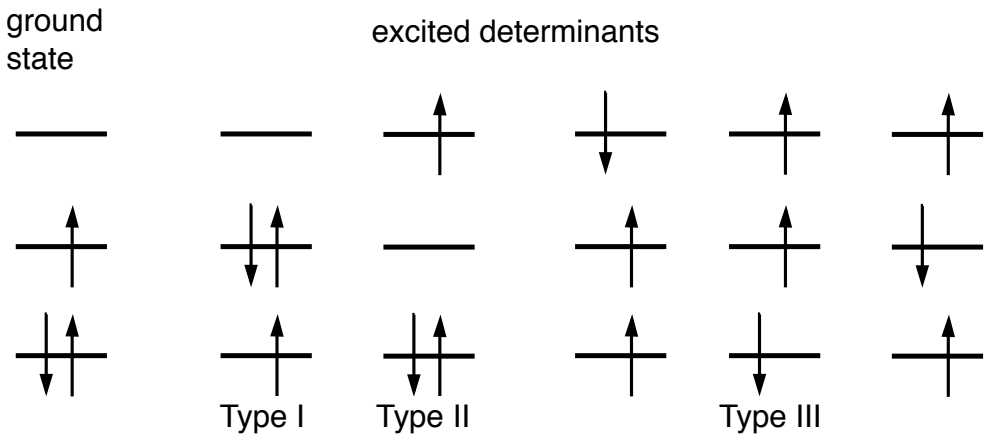

Figure 3.1.: Orbital occupation of a radical in its ground state and corresponding single excited determinants. Type I corresponds to excitations from a doubly occupied molecular orbital into the SOMO. Type II and III are excitations into the virtual space, the former exciting the electron from the SOMO, the latter exciting an electron of doubly occupied molecular orbitals. The excitation on the very right is a spin-flip case.

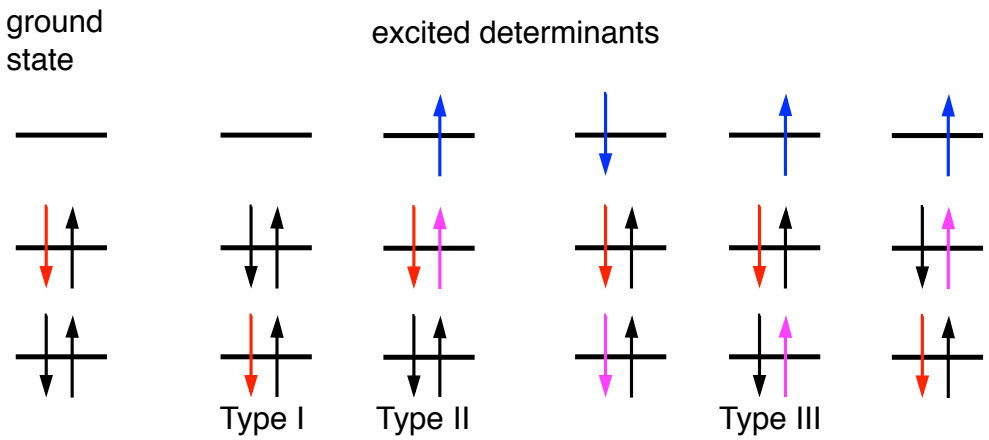

Figure 3.2.: Obtaining determinants of Fig. 3.1 from a closed-shell molecule (all non-blue arrows) through ionization. Red arrows correspond to electrons removed from the molecule. Purple (from orbital) and blue arrows (to orbital) mark an simultaneous electron excitation.

excitation energy of the radical. Such excitation energies can be expected to be well described by the IP-CCSD$[k]_M$ methods with $k > 0$ or EOMIP-CCSD, since ionization energies in the $m = \frac{1}{2}$ excitation space are correct through second order. On the other hand, excitation processes into the virtual space (Type II and III), e.g. SOMO $\rightarrow$ LUMO, require a $m = \frac{3}{2}$ ionization process, e.g. $\tau_{\mathrm{HOMO}}\tau^{\mathrm{LUMO}}_{\mathrm{HOMO}}$. Such $m = \frac{3}{2}$ dominated ionization processes are expected to be described rather poorly by IP-CCSD$[k]_M$ or EOMIP-CCSD and an extension to the $m = \frac{5}{2}$ excitation space may be required for sufficient accuracy.

To explore this issue in the context of the IP-CCSD$[k]_M$ methods, excitation energies are calculated for the H_2O^+ and the acridine radical of Ref. [60]. In Table 3.1 excitation energies calculated via IP differences from IP-CCSD$[k]_{\mathrm{CC2}}$, $k = 1$ and $k = \mathrm{f}$, are compared with CASSCF and CASPT2 excitation energies. The latter were computed with the `MOLPRO` (H_2O^+) and the `MOLCAS v.7` [61] (acridine radical) program packages.

The D_1 and D_2 states of H_2O^+ can be generated by $m = \frac{1}{2}$ ionization processes. The excitation energies calculated via the IP-CCSD$[k]_{\mathrm{CC2}}$, $k = 1$ and $k = \mathrm{f}$, IPs are, as expected, in good agreement with those obtained through CASPT2. The much higher lying D_3 state, on the other hand, corresponds to the SOMO $\rightarrow$ LUMO (Type II) excitation and requires a $m = \frac{3}{2}$ ionization process. Here, the agreement between CASPT2 and IP-CCSD$[k]_{\mathrm{CC2}}$ is much worse. Interestingly, the error is much larger for the full method than for the $k = 1$ method. Moreover, the latter finds another ionized state corresponding to $\tau_{(\mathrm{HOMO}-1)}\tau^{\mathrm{LUMO}}_{(\mathrm{HOMO}-1)}$ (with an excitation energy of 14.3 eV), which is not found by the other methods. The EOMIP-CCSD method of the `CFOUR` program package [56] yields very similar results as the $k = \mathrm{f}$ method.

The computations on the acridine radical were performed in the same geometry as in Ref. [60]. Here, the radical is neutral and the reference for the IP-CCSD$[k]_{\mathrm{CC2}}$ calculations thus a closed-shell anion. The two lowest lying excited states of the radical possess mainly SOMO $\rightarrow$ LUMO and SOMO $\rightarrow$ (LUMO+1) character, respectively, and are therefore dominated by $m = \frac{3}{2}$ ionization processes acting on the closed-shell anion. The D_3 state, on the

Table 3.1.: Excitation energies (in [eV]) of the lowest three doublet states of H_2O^+ and the acridine radical calculated with CASSCF, CASPT2, and via IP differences computed with IP-CCSD$[k]_{CC2}$, $k = 1$ and $k = \mathrm{f}$. The occupation string of the dominant configuration state function in the CASSCF wavefunction is also given. The CASSCF calculations were state averaged over four roots. For H_2O^+ and the acridine radical a full valence active space, and a 9 electrons in 10 orbitals CAS space was used, respectively. No point group symmetry was imposed.

				IP-CCSD$[k]_{CC2}$	
	Dominant CSF	CASSCF	CASPT2	$k = 1$	$k = \mathrm{f}$
			H_2O^+		
D_1	22a200	2.294	2.221	2.198	2.400
D_2	2a2200	7.282	6.889	6.855	6.999
D_3	2220a0	15.040	15.039	19.004	23.145
	2202a0			14.268	
			Acridine radical		
D_1	22220a0000	2.688	2.225	1.944	...
D_2	222200a000	3.430	2.830	2.868	...
D_3	222a200000	3.613	2.878	2.834	2.972

other hand, has (SOMO-1) $\rightarrow$ SOMO character and is well described by an $m = \frac{1}{2}$ ionization process. As is evident from Table 3.1, the $k = 1$ method yields excitation energies in good agreement with CASPT2 for all three states. For the $k = \mathrm{f}$ method, on the other hand, the D_1 and D_2 states are not found, even when utilizing the final $k = 1$ eigenvectors as start vectors in the Davidson diagonalization of the $k = \mathrm{f}$ calculation.

Generally, the $k = 1$ methods have a much increased tendency to locate $m = \frac{3}{2}$ dominated states than the $k = 0$ methods, and the more complete $k > 1$ methods. The origin of this peculiar behaviour must be the first-order correction in the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ submatrix, Eq. (2.35), since omission of the second-order correction in the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ submatrix, Eq. (2.34), does not change the behaviour of the $k = 1$ method in this respect. On the other hand, the higher-order terms of Eq. (2.38) apparently suppress this behaviour again in the $k > 1$ methods.

3.3. IP-CC3$_{CC2}$ theory

In this section the theory for extending the IP-CCSD$[k]_{\mathrm{CC2}}$ hierarchy to the $m = \frac{5}{2}$ excitation space is presented. The volume of this topic is probably enough for another thesis, therefore time limitation restricted the development in the course of this work mainly to the theory part. Some preliminary coding was already done, but not to a point where actual numbers can be presented.

The enlarged excitation space adds further elements to the Lagrangian given in Eq. (2.10). Including these elements, the new Lagrangian reads

$$\begin{aligned}\{^{2n+1}L^{(2)}(t)\}_T = \sum_{k=-n}^{n} \Bigg[& \left\langle 0 \middle| \left[\mathbf{V}(-\omega_k), \mathbf{T}^{(1)}_{\frac{1}{2}}(\omega_k) \right] \middle| \mathrm{CC} \right\rangle \\ & + \lambda^{(0)}_{\mu_m} \left\langle \tilde{\mu}'_m \middle| \left[\mathbf{V}(-\omega_k), \mathbf{T}^{(1)}_{\frac{1}{2}}(\omega_k) + \mathbf{T}^{(1)}_{\frac{3}{2}}(\omega_k) + \mathbf{T}^{(1)}_{\frac{5}{2}}(\omega_k) \right] \middle| \mathrm{CC} \right\rangle \\ & - \lambda^{(1)}_{\mu_m}(-\omega_k)\omega_k t^{(1)}_{\nu_l}(\omega_k) \left\langle \tilde{\mu}_m \middle| \tau_{\nu_l} \right| \mathrm{CC} \rangle \\ & + \lambda^{(1)}_{\mu_m}(-\omega_k) \left\langle \tilde{\mu}'_m \middle| \mathbf{V}(\omega_k) + \left[\mathbf{H}^{(0)}, \mathbf{T}^{(1)}_{\frac{1}{2}}(\omega_k) + \mathbf{T}^{(1)}_{\frac{3}{2}}(\omega_k) + \mathbf{T}^{(1)}_{\frac{5}{2}}(\omega_k) \right] \middle| \mathrm{CC} \right\rangle \Bigg] \end{aligned} \tag{3.1}$$

with particle-rank index $m = 1, 2$ for zeroth-order multipliers $\lambda^{(0)}_{\mu_m}$. For first-order multipliers $\lambda^{(1)}_{\mu_m}(\omega_k)$ and amplitudes $t^{(1)}_{\mu_m}(\omega_k)$ index $m = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$ is half-integer. The new cluster operator $\mathbf{T}^{(1)}_{\frac{5}{2}}$ is second order with respect to the fluctuation potential [7].

3.3.1. Co- and contravariant configuration state functions

Besides a new cluster operator, $\mathbf{T}^{[2](1)}_{\frac{5}{2}}$, the $m = \frac{5}{2}$ excitation space also entails corresponding co- and contravariant CSFs, $\mu_{\frac{5}{2}}$ and $\tilde{\mu}_{\frac{5}{2}}$, respectively. The covariant CSF is

$$|\mu^{ijk}_{ab}\rangle = \tau^{ijk}_{ab}|0\rangle, \tag{3.2}$$

with τ^{ijk}_{ab} already defined in Eq. (2.8). However, the six different CSFs, obtained through permuting orbital indices i, j, k, are linearly dependent:

$$\mu^{ijk}_{ab} + \mu^{ikj}_{ab} + \mu^{jik}_{ab} + \mu^{jki}_{ab} + \mu^{kij}_{ab} = -\mu^{kji}_{ab}. \tag{3.3}$$

This linear dependence effects the contravariant CSF, which is constructed as a linear combination of covariant CSFs,

$$\tilde{\mu}_{ab}^{ijk} = a\mu_{ab}^{ijk} + b\mu_{ab}^{ikj} + c\mu_{ab}^{jik} + d\mu_{ab}^{jki} + e\mu_{ab}^{kij} + f\mu_{ab}^{kji}, \tag{3.4}$$

with variables $a-f$. Solving the equation system reveals that the orthogonality condition, Eq. (1.9), cannot be fulfilled for all CSFs simultaneously. A new, additional condition is imposed by using the orthogonality condition for all overlaps except one. There is a certain arbitrariness in the choice of the latter, namely the choice of the permutation of the covariant CSF. In close analogy to LCCSD(T) of Ref. [62, Eq. (8)] the new conditions are:

$$\begin{aligned}
&\langle\tilde{\mu}_{ab}^{ijk}|\mu_{ab}^{ijk}\rangle = 1 \\
&\langle\tilde{\mu}_{ab}^{ijk}|\mu_{ab}^{ikj}\rangle = \langle\tilde{\mu}_{ab}^{ijk}|\mu_{ab}^{jik}\rangle = \langle\tilde{\mu}_{ab}^{ijk}|\mu_{ab}^{kij}\rangle = \langle\tilde{\mu}_{ab}^{ijk}|\mu_{ab}^{kji}\rangle = 0 \\
&\langle\tilde{\mu}_{ab}^{ijk}|\mu_{ab}^{jki}\rangle = -1.
\end{aligned} \tag{3.5}$$

For these conditions the corresponding contravariant CSF is

$$\tilde{\mu}_{ab}^{ijk} = \frac{1}{3}\left(\mu_{ab}^{ijk} - \mu_{ab}^{jki}\right). \tag{3.6}$$

Note, that the contravariant CSF contains the two covariant CSFs which are *not* orthogonal to $\tilde{\mu}_{ab}^{ijk}$ with corresponding signs.

3.3.2. IP-CC3$_{\mathrm{CC2}}$ Jacobian

A full inclusion of all diagrams caused by the $m = \frac{5}{2}$ excitation space is computationally very complex. Therefore, a CC3 like approach is chosen, i.e., inclusion of all terms in the $m = \frac{1}{2}, \frac{3}{2}$ amplitude equations in the time-averaged second-order Lagrangian, Eq. (3.1), from which the Jacobian is obtained. The $m = \frac{5}{2}$ amplitude equation in the Lagrangian, on the other hand, is restricted to fourth order. For further reference the method is called IP-CC3$_{\mathrm{CC2}}$. Similar to the IP-CCSD$[k]_M$ hierarchy, this acronym implies an approximation to the CC3 Jacobian on top of a CC2 ground state calculation. It is important to

note that the approximated IP-CC3 Jacobian misses the ground state triples contribution in the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ block of the Jacobian for ground state methods like MP2, CC2 or CCSD. Only ground state methods containing triples excitations can overcome the absence of this contribution.

Furthermore,the LMO basis introduces some complications due to the non-diagonality of the Fock matrix, which leads to couplings between $t^{(1)}_{\frac{5}{2}}$ amplitudes and the internal-internal block of the Fock matrix. In order to simplify the theory, such couplings are neglected in a first approach. Such an approach was already applied to the ground state method LCCSD(T0) [62, 63]. The acronym (T0) indicates that only 0th-order Fock matrix elements, i.e., diagonal elements, of the internal-internal block are coupled with the $t^{(1)}_{\mu_3}$ amplitudes. The absence of such couplings in the IP case is indicated by superscript 0 added to the acronym: IP-CC3$^0_{\mathrm{CC2}}$.

Moving on to the CC3 Jacobian, its block structure is extended with five new blocks compared to the CC2 Jacobian:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{\frac{1}{2}\frac{1}{2}} & \mathbf{A}_{\frac{1}{2}\frac{3}{2}} & \mathbf{A}_{\frac{1}{2}\frac{5}{2}} \\ \mathbf{A}_{\frac{3}{2}\frac{1}{2}} & \mathbf{A}_{\frac{3}{2}\frac{3}{2}} & \mathbf{A}_{\frac{3}{2}\frac{5}{2}} \\ \mathbf{A}_{\frac{5}{2}\frac{1}{2}} & \mathbf{A}_{\frac{5}{2}\frac{3}{2}} & \mathbf{A}_{\frac{5}{2}\frac{5}{2}} \end{pmatrix} \tag{3.7}$$

$$= \begin{pmatrix} \left\langle \tilde{\mu}_{\frac{1}{2}} \middle| \left[\hat{\mathbf{H}}, \tau_{\nu_{\frac{1}{2}}}\right] \exp(\mathbf{T}_2^{(0)}) \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{1}{2}} \middle| \left[\hat{\mathbf{H}}, \tau_{\nu_{\frac{3}{2}}}\right] \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{1}{2}} \middle| \left[\hat{\mathbf{W}}, \tau_{\nu_{\frac{5}{2}}}\right] \middle| 0 \right\rangle \\ \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{W}} + [\hat{\mathbf{F}} + \hat{\mathbf{W}}, \mathbf{T}_2^{(0)}], \tau_{\nu_{\frac{1}{2}}}\right] \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{F}} + \hat{\mathbf{W}} + [\hat{\mathbf{W}}, \mathbf{T}_2^{(0)}], \tau_{\nu_{\frac{3}{2}}}\right] \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{3}{2}} \middle| \left[\hat{\mathbf{F}} + \hat{\mathbf{W}}, \tau_{\nu_{\frac{5}{2}}}\right] \middle| 0 \right\rangle \\ \left\langle \tilde{\mu}_{\frac{5}{2}} \middle| \left[[\hat{\mathbf{W}}, \mathbf{T}_2^{(0)}], \tau_{\nu_{\frac{1}{2}}}\right] \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{5}{2}} \middle| \left[\hat{\mathbf{W}}, \tau_{\nu_{\frac{3}{2}}}\right] \middle| 0 \right\rangle & \left\langle \tilde{\mu}_{\frac{5}{2}} \middle| \left[\mathbf{F}, \tau_{\nu_{\frac{5}{2}}}\right] \middle| 0 \right\rangle \end{pmatrix}$$

Diagrams of the $\mathbf{A}_{\frac{1}{2}\frac{5}{2}}$, $\mathbf{A}_{\frac{3}{2}\frac{5}{2}}$, $\mathbf{A}_{\frac{5}{2}\frac{5}{2}}$, $\mathbf{A}_{\frac{5}{2}\frac{3}{2}}$, $\mathbf{A}_{\frac{5}{2}\frac{1}{2}}$ blocks of the Jacobian are illustrated in Figures B.5-B.9, respectively. The new $\mathbf{A}_{\frac{1}{2}\frac{5}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{5}{2}}$ blocks contain already all possible diagrams within the given excitation space. Blocks with contravariant $\langle\tilde{\mu}_{\frac{5}{2}}|$ `bra` functions, on the other hand, are restricted to fourth order with respect to the fluctuation potential in the Lagrangian, Eq. (3.1). Higher order terms are omitted in this approach. From here, IP-CC3$_{\mathrm{CC2}}$ can get systematically improved as it was done for regular IP-CC2. However, already the most basic approach is computationally very demanding and higher

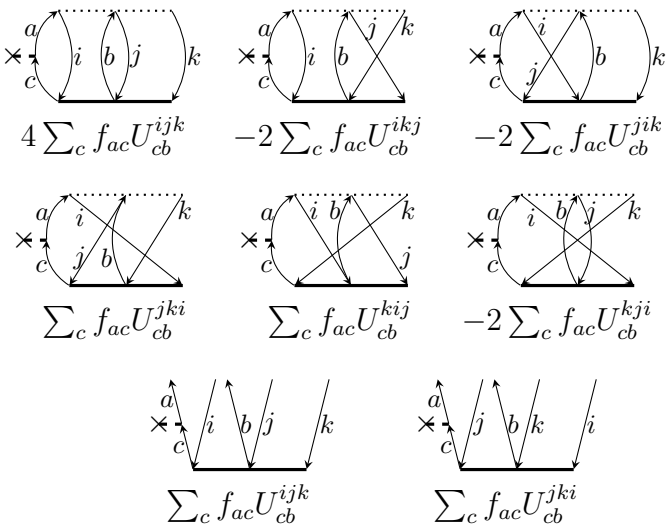

Figure 3.3.: All six covariant and the two remaining contravariant diagrams of the $\mathbf{A}_{\frac{5}{2}\frac{5}{2}}$ block of the Jacobian, with the Fock operator explicitly connected to index a of the `bra` function.

order approaches highly undesirable.

So far, the process of finding all diagrams was automated. The underlying program determined all covariant diagrams for a given commutator of the Jacobian. Unfortunately, it can not handle the $m = \frac{5}{2}$ excitation space yet and it is necessary to find the diagrams by hand. For large excitation spaces, however, the number of diagrams is quite overwhelming. Figure 3.3 provides a small insight into this issue. The figure reveals all six covariant diagrams of the $\mathbf{A}_{\frac{5}{2}\frac{5}{2}}$ block of the Jacobian, where the Fock operator is explicitly connected to index a of the `bra` side. These six diagrams reflect all possible permutations of $\langle\mu^{ijk}_{ab}|$ with the `ket` side and are easily distinguishable by comparing the triples of trial vector amplitudes U. Of course there are four other possibilities for the Fock operator to connect to the `bra` side, namely indices b, i, j and k. Each connection contributes another six diagrams due to permutations with the `ket` side, resulting in a total of thirty diagrams for the $\mathbf{A}_{\frac{5}{2}\frac{5}{2}}$ block only, which is one of the easier blocks to evaluate. When projecting onto the contravariant

`bra` CSF instead, Eq. (3.6), it turns out that only two permutations with the `ket` side remain. For the example given above (Fock operator connected to index a), the diagrammatic representation of these two permutations is also shown in Figure 3.3, now drawn as contravariant diagrams as introduced in Section 1.3. Corresponding triples of the trial vector amplitudes reveal the two remaining permutations: ijk and jki, which are exactly the ones of the covariant `ket` CSFs chosen to be nonorthogonal to the contravariant CSF (cf. Eq. (3.5)). This is an important finding, because it is now clear which permutations remain for a certain fragment of the Hamiltonian, when projecting against the contravariant CSF, $\langle\tilde{\mu}_{ab}^{ijk}|$. The usefulness of contravariant CSFs becomes evident, too, since the number of diagrams is reduced considerably and thus the number of expressions contributing to the working equation. In case of the $\mathbf{A}_{\frac{5}{2}\frac{5}{2}}$ block of the Jacobian thirty covariant diagrams are replaced by ten contravariant ones. Furthermore, exchange diagrams corresponding to permutation jki are easily obtainable by just renaming the occupied indices of corresponding coulomb diagrams (those with permutation ijk) properly: $i \rightarrow j, j \rightarrow k$ and $k \rightarrow i$.

What remains is the task of finding all coulomb diagrams. Since coulomb diagrams correspond to permutation ijk, indices a and i share an electron, indices b and j another one and index k belongs to a third one. This applies to all vertices and, consequently, indices a and i form one loop, indices b and j a second loop and index k another one. Loops in contravariant diagrams are easier to detect by drawing them closed like the covariant ones (the purpose of drawing open contravariant diagrams is to be able to use the same diagrammatic rules for co- and contravariant CSFs). The contravariant coulomb diagram of Figure 3.3 (diagram 7) then looks like the first one in this figure. There, the three different loops are obvious. If the contravariant exchange diagram (diagram 8), on the other hand, is closed, it looks like the fourth diagram. Note that different contravariant CSFs yield different contravariant diagrams, depending on the selected conditions of Eq. (3.5).

The $\mathbf{A}_{\frac{1}{2}\frac{5}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{5}{2}}$ blocks of the Jacobian add the following terms to the full CCSD matrix trial-vector products given in Eqs. 2.29, and 2.30, 2.36, 2.39,

2.42, respectively:

$$^{\mathrm{T0}}\Delta\mathfrak{v}^{i} = (lc|md)\left(2u_{cd}^{lmi} - u_{cd}^{lim} - u_{cd}^{iml} + u_{dc}^{lmi} - \frac{1}{2}u_{dc}^{lim} - \frac{1}{2}u_{dc}^{iml}\right) \tag{3.8}$$

$$^{\mathrm{T0}}\Delta\mathfrak{v}_{a}^{ij} = (ac\hat{|}kb)\bar{u}_{cb}^{ikj} + S_{aa'}\left(\hat{f}_{kb}\bar{u}_{a'b}^{ikj} - (kb\hat{|}lj)\bar{u}_{a'b}^{ikl} - (kb\hat{|}li)\bar{u}_{a'b}^{lkj}\right), \tag{3.9}$$

with

$$\bar{u}_{ab}^{ijk} = 2u_{ab}^{ijk} - u_{ba}^{ijk} - u_{ab}^{ikj}. \tag{3.10}$$

Finally, the third row of the Jacobian entails a third matrix trial-vector product,

$$\mathfrak{v}_{ab}^{ijk} = Q_{ab}^{ijk} - Q_{ab}^{jki} + W_{ab}^{ijk} - W_{ab}^{jki}, \tag{3.11}$$

with the new intermediates

$$\begin{aligned}
Q_{ab}^{ijk} &= f_{ac}u_{cb'}^{ijk}S_{b'b} + S_{aa'}u_{ac}^{ijk}f_{bc} - S_{aa'}\left(u_{ab}^{ljk}f_{li} + u_{ab}^{ilk}f_{lj} + u_{ab}^{ijl}f_{lk}\right)S_{b'b}, \\
W_{ab}^{ijk} &= \Big\{-S_{aa'}\left(t_{a'c}^{ij}(bc\hat{|}lk) + t_{a'c}^{ik}(bj\hat{|}lc)\right) - \left((ac\hat{|}lk)t_{cb'}^{ij} + (ai\hat{|}lc)t_{cb'}^{kj}\right)S_{b'b} \\
&\quad + S_{aa'}\left(t_{a'b'}^{im}(mj\hat{|}lk) + t_{a'b'}^{mj}(mi\hat{|}lk)\right)S_{b'b}\Big\}u^{l} \\
&\quad - S_{aa'}\left(u_{a'}^{il}(bj\hat{|}lk) + u_{a'}^{lk}(bj\hat{|}li)\right) - \left((ai\hat{|}lk)u_{b'}^{jl} + (ai\hat{|}lj)u_{b'}^{lk}\right)S_{b'b} \\
&\quad + (ac\hat{|}bj)u_{c}^{ik} + (bc\hat{|}ai)u_{c}^{jk}.
\end{aligned} \tag{3.12}$$

A similar equation with similar intermediates was found for LCCSD(T) [62, Eq. (17)]. There, the equation

$$Q_{rst}^{ijk} - Q_{rst}^{jki} + W_{rst}^{ijk} - W_{rst}^{jki} = 0 \tag{3.13}$$

can be simplified to

$$Q_{rst}^{ijk} + W_{rst}^{ijk} = 0. \tag{3.14}$$

Further work is necessary to verify that this is also possible for the IPs.

So far, theories developed for the EE case were entirely adaptable to the IP case. However, IP-CC3$_{\mathrm{CC2}}$ cannot follow the entire scheme of Ref. [62] anymore. For standard excitation energies the indistinguishability of electrons

enables six equivalent expressions for the cluster operator:

$$\mathbf{T}_{abc}^{ijk} = \mathbf{T}_{acb}^{ikj} = \mathbf{T}_{bac}^{jik} = \mathbf{T}_{bca}^{jki} = \mathbf{T}_{cab}^{kij} = \mathbf{T}_{cba}^{kji}. \tag{3.15}$$

These six expressions cover all electron permutations, which enable the transformation of an arbitrary index order of indices i, j and k into index order ijk. This is used to write the working equations such that the contributions of triple pairs to the (T0) energy correction are calculated for every triple ijk separately. Such a scheme avoids the storage of all triples amplitudes, because after evaluating the amplitude for a certain triple ijk, the amplitude is immediately used to calculate the contribution of this triple to the (T0) energy correction and is discarded afterwards, since it is not needed anymore. Furthermore, the equations can be written such that a triangular triples list with $i \geq j \geq k$ can be used, which reduces the number of triples considerably.

The IP case, on the other hand, entails only the following identity in the cluster operator:

$$\mathbf{T}_{ab}^{ijk} = \mathbf{T}_{ba}^{jik}. \tag{3.16}$$

As a consequence, there is no convenient way to formulate the equations such that the same triple ijk is used for all intermediates and it becomes necessary to handle several triples amplitudes at the same time. In order to obtain an efficient program an alternative formulation is necessary.

Besides the triples amplitudes the second biggest quantity are the three-external integrals occurring in intermediate W_{ab}^{ijk} and in matrix trial-vector product ${}^{\mathrm{T0}}\Delta\mathfrak{v}_a^{ij}$. The LCCSD(T0) formalism, which treats the triples amplitudes pair by pair, requires reading of these integrals for every triples pair anew. Alternatively, one can drive the formalism by reading these integrals only once. This, however, implicates the storage of all triples amplitudes. Such an implementation was also realized by Schütz, however it is not published yet.

A similar approach is probably necessary for IP-CC3$_{\mathrm{CC2}}$ to obtain an efficient implementation, however, the storage of all amplitudes is very undesirable, due to enormous disk requirements. Therefore, a formalism is needed which

does not require all amplitudes to be stored, but at the same time minimizes the number of three-external integral readings. In the course of this work a minor part was already implemented, namely the contraction of the density fitted integral parts building the three-external integral $(ab|ci)$. In order to reduce the size of this integral, but also the size of the $t^{(1)}_{\mu_{\frac{5}{2}}}$ amplitudes, further approximations are necessary. A short overview is given in the next section.

3.3.3. Local approximations

By just looking at the sheer amount of triples amplitudes it gets immediately clear why (local) approximations are absolutely necessary when it comes to bigger systems: in cc-pVDZ basis the unrestricted $t^{(1)}_{\mu_{\frac{5}{2}}}$ amplitudes of the D21L6 molecule used earlier (c.f. Fig. 2.5) contain $2 \cdot 10^{12}$ entries, which require 15052 GB (14.7 TB) of memory. These numbers also emphasize the importance of avoiding the storage of all triples amplitudes. The fully unrestricted three-external integral, on the other hand, needs a little less memory: 832 GB.

Two restrictions are imposed on the three-external integral $(ab|ci)$. The first one is addressed to the distance between atoms c and LMOs i. Due to the exponential decay of the integral only those with atoms c close to LMOs i will have a non-negligible value. There is already a quantity which contains only close centers for any given LMO, i.e., the ground state orbital domains. These domains are imposed on index c. A restriction of indices a and b was introduced earlier on the three-index integral $(ab|P)$ (cf. Sec. 2.4.4) and is reused here. Center pairs A, B are neglected if their maximum overlap S_{AB} is below a given threshold. A test calculation with the same parameters used for D21L6 in Sec. 2.5 reveals a decrease of integral size by one order of magnitude to 58 GB.

Even more important are tools which reduce the size of the $t^{(1)}_{\mu_{\frac{5}{2}}}$ amplitudes. As already known from the $m = \frac{3}{2}$ excitation space, a pair list is applied to the occupied indices. Of course the $m = \frac{5}{2}$ excitation space requires a suitable triples pair list, which prevents certain triples ijk to be calculated. Due to the additional virtual index in $m = \frac{5}{2}$ quantities, a reduction of the virtual

space becomes important, too. To this, state-specific excited state domains are employed. In the course of this work, the module, which treats state-specific pair lists, was extended to be able to deal with such domains. The domain information is obtained in close analogy to the pair lists. The criterion is

$$d_A^{ij} = \sum_{a \in A, a'} r_a^{ij}(\bar{I}) S_{aa'} r_{a'}^{ij}(\bar{I}). \tag{3.17}$$

The upper equation does not employ the Einstein convention. From d_A^{ij} the important centers for pair ij are extracted, by (1) normalizing d_A^{ij} to unity ($\sum_A d_A^{ij} = 1$), (2) sorting these values according to decreasing magnitude, and (3) adding d_A^{ij} with largest values, until a given domain threshold is reached. Right now, ground state pair domains are added by default and augmented with important centers. If already the sum $\sum_A d_A^{ij}$ of centers in the ground state pair domain exceeds the domain threshold, excited and ground state pair domains coincide. Excited state domains which are smaller than ground state domains are not allowed.

The presented and implemented approximations are first steps towards the full program and only test calculations will reveal their suitability. From todays point of view these approximations seem appropriate, however, the method might be exceptionally prone to one or another such that adjustments may become necessary.

4. Local coupled cluster electron affinities

The original topic of this thesis, besides the implementation of ionization potentials, included also the implementation of electron affinities. Since the basic IP-LCC2 approach was not satisfactory all the effort went into optimizing the IP program. After recognizing that ionization potentials can also be used to calculate excitation energies of radicals, new interest in an improvement in the methodology awoke, since different types of excitations have different requirements for a sufficient accuracy. This left little to no time for programming EAs. However, working equations were already developed and are presented in the following.

4.1. Introduction

The theory for electron affinities is very similar to IPs and shares some advantages with it. For example, the ground state calculation remains unaffected by the choice of the excited state and already implemented ground state methods can be used. Furthermore, $m = \frac{1}{2} \oplus \frac{3}{2}$ excited state quantities of EAs and IPs have only one or three indices, compared to the electronically excited state with two and four indices (up to CCSD). However, EAs are computationally more expensive than IPs, because the excess index, which only destroys or only creates a particle, is an occupied index for IPs but virtual for EAs. Finally, from a programmers perspective the implementation of EAs is fairly easy, because the program structure, which can handle the $m = \frac{1}{2} \oplus \frac{3}{2}$ excitation space,

is already implemented for the IP case. Of course this structure has to be adjusted to the EA case, but it is now easier to identify where adjustments are necessary.

In section 2.2 the derivation of IP-CCSD$[k]_M$ methods started by giving the linear response function, Eq. (2.3). A formal operator $\mathbf{Y}$ was introduced, which either destroys a particle or creates one. Thus, the final linear response function contains poles for ionization energies as well as electron affinities. In this chapter the second part of $\langle\langle X; Y \rangle\rangle$ is discussed in a nutshell, containing the poles $\omega_{\bar{A}}$.

4.2. Theory

Most of the equations developed for the IP case are also applicable to EAs. Adjustments are solely necessary for equations with explicit reference to orbitals, which are outlined in the following.

As mentioned in the introduction, excited state quantities of EAs have one excess index creating a new particle in the virtual space without a corresponding annihilation of a particle. Therefore, the excited state cluster operators given in Eq. (2.6) are substituted by their EA counterparts:

$$\begin{aligned} \mathbf{T}^{(1)}_{\frac{1}{2}}(t) &= t^{(1)}_{\mu_{\frac{1}{2}}}(t)\tau_{\mu_{\frac{1}{2}}} = t_a(t)\tau^a, \\ \mathbf{T}^{(1)}_{\frac{3}{2}}(t) &= t^{(1)}_{\mu_{\frac{3}{2}}}(t)\tau_{\mu_{\frac{3}{2}}} = t^i_{ab}(t)\tau^{ab}_i, \end{aligned} \tag{4.1}$$

with the new spin-adapted operators

$$\tau^a = a^\dagger_{a\alpha}, \qquad \tau^{ab}_i = \tau^a_i \tau^b. \tag{4.2}$$

The contravariant `bra` CSFs are also virtually identical to those given in

Eq. (2.9), again, only the orbital indices differ:

$$\begin{aligned}\langle\tilde{\mu}_{\frac{1}{2}}| &= \langle\tilde{\Phi}^a| = \langle 0|(\tau^a)^\dagger,\\ \langle\tilde{\mu}_{\frac{3}{2}}| &= \langle\tilde{\Phi}_i^{ab}| = \frac{1}{3}\langle 0|(2\tau_i^{ab} + \tau_i^{ba})^\dagger.\end{aligned} \tag{4.3}$$

Furthermore, the Jacobian with its general expressions given in Eq. (2.27) is the same for electron affinities. Even the diagrams are similar to IP diagrams and it is enough to reverse the direction of the arrows in Figures 2.1 and B.1, B.2 to obtain EA-CC2 diagrams and the diagrams of the first extension EA-CCSD[1]$_M$, respectively. The final EA-CC2 working equations for the right matrix trial-vector product then reads:

$$\mathfrak{v}_a = \hat{f}_{ab}u_b - S_{aa'}Z_{a'c}u_c + f_b^j\tilde{u}_{ba'}^j S_{a'a} + (ac\hat{|}P)^{\mathrm{fc}}W_c^P \tag{4.4}$$

$$\mathfrak{v}_{ab}^i = c_P^{ai}\hat{B}_b^P + f_{ac}u_{cb'}^i S_{b'b} + f_{bc}u_{a'c}^i S_{a'a} - f_{ji}u_{a'b'}^j S_{a'a}S_{b'b}, \tag{4.5}$$

with the new intermediates

$${}^{\mathrm{fc}}W_a^P = c_{jc}^P\tilde{u}_{ca}^j, \qquad \hat{B}_a^P = (ac\hat{|}P)u_c. \tag{4.6}$$

The Z_{ac} intermediate was already introduced in Eq. (2.40) for IP-CCSD[2]$_M$, and, therefore, requires no extra coding.

It is very likely that CC2 electron affinities are as inaccurate as CC2 IPs. Orbital relaxation effects play probably an even bigger role when adding an extra electron to a molecule. Consequently, higher order methods are necessary for a better prediction of EAs. Following the same scheme used for IPs, higher-order diagrams are added to the CC2 Jacobian, while still sticking to the $m = \frac{1}{2} \oplus \frac{3}{2}$ excitation space. Again, the $\mathbf{A}_{\frac{1}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{1}{2}\frac{3}{2}}$ submatrices contain already all possible diagrams. Furthermore, the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ and $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ submatrices are only correct to first- and zeroth-order with respect to the fluctuation potential, respectively. Increasing the order of these two submatrices by one is a first approach to the full CCSD Jacobian. The corresponding acronym is EA-CCSD[1]$_M$, allowing for a flexible ground state calculation, M. Since the

undressed Fock operator is replaced by the dressed one in $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$, the undressed Fock matrix elements in Eq. (4.5) have to be substituted by their dressed counterparts. The new higher-order terms add the following expressions to the EA-CCSD[0]$_M$ matrix trial-vector product given in Eq. (4.5):

$$\begin{aligned}
{}^{[1]}\Delta\mathfrak{v}^i_{ab} = &-S_{aa'}t^{ij}_{a'c}\hat{Y}^j_{bc} - S_{bb'}t^{ij}_{cb'}\hat{Y}^j_{ac} + S_{aa'}V^P_{ia'}\hat{B}^P_b + S_{aa'}S_{bb'}t^{kj}_{ab}\hat{Y}^{ijk} \\
&+ (ai\hat{|}P)^{\mathrm{fc}}W^P_{b'}S_{b'b} + (bc\hat{|}P)\hat{c}^P_{ad}u^i_{dc} \\
&- (ac\hat{|}P)\hat{c}^P_{ji}u^j_{cb'}S_{b'b} - (bc\hat{|}P)\hat{c}^P_{ji}u^j_{a'c}S_{a'a},
\end{aligned} \tag{4.7}$$

with the new intermediates

$$\begin{aligned}
\hat{Y}^i_{ab} &= (ab\hat{|}P)^{\mathrm{fc}}B^P_i, & \hat{Y}^{ijk} &= {}^{\mathrm{int}}B^P_j\hat{c}^P_{ki}, \\
{}^{\mathrm{fc}}B^P_i &= c^P_{ic}u_c, & {}^{\mathrm{int}}B^P_i &= (ic|P)u_c.
\end{aligned} \tag{4.8}$$

The first line of Eq. (4.7) summarizes extra terms of the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ block, the remaining ones belong to the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ block of the Jacobian. The new intermediate $\hat{Y}^{ijk}$ is much easier than its three-external IP counterpart, $\hat{Y}_{abc}$ (cf. Eq. (2.37)). However, this slight advantage is immediately negated by one term from the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ block, namely the second one containing the four-external integral, split into two three-index integrals by density fitting. Two disadvantages come along with this expression: 1) the two-external fit coefficient has to be created, which could be avoided so far, and 2) the contractions between the matrix elements are extremely inconvenient. Computationally it is probably still best to build the four-external integral $(ab\hat{|}cd)$ outside the Davidson diagonalization.

Just like IP-CCSD[1]$_M$, EA-CCSD[1]$_M$ diagrams scale nominally as $\mathcal{O}(\mathcal{N}^5)$, and, therefore, the scaling is one magnitude higher than EA-CCSD[0]$_M$. Local correlation techniques should be introduced to counter the higher scaling and reduce the computational cost. Since the three-index objects have one occupied and two external indices, excited state pair lists make no sense. Instead, state-specific excited state orbital domains can be used to restrict the virtual space. A first approach could be a Löwdin alike analysis of $r^i_{ab}(\bar{I})$, which is already

used for LT-DF-LCC2 [19]:

$$d_A^i(\bar{I}) = \sum_{a\in[A],bcc'd} S_{ab}^{\frac{1}{2}} r_{bc}^i(\bar{I}) S_{cc'} r_{dc'}^i(\bar{I}) S_{da}^{\frac{1}{2}}. \tag{4.9}$$

The upper summation does not employ the Einstein convention. Such an analysis requires a preceding EA-CCSD$[0]_M$ calculation which delivers the state-specific eigenvectors $r_{ab}^i(\bar{I})$ to be analyzed. From d_A^i state-specific orbital domains can be determined by first normalizing d_A^i to unity and adding centers A afterwards, until the sum $\sum_A d_A^i$ reaches a certain threshold. Test calculations will reveal if such domains are sufficient enough for a proper description of electron attached states. Another alternative is to start from ground state orbital domains and augment these with centers A, until the sum $\sum_A d_A^i$ reaches a certain threshold. Or it might even be necessary to have two different sets of orbital domains, one set for the electron which is excited within the molecule, the other set for the electron which is added to the molecule. This can only be answered once the program is available.

5. Summary

The goal of this work was to implement an ab initio method for the calculation of ionization potentials with good accuracy and cost effectiveness even for big-seized molecules. This was achieved within the linear response coupled cluster theory. As has already been pointed out, the coupled cluster ansatz is very versatile, so far defining the gold standard, and has already been proven successful studying (particle conserving) excited state properties of extended molecular systems. The disadvantages coming along with an ab initio method, mainly high computational cost, are reduced by proper approximations. Particularly noteworthy are the local ansatz and corresponding approximations, as well as density fitting. The former reduces the overall scaling of a method, the latter the prefactor.

It was found that the accuracy provided by a pure IP-CC2 ansatz is not sufficient. Improvements were realized by augmenting the CC2 Jacobian with higher-order contributions, but keeping the ground state on the level of MP2 or CC2. Already the first extension ($k = 1$) has a huge impact on the resulting ionization potentials, changing them in all cases from insufficient to an acceptable accuracy. Further enhancements ($k = 2, \mathrm{f}$) not only lead to an increased accuracy, but also to faster convergence. This behaviour makes IP-CCSD$_{\mathrm{CC2}}$ the method of choice.

Furthermore, it was illustrated that it is naturally possible to calculate excitation energies of radicals from ionization potentials. Different radical excitation processes (Type I,II,III), however, place different requirements on the IP calculations. The calculation of Type II and III excitations energies with ionization potentials requires a more accurate description of $m = \frac{3}{2}$ ionization processes than offered by the IP-CCSD$[k]_M$ methods. For this purpose, the

concept of adding higher-order terms to the Jacobian is continued, which is accompanied by a larger excitation space. Preliminary work for an extension to the $m = \frac{1}{2} \oplus \frac{3}{2} \oplus \frac{5}{2}$ excitation space was presented and first implementations were already realized.

Finally, first steps towards a linear response coupled cluster method were illustrated for the calculation of electron affinities. Similarities to the IP case eased the process of developing the underlying theory. Working equations are presented for EA-CC2 as well as the next higher-order method EA-CCSD$[1]_M$. In analogy to the IP case, the additional higher-order terms of EA-CCSD$[1]_M$ increase the computational complexity. The dissertation concludes with a suggestion on possible local approximations for the virtual space in order to counter the increased computational cost.

A. Supplementary information for test calculations

This section summarizes ionization potentials of several test calculations. Some of these data are used for the tables and figures in the main text.

Tables A.1 and A.2 contain canonical and local ionization potentials, the latter with corresponding local errors. Different thresholds for the excited state pair lists were tested and the results are listed in Table A.3.

The influence of different ij- and AB-thresholds is summarized in Tables A.4-A.8. The AB-threshold effects the $(ab|\hat{P})$ integral, which, in turn, is contracted on the one hand with the fit coefficients $\hat{c}_P^{ij}$, forming the intermediate *oovv*, and on the other hand for building the intermediate Y_{abc}. The effect of the threshold on these two quantities is examined separately on two bigger sized systems, namely a chain of six glycine units [64] and the *d* 10-arylisoalloxazine derivative [65, Fig. 2]. It is not recommended to use stricter AB-thresholds than 10^{-2} for both intermediates, although the Y_{abc} intermediate seems more robust to harsher criteria. Since the same threshold is suggested for both intermediates, the program was changed to not distinguish between two different AB-thresholds anymore.

Finally, Table A.9 gives more details regarding calculations on the D21L6 molecule than the one included in the main text (Tab. 2.2).

Table A.1.: Ionization potentials (in [eV]) for IP-CCSD$[k]_M$, HF, EOMIP-CCSD and Δ-methods. No local approximations are invoked in these calculations, thus, ground state domains and pair lists P_0 and $P_{\bar{I}}$ remain full for ADC(2) and all IP-CCSD$[k]_M$ methods. The AB- and ij-thresholds are set to 10^{-8}.

			IP-CCSD$[k]_M$										
			$k=0$		$k=1$		$k=2$		$k=f$				
Molecule	State	HF	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2	EOMIP-CCSD	ΔMP2	ΔCCSD	ΔCCSD(T)
Water	D_0	13.320	10.826	10.847	11.574	11.614	11.611	11.652	11.651	11.666	11.996	11.806	11.830
	D_1	14.991	12.922	12.947	13.576	13.617	13.611	13.652	13.652	13.689	13.966	13.817	13.842
	D_2	19.053	17.941	17.966	18.378	18.416	18.422	18.461	18.460	18.524	18.735	18.587	18.602
Furan	D_0	8.574	8.442	8.390	8.626	8.569	8.778	8.728	8.730	8.722	8.751	8.679	8.703
	D_1	10.729	9.523	9.549	9.982	10.000	10.148	10.175	10.174	10.076	10.098	10.177	10.127
	D_2	14.585	11.695	11.715	12.787	12.784	12.921	12.994	12.991	12.885	13.603	13.070	13.028
	D_3	15.301	12.867	12.779	12.850	12.859[a]	13.654	13.606	13.607	13.600			
	D_4	15.611	13.633	13.520	13.507	13.461	14.225	14.111	14.113	14.176	13.981	14.343	14.142
	D_5	16.509	13.754	13.761	14.053[b]	13.940[b]	14.717	14.806	14.804	14.685			
Pyridin	D_0	9.328[c]	8.346	8.304	9.278	9.265	9.481	9.466	9.467	9.393	9.573[d]	9.404	9.396
	D_1	10.272[c]	9.167	9.134	9.397	9.368	9.569	9.544	9.545	9.459	9.687[d]	9.468	9.502
	D_2	11.314[c]	9.774	9.756	10.064	10.055	10.273	10.268	10.269	10.148	10.259	10.198	10.200
	D_3	14.000	11.923	11.827	12.536	12.447	12.701	12.610	12.612	12.625			
Phenol	D_0	8.495	7.836	7.798	8.223	8.187	8.392	8.359	8.360	8.258	8.386	8.341	8.327
	D_1	9.225	8.813	8.752	9.062	9.001	9.259	9.201	9.203	9.135			
	D_2	12.980	10.848	10.799	11.181	11.136	11.657	11.636	11.636	11.603			
	D_3	13.372	11.098	11.004	11.806	11.710	11.968	11.870	11.873	11.880	11.336	12.164	11.859
Adenin	D_0	8.223	7.607	7.555	8.006	7.965	8.205	8.169	8.170	7.995	8.339	8.093	8.046
	D_1	9.959	7.965	7.898	8.933	8.919	9.142	9.129	9.130	8.997			
	D_2	10.991	8.821	8.780	9.209	9.175	9.461	9.438	9.440	9.229			
	D_3	11.388	9.000	8.934	9.958	9.941	10.188	10.174	10.174				
Thymin	D_0	9.459	8.392	8.271	8.799	8.671	8.980	8.852	8.854	8.781	8.867	8.765	8.717
	D_1	11.570	8.476	8.449	9.448	9.562	9.729	9.857	9.849	9.688			
	D_2	11.962	9.458	9.408	9.924	9.899	10.214	10.207	10.206	10.054			
	D_3	12.814	9.461	9.453	10.322	10.449	10.620	10.762	10.753				

Table continues on next page

Table A.1.: Continued from previous page

			IP-CCSD[k]$_M$										
			$k=0$		$k=1$		$k=2$		$k=f$				
Molecule	State	HF	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2	EOMIP-CCSD	ΔMP2	ΔCCSD	ΔCCSD(T)
Uracil	D_0	9.870	8.559	8.538	9.171	9.045	9.366	9.242	9.242	9.171	9.273	9.155	9.097
	D_1	11.622	8.802[e]	8.681[e]	9.508	9.632	9.789	9.927	9.919	9.762			
	D_2	12.000	9.559	9.508	10.025	10.004	10.301	10.297	10.296	10.145			
	D_3	12.934	9.594	9.589	10.470	10.599	10.756	10.900	10.891				
DMABN	D_0	8.016	6.851	6.793	7.390	7.334	7.584	7.532	7.533	7.373	7.474	7.462	7.414
	D_1	9.753	9.058	8.951	9.320	9.207	9.598	9.492	9.494	9.407			
	D_2	11.211	9.548	9.494	9.828	9.778	10.266	10.244	10.244	10.136			
	D_3	12.120	10.965	10.841	11.160	11.073	11.629	11.546	11.547	11.476			
	D_4	12.977[f]	11.068	10.959	11.478	11.389	12.120	12.073	12.074	12.026	12.014	11.721	11.687
	D_5	13.698	11.600	11.515	11.694[b]	11.635[b]	12.183	12.133	12.133	12.045			
HPA	D_0	8.381	7.640	7.576	8.053	7.992	8.229	8.169	8.170	8.048	8.136	8.098	8.081
	D_1	9.363	8.828	8.750	9.076	8.996	9.296	9.221	9.223	9.149			
	D_2	12.191	9.216	9.246	10.034	10.149	10.232	10.352	10.348	10.232			
	D_3	12.863	10.612	10.544	11.030	10.966	11.464	11.418	11.418	11.519			
P-cresol	D_0	8.191	7.523	7.484	7.920	7.885	8.086	8.053	8.054	7.942	8.069	8.020	8.012
	D_1	9.206	8.703	8.644	8.954	8.896	9.165	9.111	9.112	9.040			
	D_2	12.683	10.519	10.477	10.888	10.850	11.338	11.322	11.322	11.294			
	D_3	13.138	10.854	10.758	11.560	11.464	11.724	11.626	11.628	11.633	11.098	11.911	11.606
N-acetylglycine	D_0	10.748[d]	8.258	8.264	9.153	9.235[d]	9.359	9.411[d]	9.411[d]	9.343[d]	9.561	9.253	9.267
	D_1	11.470	8.777	8.729	9.266	9.269	9.438	9.482	9.475	9.357	9.614	9.375	9.338
	D_2	12.836	9.869	9.909	10.730	10.864	10.906	11.043	11.039	10.912			
	D_3	13.448	11.364	11.389	11.902	11.980	12.103	12.193	12.190	12.052			
1-phenylpyrrole	D_0	8.144	7.449	7.442	7.812	7.755	8.067	8.021	8.022	7.943	7.986	7.861	7.877
	D_1	8.343	7.599[e]	7.536[e]	7.908	7.904	8.118	8.119	8.120	7.943			
	D_2	9.410	8.911	8.886	9.155	9.139	9.393	9.384	9.385	9.289			
	D_3	10.256	9.265	9.227	9.588	9.561	9.863	9.845	9.846				

Table continues on next page

Table A.1.: Continued from previous page

			IP-CCSD[k]$_M$										
			$k=0$		$k=1$		$k=2$		$k=f$				
Molecule	State	HF	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2	EOMIP-CCSD	ΔMP2	ΔCCSD	ΔCCSD(T)
Propanamide	D_0	11.004[d]	8.180	8.232	9.089	9.254	9.271	9.440	9.433	9.310	9.687	9.288	9.319
	D_1	11.271[d]	9.175	9.183	9.616	9.652	9.787	9.829	9.828	9.752	10.062	9.833	9.804
	D_2	13.298	11.888	11.840	12.072	12.022	12.261	12.210	12.209	12.311			
	D_3	13.940	12.205	12.155	12.650	12.604	12.780	12.733	12.733	12.828			
Tyrosin	D_0	8.346	7.548	7.494	7.982	7.934	8.161	8.114	8.115	7.986	8.102	8.054	8.035
	D_1	9.265	8.497	8.502	8.966	8.904	9.189	9.132	9.133	9.052			
	D_2	11.372	8.712	8.646	9.303	9.365	9.500	9.562	9.559				
	D_3	12.251	9.869	9.794	10.419	10.367	10.588	10.542	10.541				
trans-urocanic acid	D_0	8.346	7.911	7.820	8.152	8.060	8.377	8.287	8.289	8.211	8.398	8.195	8.176
	D_1	10.449	8.611	8.673	9.374	9.533	9.715	9.909	9.903	9.766	10.255	9.985	9.995
	D_2	11.834[g]	9.160	9.084	9.615	9.548[g]	10.027	9.968	9.969	9.847			
	D_3	11.981	9.460	9.384	10.078	10.023	10.292	10.242	10.243	10.147			

[a] D_3 of the method corresponds to D_2 of ΔCCSD(T)

[b] D_5 of the method corresponds to D_4 of ΔCCSD(T)

[c] D_0, D_1, D_2 of the method correspond to D_1,D_2,D_0 of ΔCCSD(T), respectively

[d] D_1 of the method corresponds to D_0 of ΔCCSD(T) and vice versa

[e] D_1 of the method corresponds to D_0 of ΔCCSD(T)

[f] D_6 of HF is the corresponding state to D_4 of ΔCCSD(T). Its energy is 14.463 eV.

[g] D_2 of the method corresponds to D_1 of ΔCCSD(T)

Table A.2.: Ionization potentials and related local errors (in parenthesis) in [eV] for the IP-CCSD$[k]_M$ methods. Local approximations are invoked in these calculations: the ground state pair list P_0 is restricted according to a distance criterion of $R_0 = 10$ bohr. Orbital domains are build according to the Boughton-Pulay procedure with a criterion of BP=0.98, and extended by the next nearest neighbours with `iext=1`. Furthermore, for $k \geq 1$, $\kappa_e = 0.99$ restricts the excited state pair lists $P_{\bar{I}}$. The AB- and ij-thresholds are 10^{-2} and 10^{-4}, respectively.

		IP-CCSD$[k]_M$						
		$k=0$		$k=1$		$k=2$		$k=f$
Molecule	State	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2
Water	D_0	10.826 (0.000)	10.847 (0.000)	11.592 (0.018)	11.632 (0.018)	11.628 (0.017)	11.669 (0.017)	11.668 (0.017)
	D_1	12.922 (0.000)	12.947 (0.000)	13.581 (0.005)	13.622 (0.005)	13.616 (0.005)	13.657 (0.005)	13.656 (0.004)
	D_2	17.941 (0.000)	17.966 (0.000)	18.390 (0.012)	18.428 (0.012)	18.433 (0.011)	18.472 (0.011)	18.471 (0.011)
Furan	D_0	8.442 (0.000)	8.389 (0.001)	8.635 (0.009)	8.581 (0.012)	8.787 (0.009)	8.736 (0.008)	8.738 (0.008)
	D_1	9.523 (0.000)	9.549 (0.000)	9.996 (0.014)	10.010 (0.010)	10.159 (0.011)	10.182 (0.007)	10.181 (0.007)
	D_2	11.692 (0.003)	11.713 (0.002)	12.802 (0.015)	12.850 (0.066)[a]	12.925 (0.004)	12.999 (0.005)	12.996 (0.005)
	D_3	12.866 (0.001)	12.778 (0.001)	12.921 (0.071)[a]	12.874 (0.015)	13.657 (0.003)	13.607 (0.001)	13.608 (0.001)
	D_4	13.632 (0.001)	13.518 (0.002)	13.520 (0.013)	13.470 (0.009)	14.228 (0.003)	14.114 (0.003)	14.117 (0.004)
	D_5	13.754 (0.000)	13.760 (0.001)	14.065 (0.012)	13.952 (0.012)	14.730 (0.013)	14.813 (0.007)	14.811 (0.007)
Pyridin	D_0	8.343 (0.003)	8.300 (0.004)	9.280 (0.002)	9.265 (0.000)	9.474 (0.007)	9.457 (0.009)	9.458 (0.009)
	D_1	9.167 (0.000)	9.134 (0.000)	9.405 (0.008)	9.378 (0.010)	9.576 (0.007)	9.551 (0.007)	9.553 (0.008)
	D_2	9.774 (0.000)	9.756 (0.000)	10.077 (0.013)	10.068 (0.013)	10.283 (0.010)	10.277 (0.009)	10.278 (0.009)
	D_3	11.921 (0.002)	11.824 (0.003)	12.551 (0.015)	12.461 (0.014)	12.704 (0.003)	12.612 (0.002)	12.614 (0.002)
Phenol	D_0	7.834 (0.002)	7.795 (0.003)	8.234 (0.011)	8.198 (0.011)	8.401 (0.009)	8.367 (0.008)	8.368 (0.008)
	D_1	8.813 (0.000)	8.752 (0.000)	9.071 (0.009)	9.009 (0.008)	9.265 (0.006)	9.207 (0.006)	9.208 (0.005)
	D_2	10.846 (0.002)	10.797 (0.002)	11.199 (0.018)	11.153 (0.017)	11.667 (0.010)	11.646 (0.010)	11.646 (0.010)
	D_3	11.095 (0.003)	11.000 (0.004)	11.820 (0.014)	11.725 (0.015)	11.971 (0.003)	11.873 (0.003)	11.876 (0.003)

Table continues on next page

Table A.2.: Continued from previous page

		IP-CCSD$[k]_M$						
		$k=0$		$k=1$		$k=2$		$k=f$
Molecule	State	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2
Adenine	D_0	7.605 (0.002)	7.553 (0.002)	8.014 (0.008)	7.973 (0.008)	8.209 (0.004)	8.171 (0.002)	8.173 (0.003)
	D_1	7.958 (0.007)	7.891 (0.007)	8.935 (0.002)	8.918 (0.001)	9.131 (0.011)	9.116 (0.013)	9.116 (0.014)
	D_2	8.820 (0.001)	8.778 (0.002)	9.222 (0.013)	9.186 (0.011)	9.465 (0.004)	9.442 (0.004)	9.443 (0.003)
	D_3	8.994 (0.006)	8.928 (0.006)	9.962 (0.004)	9.944 (0.003)	10.179 (0.009)	10.162 (0.012)	10.162 (0.012)
Thymine	D_0	8.390 (0.002)	8.269 (0.002)	8.807 (0.008)	8.679 (0.008)	8.981 (0.001)	8.853 (0.001)	8.855 (0.001)
	D_1	8.473 (0.003)	8.445 (0.004)	9.459 (0.011)	9.574 (0.012)	9.723 (0.006)	9.848 (0.009)	9.841 (0.008)
	D_2	9.456 (0.002)	9.405 (0.003)	9.938 (0.014)	9.919 (0.020)	10.217 (0.003)	10.210 (0.003)	10.209 (0.003)
	D_3	9.460 (0.001)	9.451 (0.002)	10.329 (0.007)	10.458 (0.009)	10.611 (0.009)	10.755 (0.007)	10.746 (0.007)
Uracil	D_0	8.557 (0.002)	8.535 (0.003)	9.183 (0.012)	9.059 (0.014)	9.372 (0.006)	9.248 (0.006)	9.249 (0.007)
	D_1	8.801 (0.001)	8.680 (0.001)	9.518 (0.010)	9.644 (0.012)	9.784 (0.005)	9.919 (0.008)	9.911 (0.008)
	D_2	9.557 (0.002)	9.506 (0.002)	10.044 (0.019)	10.024 (0.020)	10.307 (0.006)	10.302 (0.005)	10.301 (0.005)
	D_3	9.592 (0.002)	9.587 (0.002)	10.479 (0.009)	10.609 (0.010)	10.753 (0.003)	10.895 (0.005)	10.886 (0.005)
DMABN	D_0	6.845 (0.006)	6.784 (0.009)	7.395 (0.005)	7.336 (0.002)	7.579 (0.005)	7.523 (0.009)	7.524 (0.009)
	D_1	9.056 (0.002)	8.949 (0.002)	9.329 (0.009)	9.216 (0.009)	9.599 (0.001)	9.493 (0.001)	9.494 (0.000)
	D_2	9.547 (0.001)	9.494 (0.000)	9.853 (0.025)[a]	9.806 (0.028)[a]	10.270 (0.004)	10.246 (0.002)	10.247 (0.003)
	D_3	10.964 (0.001)	10.842 (0.001)	11.191 (0.031)	11.107 (0.034)	11.633 (0.004)	11.550 (0.004)	11.551 (0.004)
	D_4	11.063 (0.005)	10.955 (0.004)	11.515 (0.037)[a]	11.431 (0.042)[a]	12.111 (0.009)	12.062 (0.011)	12.063 (0.011)
	D_5	11.599 (0.001)	11.516 (0.001)	11.709 (0.015)	11.651 (0.016)	12.190 (0.007)	12.139 (0.006)	12.139 (0.006)

Table continues on next page

Table A.2.: Continued from previous page

		IP-CCSD$[k]_M$						
		$k=0$		$k=1$		$k=2$		$k=f$
Molecule	State	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2
HPA	D_0	7.632 (0.008)	7.566 (0.010)	8.052 (0.001)	7.988 (0.004)	8.221 (0.008)	8.158 (0.011)	8.159 (0.011)
	D_1	8.827 (0.001)	8.749 (0.001)	9.082 (0.006)	9.003 (0.007)	9.296 (0.000)	9.223 (0.002)	9.224 (0.001)
	D_2	9.212 (0.004)	9.240 (0.006)	10.038 (0.004)	10.150 (0.001)	10.225 (0.007)	10.342 (0.010)	10.338 (0.010)
	D_3	10.612 (0.000)	10.545 (0.001)	11.041 (0.011)	10.975 (0.009)	11.464 (0.000)	11.416 (0.002)	11.416 (0.002)
p-cresol	D_0	7.516 (0.007)	7.475 (0.009)	7.925 (0.005)	7.888 (0.003)	8.086 (0.000)	8.052 (0.001)	8.052 (0.002)
	D_1	8.702 (0.001)	8.643 (0.001)	8.962 (0.008)	8.904 (0.008)	9.169 (0.004)	9.116 (0.005)	9.117 (0.005)
	D_2	10.521 (0.002)	10.478 (0.001)	10.903 (0.015)	10.865 (0.015)	11.347 (0.009)	11.330 (0.008)	11.330 (0.008)
	D_3	10.849 (0.005)	10.754 (0.004)	11.573 (0.013)	11.477 (0.013)	11.724 (0.000)	11.625 (0.001)	11.627 (0.001)
N-acetylglycine	D_0	8.256 (0.002)	8.261 (0.003)	9.162 (0.009)	9.240 (0.005)	9.356 (0.003)	9.411 (0.000)	9.411 (0.000)
	D_1	8.775 (0.002)	8.727 (0.002)	9.272 (0.006)	9.275 (0.006)	9.438 (0.000)	9.475 (0.007)	9.468 (0.007)
	D_2	9.866 (0.003)	9.905 (0.004)	10.735 (0.005)	10.865 (0.001)	10.900 (0.006)	11.035 (0.008)	11.030 (0.009)
	D_3	11.361 (0.003)	11.386 (0.003)	11.923 (0.021)	12.003 (0.023)	12.110 (0.007)	12.201 (0.008)	12.198 (0.008)
1-phenylpyrrole	D_0	7.447 (0.002)	7.440 (0.002)	7.823 (0.011)	7.765 (0.010)	8.070 (0.003)	8.023 (0.002)	8.024 (0.002)
	D_1	7.597 (0.002)	7.534 (0.002)	7.915 (0.007)	7.910 (0.006)	8.118 (0.000)	8.118 (0.001)	8.119 (0.001)
	D_2	8.909 (0.002)	8.884 (0.002)	9.162 (0.007)	9.146 (0.007)	9.395 (0.002)	9.385 (0.001)	9.386 (0.001)
	D_3	9.262 (0.003)	9.224 (0.003)	9.595 (0.007)	9.568 (0.007)	9.862 (0.001)	9.842 (0.003)	9.843 (0.003)
Propanamide	D_0	8.177 (0.003)	8.228 (0.004)	9.095 (0.006)	9.259 (0.005)	9.269 (0.002)	9.435 (0.005)	9.429 (0.004)
	D_1	9.173 (0.002)	9.181 (0.002)	9.628 (0.012)	9.664 (0.012)	9.795 (0.008)	9.835 (0.006)	9.834 (0.006)
	D_2	11.885 (0.003)	11.837 (0.003)	12.085 (0.013)	12.034 (0.012)	12.266 (0.005)	12.215 (0.005)	12.215 (0.006)
	D_3	12.202 (0.003)	12.151 (0.004)	12.657 (0.007)	12.612 (0.008)	12.779 (0.001)	12.731 (0.002)	12.731 (0.002)

Table continues on next page

Table A.2.: Continued from previous page

		IP-CCSD$[k]_M$						
		$k=0$		$k=1$		$k=2$		$k=f$
Molecule	State	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2
Tyrosine	D_0	7.547 (0.001)	7.492 (0.002)	7.985 (0.003)	7.935 (0.001)	8.157 (0.004)	8.109 (0.005)	8.109 (0.006)
	D_1	8.492 (0.005)	8.496 (0.006)	8.970 (0.004)	8.906 (0.002)	9.189 (0.000)	9.130 (0.002)	9.131 (0.002)
	D_2	8.713 (0.001)	8.646 (0.000)	9.303 (0.000)	9.364 (0.001)	9.490 (0.010)	9.550 (0.012)	9.547 (0.012)
	D_3	9.859 (0.010)	9.782 (0.012)	10.413 (0.006)	10.358 (0.009)	10.575 (0.013)	10.524 (0.018)	10.524 (0.017)
trans-urocanic acid	D_0	7.908 (0.003)	7.816 (0.004)	8.162 (0.010)	8.070 (0.010)	8.380 (0.003)	8.289 (0.002)	8.290 (0.001)
	D_1	8.611 (0.000)	8.674 (0.001)	9.389 (0.015)	9.558 (0.025)[a]	9.712 (0.003)	9.905 (0.004)	9.899 (0.004)
	D_2	9.153 (0.007)	9.074 (0.010)	9.641 (0.026)[a]	9.562 (0.014)	10.035 (0.008)	9.975 (0.007)	9.975 (0.006)
	D_3	9.459 (0.001)	9.383 (0.001)	10.081 (0.003)	10.024 (0.001)	10.282 (0.010)	10.229 (0.013)	10.230 (0.013)

[a] State with substantial m= $\frac{3}{2}$ character (at least a weight of 10%)

Table A.3.: IP-CCSD$_{CC2}$ ionization potentials (in [eV]) for tyrosine with local approximations for the ground state (P_0: $R_0 = 10$ bohr, orbital domains: BP$= 0.98$, in one case extended with `iext=1`). The AB- and ij-threshold are set to 10^{-2} and 10^{-4}, respectively. Different κ_e values are tested. *Can* gives canonical IPs (un-restricted pair lists P_0 and $P_{\bar{I}}$, AB- and ij-threshold set to 10^{-8}).

	Can	0.999	0.995	0.99	0.99 `iext=1`	0.95	0.9
D_0	8.115	8.038	8.039	8.039	8.109	8.039	8.034
D_1	9.133	9.076	9.076	9.077	9.131	9.083	9.089
D_2	9.559	9.493	9.493	9.493	9.547	9.472	9.438
D_3	10.541	10.448	10.448	10.448	10.524	10.436	10.413
D_4	10.986	10.936	10.936	10.935	10.976	10.931	10.932
D_5	11.410	11.343	11.343	11.342	11.401	11.336	11.295
D_6	11.682	11.603	11.602	11.600	11.668	11.584	11.503

Table A.4.: IP-CCSD$_{CC2}$ ionization potentials (in [eV]) for tyrosine with local approximations for the ground state (P_0: $R_0 = 10$ bohr, orbital domains: BP$= 0.98$, `iext=1`) and un-truncated pair lists $P_{\bar{I}}$. The AB-threshold is set to zero. Different ij-thresholds are tested. *Scr. Int Len* is the number of entries for integral part $(ij\hat{|}P)$ after screening with ij-threshold, while *Unscr. Int Len* is the number of entries for $(ij\hat{|}P)$ without screening according to the ij-threshold. *% Unscr. Int* is the ratio *Scr. Int Len* to *Unscr. Int Len*.

	0.0	10^{-6}	$5*10^{-6}$	10^{-5}	$5*10^{-5}$	10^{-4}	$5*10^{-4}$	10^{-3}	$5*10^{-3}$
D_0	8.105	8.105	8.105	8.106	8.108	8.109	8.118	8.129	8.181
D_1	9.127	9.127	9.127	9.127	9.128	9.130	9.130	9.133	9.133
D_2	9.550	9.550	9.549	9.548	9.548	9.549	9.548	9.549	9.570
D_3	10.525	10.525	10.525	10.525	10.525	10.525	10.526	10.524	10.513
D_4	10.979	10.979	10.979	10.979	10.979	10.978	10.980	10.982	10.970
D_5	11.404	11.404	11.404	11.405	11.404	11.403	11.392	11.417	11.384
D_6	11.671	11.671	11.671	11.671	11.671	11.671	11.668	11.663	11.643
Scr. Int Len	1080450	1032822	958734	902286	762930	711774	563598	503622	357210
Unscr. Int Len	1080450	1080450	1080450	1080450	1080450	1080450	1080450	1080450	1080450
% Unscr. Int	1.00	0.96	0.89	0.84	0.71	0.66	0.52	0.47	0.33

Table A.5.: IP-CCSD[1]$_{CC2}$ ionization potentials (in [eV]) for a linear chain of six glycine units with local approximations for the ground state (P_0: $R_0 = 10$ bohr, orbital domains: BP= 0.98, `iext=1`) and un-truncated pair lists $P_{\bar{I}}$. The AB-threshold for the intermediate $\hat{Y}_{abc}$ is set to zero and the ij-threshold to 10^{-4}. Different AB-thresholds for intermediate *oovv* are tested. *Len oovv max* is the highest possible number of entries, *Len oovv* is the actual number of entries and *% oovv* is the ratio *Len oovv* to *Len oovv max*.

	0.0	10^{-3}	$3 * 10^{-3}$	10^{-2}	$3 * 10^{-2}$	10^{-1}
D_0	8.713	8.713	8.713	8.713	8.713	not conv
D_1	8.806	8.806	8.806	8.806	8.807	not conv
D_2	9.164	9.164	9.164	9.164	9.165	not conv
D_3	9.334	9.334	9.334	9.334	9.334	not conv
D_4	9.442	9.442	9.442	9.442	9.443	not conv
D_5	9.613	9.613	9.613	9.613	9.613	not conv
D_6	9.709	9.709	9.709	9.709	9.711	not conv
Len oovv	323441590	203901604	170896512	139690768	106123180	76189444
Len oovv max	992250000	992250000	992250000	992250000	992250000	992250000
% oovv	0.33	0.21	0.17	0.14	0.11	0.08

Table A.6.: IP-CCSD[1]$_{CC2}$ ionization potentials (in [eV]) for the 10-arylisoalloxazine derivative with local approximations for the ground state (P_0: $R_0 = 10$ bohr, orbital domains: BP= 0.98, `iext=1`) and un-truncated pair lists $P_{\bar{I}}$. The AB-threshold for the intermediate $\hat{Y}_{abc}$ is set to zero and the ij-threshold to 10^{-4}. Different AB-thresholds for intermediate *oovv* are tested. *Len oovv max* is the highest possible number of entries, *Len oovv* is the actual number of entries and *% oovv* is the ratio *Len oovv* to *Len oovv max*.

	0.0	10^{-3}	$3 * 10^{-3}$	10^{-2}	$3 * 10^{-2}$	10^{-1}
D_0	7.949	7.949	7.949	7.949	6.193	not conv
D_1	8.601	8.601	8.601	8.601	7.446	not conv
D_2	8.777	8.777	8.777	8.777	7.715	not conv
D_3	9.068	9.068	9.068	9.068	7.944	not conv
D_4	9.101	9.101	9.101	9.101	8.595	not conv
D_5	9.281	9.281	9.281	9.281	8.751	not conv
D_6	9.386	9.386	9.386	9.386	8.781	not conv
Len oovv	582885252	573933942	524778066	405006300	286367904	174261420
Len oovv max	1270352164	1270352164	1270352164	1270352164	1270352164	1270352164
% oovv	0.46	0.45	0.41	0.32	0.23	0.14

Table A.7.: IP-CCSD[1]$_{CC2}$ ionization potentials (in [eV]) for a linear chain of six glycine units with local approximations for the ground state (P_0: $R_0 = 10$ bohr, orbital domains: BP= 0.98, `iext=1`) and un-truncated pair lists $P_{\bar{I}}$. The AB-threshold for the intermediate *oovv* is set to zero and the ij-threshold to 10^{-4}. Different AB-thresholds for intermediate $\hat{Y}_{abc}$ are tested. *Len abc max* is the highest possible number of entries, *Len abc* is the actual number of entries and *% abc* is the ratio *Len abc* to *Len abc max*.

	0.0	10^{-3}	$3*10^{-3}$	10^{-2}	$3*10^{-2}$	10^{-1}
D_0	8.713	8.713	8.713	8.713	8.713	8.712
D_1	8.806	8.806	8.806	8.806	8.806	8.805
D_2	9.164	9.164	9.164	9.164	9.164	9.163
D_3	9.334	9.334	9.334	9.334	9.334	9.332
D_4	9.442	9.442	9.442	9.442	9.441	9.441
D_5	9.613	9.613	9.613	9.613	9.613	9.611
D_6	9.709	9.709	9.709	9.709	9.709	9.708
Len abc	63660600	41782966	34877224	28448640	21475180	15290532
Len abc max	91125000	91125000	91125000	91125000	91125000	91125000
% abc	0.70	0.46	0.38	0.31	0.24	0.17

Table A.8.: IP-CCSD[1]$_{CC2}$ ionization potentials (in [eV]) for the 10-arylisoalloxazine derivative with local approximations for the ground state (P_0: $R_0 = 10$ bohr, orbital domains: BP= 0.98, `iext=1`) and un-truncated pair lists $P_{\bar{I}}$. The AB-threshold for the intermediate *oovv* is set to zero and the ij-threshold to 10^{-4}. Different AB-thresholds for intermediate $\hat{Y}_{abc}$ are tested. *Len abc max* is the highest possible number of entries, *Len abc* is the actual number of entries and *% abc* is the ratio *Len abc* to *Len abc max*.

	0.0	10^{-3}	$3*10^{-3}$	10^{-2}	$3*10^{-2}$	10^{-1}
D_0	7.949	7.949	7.949	7.949	7.948	7.944
D_1	8.601	8.601	8.601	8.601	8.600	8.599
D_2	8.777	8.777	8.777	8.777	8.776	8.774
D_3	9.068	9.068	9.068	9.068	9.068	9.067
D_4	9.101	9.101	9.101	9.101	9.101	9.097
D_5	9.281	9.281	9.281	9.281	9.281	9.279
D_6	9.386	9.386	9.386	9.386	9.386	9.384
Len abc	126506008	124563268	113894764	87900200	62151616	37820680
Len abc max	126506008	126506008	126506008	126506008	126506008	126506008
% abc	1.00	0.98	0.90	0.69	0.49	0.30

Table A.9.: Results for D21L6. Locality for the ground state (R_0: 10 bohr, BP: 0.98, `iext=1`) and excited states (κ_e: 0.99, AB-threshold: 10^{-2}, ij-threshold: 10^{-4}) was utilized. Timings for ground-state, the initial diagonalization (Dav 1), and the second diagonalization with additional terms included (Dav 2) are all given in [h], whereas averaged timings for building the Jacobian trial vector products are given in [s]. Iter gives the number of iterations until convergence was reached for the initial (extension 1) and the second diagonalization (extension 2). In parenthesis the number of Jacobian × trial vector products are given, which were computed during the iterative diagonalizations. $N_{\bar{I}}$ denotes the initial number of pairs for the second diagonalization and the final number of pairs when the state converged. The calculations were performed in parallel mode on seven AMD 6180 SE cores @ 2.5 GHz.

		IP-CCSD$[k]_M$						
		$k=0$		$k=1$		$k=2$		$k=f$
	State	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2
Ionization energy [eV]	D_0	5.414	5.317	6.082	5.968	6.354	6.257	6.258
	D_1	6.516	6.400	6.935	6.817	7.285	7.186	7.188
	D_2	7.363	7.316	7.811[a]	7.639[a]	8.165	8.145	8.145
	D_3	7.533	7.451	7.888	7.847	8.305	8.239	8.241
	D_4	7.891	7.862	8.179[a]	8.021[a]	8.703	8.700	8.701
	D_5	8.237	8.146	8.436	8.340	8.860	8.793	8.794
	D_6	8.271	8.195	8.476	8.415	8.997	8.923	8.924
Jacobu 1 elapsed [s]	——	12.89	15.04	14.57	16.12	13.60	15.04	15.51
Iter 1	——	28 (143)	29 (134)	29 (143)	29 (134)	29 (143)	29 (134)	29 (134)
Dav 1 elapsed [h]	——	3.9	3.7	4.2	4.0	4.0	4.0	4.0

Table continues on next page

[a] ionized state dominated by m=$\frac{3}{2}$ character

Table A.9.: Continued from previous page

		IP-CCSD[k]$_M$						
		$k=0$		$k=1$		$k=2$		$k=f$
	State	ADC(2)	M=CC2	M=MP2	M=CC2	M=MP2	M=CC2	M=CC2
Jacobu 2 elapsed [s]	D_0	——	——	170.08	172.06	242.33	340.21	363.49
	D_1	——	——	147.77	161.29	217.21	290.30	321.57
	D_2	——	——	170.16	223.65	191.13	220.41	238.25
	D_3	——	——	141.62	163.69	188.57	219.58	226.32
	D_4	——	——	230.02	290.81	181.33	205.77	216.12
	D_5	——	——	216.40	269.57	180.78	212.60	217.13
	D_6	——	——	211.81	260.50	187.22	217.90	223.93
Iter 2	——	——	——	33 (170)	34 (138)	34 (102)	31 (137)	31 (137)
Dav 2 elapsed [h]	——	——	——	17.9	22.8	12.4	15.1	15.6
$N_{\bar{I}}$	D_0	——	——	1042/1623	1078/1476	1042/1368	1078/1742	1078/1745
	D_1	——	——	1025/1147	1046/1102	1025/1323	1046/1656	1046/1653
	D_2	——	——	763/1065	788/1024	763/1799	788/1662	788/1659
	D_3	——	——	1207/1450	1217/1407	1207/1777	1217/1684	1217/1685
	D_4	——	——	1104/2331	1113/2124	1104/1751	1113/1633	1113/1636
	D_5	——	——	951/2323	713/2288	951/1649	713/1528	713/1526
	D_6	——	——	1117/2217	1201/2302	1117/2399	1201/2489	1201/2489
HF+GS elapsed [h]	——	1.8	5.2	1.8	5.1	1.9	5.0	5.1
ES elapsed [h]	——	4.4	4.3	22.7	27.5	17.1	19.8	20.5
Total elapsed time [h]	——	6.2	9.5	24.5	32.6	18.9	24.9	25.6

B. Coupled Cluster diagrams

In the course of this work, diagrammatic techniques are used to obtain programmable working equations. The rules, which are used to translate diagrams into mathematical expressions, are presented in Section 1.3 of the Introduction. Diagrams for a pure CC2 Jacobian are already illustrated in Fig. 2.1 in the main text. In this section, higher-order diagrams, which are added on top of the CC2 Jacobian, are presented. This includes all extra diagrams of the $m = \frac{3}{2}$ excitation space, as well as those introduced by going to the $m = \frac{5}{2}$ excitation space.

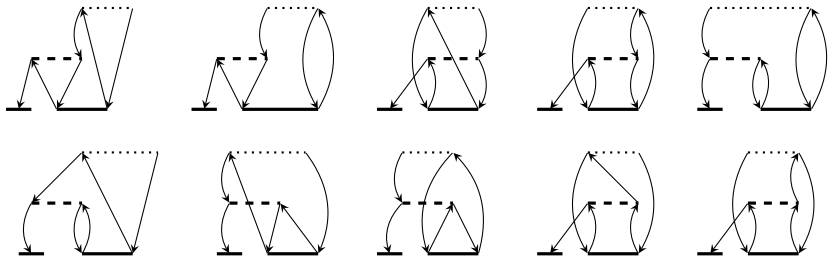

Figure B.1.: Extra second-order diagrams of the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ block of the Jacobian. Corresponding expressions are found in Eq. (2.36).

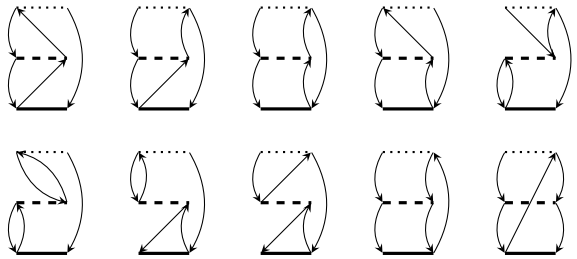

Figure B.2.: Extra first-order diagrams of the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ block of the Jacobian. Corresponding expressions are found in Eq. (2.36).

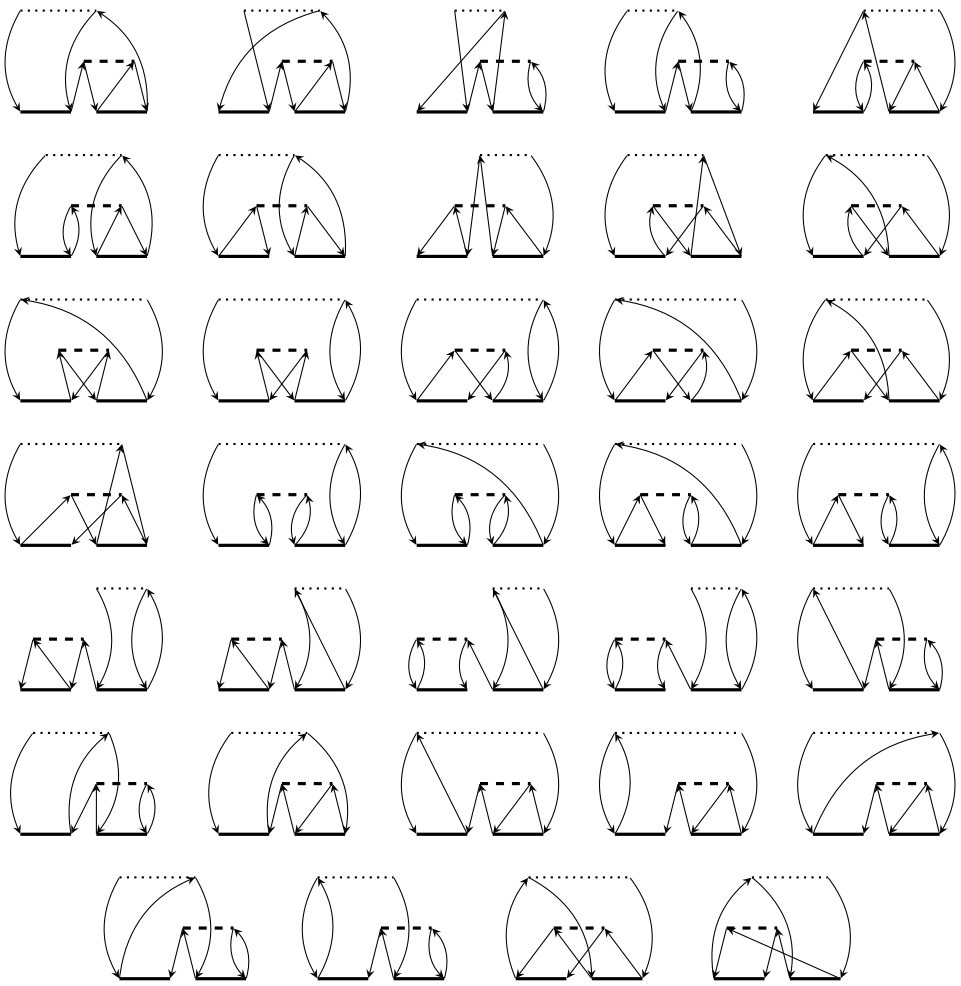

Figure B.3.: Extra second-order diagrams of the $\mathbf{A}_{\frac{3}{2}\frac{3}{2}}$ block of the Jacobian. Corresponding expressions are summarized in Eq. (2.39).

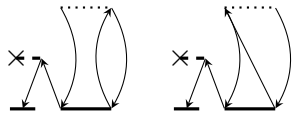

Figure B.4.: Extra fourth-order diagrams of the $\mathbf{A}_{\frac{3}{2}\frac{1}{2}}$ block of the Jacobian. Corresponding expressions are summarized in Eq. (2.42).

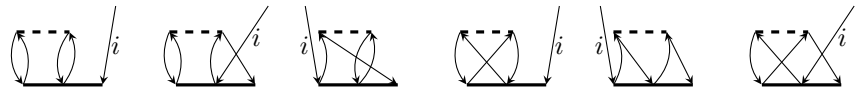

Figure B.5.: Contravariant diagrams of the $\mathbf{A}_{\frac{1}{2}\frac{5}{2}}$ block of the Jacobian resulting in Eq. (3.8).

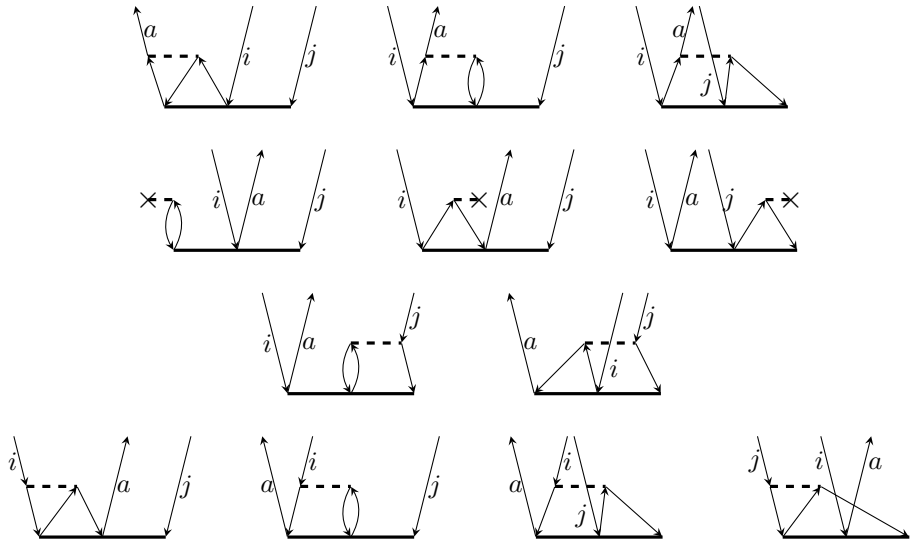

Figure B.6.: Contravariant diagrams of the $\mathbf{A}_{\frac{3}{2}\frac{5}{2}}$ block of the Jacobian resulting in Eq. (3.9).

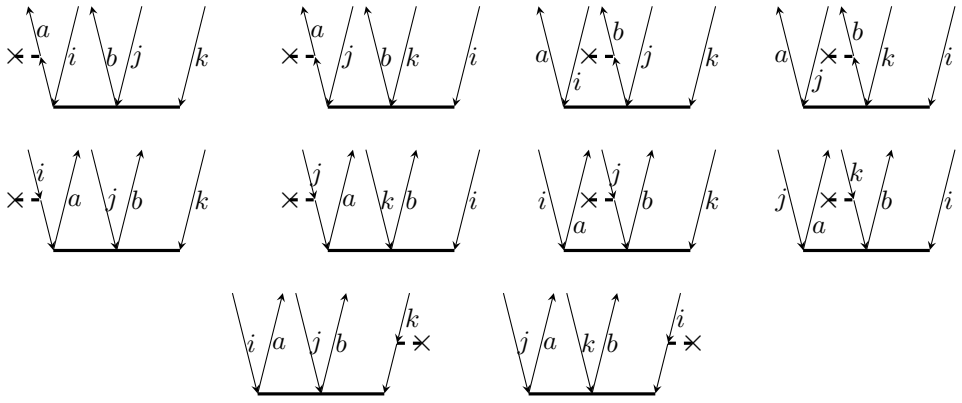

Figure B.7.: Contravariant diagrams of the $\mathbf{A}_{\frac{5}{2}\frac{5}{2}}$ block of the Jacobian contributing to the $\mathbf{Q}$ intermediate in Eq. (3.11).

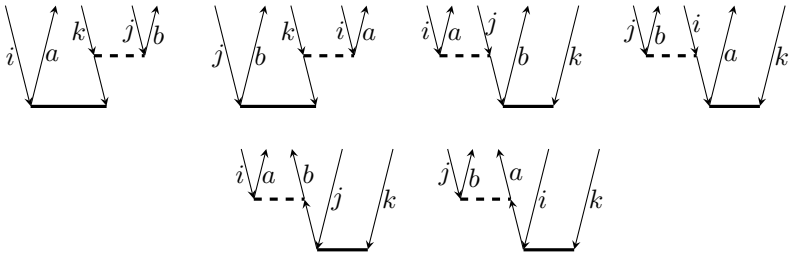

Figure B.8.: Contravariant diagrams of the $\mathbf{A}_{\frac{5}{2}\frac{3}{2}}$ block of the Jacobian contributing to the $\mathbf{W}^{ijk}$ intermediate in Eq. (3.11). Every diagram has a corresponding exchange diagram, which are not drawn explicitly. Exchange diagrams are obtained by renaming indices $i \to j$, $j \to k$ and $k \to i$.

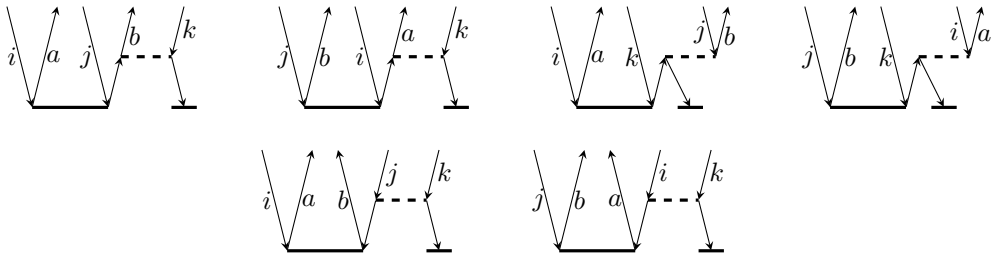

Figure B.9.: Contravariant diagrams of the $\mathbf{A}_{\frac{5}{2}\frac{1}{2}}$ block of the Jacobian contributing to the $\mathbf{W}^{ijk}$ intermediate in Eq. (3.11). Every diagram has a corresponding exchange diagram, which are not drawn explicitly. Exchange diagrams are obtained by renaming indices $i \to j$, $j \to k$ and $k \to i$.

Bibliography

[1] E. J. Baerends and O. V. Gritsenko. *J. Phys. Chem. A*, **101**, 5383, (1997).

[2] R. G. Parr and Y. Weitao. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, (1994).

[3] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, (1996).

[4] C. Møller and M. S. Plesset. *Phys. Rev.*, **46**, 618, (1934).

[5] R. J. Bartlett and J. Noga. *Chem. Phys. Lett.*, **150**, 29, (1988).

[6] R. J. Bartlett, S. A. Kucharski, and J. Noga. *Chem. Phys. Lett.*, **155**, 133, (1989).

[7] D. Kats, D. Usvyat, and M. Schütz. *Phys. Rev. A*, **83**, 062503, (2011).

[8] G. Wälz, D. Kats, D. Usvyat, T. Korona, and M. Schütz. *Phys. Rev. A*, **86**, 052519, (2012).

[9] O. Christiansen, H. Koch, and P. Jørgensen. *Chem. Phys. Lett.*, **243**, 409, (1995).

[10] G. D. Purvis and R. J. Bartlett. *J. Chem. Phys.*, **76**, 1910, (1982).

[11] H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Merás, and T. Helgaker. *J. Chem. Phys.*, **106**, 1808, (1997).

[12] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. *Chem. Phys. Lett.*, **157**, 479-483, (1989).

[13] R. J. Bartlett, J. D. Watts, S. A. Kucharski, and J. Noga. *Chem. Phys. Lett.*, **165**, 513-522, (1990).

[14] J. Noga and R. J. Bartlett. *J. Chem. Phys.*, **86**, 7041, (1987).

[15] G. E. Scuseria and H. F. Schäfer. *Chem. Phys. Lett*, **152**, 382, (1988).

[16] J. A. Pople, J. S. Binkley, and R. Seeger. *Int. J. Quantum Chem. Symp.*, **10**, 1-19, (1976).

[17] R. J. Bartlett. *Ann. Rev. Phys. Chem.*, **32**, 359-401, (1981).

[18] D. Kats, T. Korona, and M. Schütz. *J. Chem. Phys.*, **125**, 104106, (2006).

[19] D. Kats and M. Schütz. *J. Chem. Phys.*, **131**, 124117, (2009).

[20] D. Kats, T. Korona, and M. Schütz. *J. Chem. Phys.*, **127**, 064107, (2007).

[21] D. Kats and M. Schütz. *Z. Phys. Chem.*, **224**, 601, (2010).

[22] K. Freundorfer, D. Kats, T. Korona, and M. Schütz. *J. Chem. Phys.*, **133**, 244110, (2010).

[23] K. Ledermüller, D. Kats, and M. Schütz. *J. Chem. Phys.*, **139**, 084111, (2013).

[24] K. Ledermüller and M. Schütz. *J. Chem. Phys.*, **140**, 164113, (2014).

[25] P. Pulay. *Chem. Phys. Lett.*, **100**, 151, (1983).

[26] J. Pipek and P. G. Mezey. *J. Chem. Phys.*, **90**, 4916, (1989).

[27] S. F. Boys. Localized orbitals and localized adjustment functions. In *Quantum Theory of Atoms, Molecules, and the Solid State, A Tribute to John C. Slater*, page 253. Academic Press, New York, (1966).

[28] J. W. Boughton and P. Pulay. *J. Comput. Chem.*, **14**, 736, (1993).

[29] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz. *Comput. Mol. Sci.*, **2**, 242, (2012).

[30] B. I. Dunlap. *Phys. Chem. Chem. Phys.*, **2**, 2113, (2000).

[31] E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martínez. *J. Chem. Phys.*, **138**, 124111, (2013).

[32] E. G. Hohenstein, S. I. L. Kokkila, R. M. Parrish, and T. J. Martínez. *J. Phys. Chem. B*, **117**, 12972, (2013).

[33] G. Wälz, D. Usvyat, T. Korona, and M. Schütz. *J. Chem. Phys.*, **144**, 084117, (2016).

[34] O. Christiansen, P. Jørgensen, and C. Hättig. *Int. J. Quantum Chem.*, **68**, 1, (1998).

[35] M. Nooijen and J. G. Snijders. *Int. J. Quantum Chem.*, **48**, 15, (1993).

[36] J. F. Stanton and J. Gauss. *J. Chem. Phys.*, **101**, 8938, (1994).

[37] J. F. Stanton and J. Gauss. *J. Chem. Phys.*, **103**, 1064, (1995).

[38] J. F. Stanton and J. Gauss. *Theor. Chim. Acta*, **93**, 303, (1996).

[39] J. F. Stanton and J. Gauss. *J. Chem. Phys.*, **111**, 8785, (1999).

[40] M. Musiał, S. A. Kucharski, and R. J. Bartlett. *J. Chem. Phys.*, **118**, 1128, (2003).

[41] Y. J. Bomble, J. C. Saeh, J. F. Stanton, P. G. Szalay, M. Kallay, and J. Gauss. *J. Chem. Phys.*, **122**, 154107, (2005).

[42] J. R. Gour and Piotr Piecuch. *J. Chem. Phys.*, **125**, 234107, (2006).

[43] M. Kamiya and S. Hirata. *J. Chem. Phys.*, **125**, 074111, (2006).

[44] P. U. Manohar, J. F. Stanton, and A. I. Krylov. *J. Chem. Phys.*, **131**, 114112, (2009).

[45] A. Landau, K. Khistyaev, S. Dolgikh, and A. I. Krylov. *J. Chem. Phys.*, **132**, 014109, (2010).

[46] E. R. Davidson. *J. Comput. Phys.*, **17**, 87, (1975).

[47] K. Hirao and H. Nakatsuji. *J. Comput. Phys.*, **45**, 246, (1982).

[48] B. Helmich and C. Hättig. *Comp. Theo. Chem*, **1040**, 35, (2014).

[49] J. Schirmer. *Phys. Rev. A*, **26**, 2395, (1982).

[50] A. B. Trofimov and J. Schirmer. *J. Phys. B*, **28**, 2299, (1995).

[51] O. Christiansen, H. Koch, and P. Jørgensen. *J. Chem. Phys.*, **103**, 7429, (1995).

[52] D. Kats. *J. Chem. Phys.*, **141**, 244101, (2014).

[53] T. H. Dunning Jr. *J. Chem. Phys.*, **90**, 1007, (1989).

[54] F. Weigend, A. Köhn, and C. Hättig. *J. Ch*, **116**, 3175, (2002).

[55] F. Weigend. *Phys. Chem. Chem. Phys.*, **4**, 4285, (2002).

[56] CFOUR, a quantum chemical program package written by J.F. Stanton, J. Gauss, M.E. Harding, P.G. Szalay with contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, Y.J. Bomble, L. Cheng, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W.J. Lauderdale, F. Lipparini, D.A. Matthews, T. Metzroth, L.A. Mück, D.P. O'Neill, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

[57] J.-H. Yum, D. P. Hagberg, S.-J. Moon, K. M. Karlsson, T. Marinado, L. Sun, A. Hagfeldt, M. K. Nazeeruddin, and M. Grätzel. *Angewandte Chemie, International Edition*, **48**, 1576, (2009).

[58] P. G. Szalay and J. Gauss. *J. Chem. Phys.*, **112**, 4027, (2000).

[59] M. Schütz. *Proposal*, (2016).

[60] X. Liu, T. N. V. Karsili, A. L. Sobolewski, and W. Domcke. *J. Phys. Chem. B*, **119**, 10664, (2015).

[61] G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, and L. Seijo. *Comput. Mater. Sci.*, **28**, 222, (2003).

[62] M. Schütz. *J. Chem. Phys.*, **113**, 9986, (2000).

[63] M. Schütz and H.-J. Werner. *Chem. Phys. Lett.*, **318**, 370, (2000).

[64] M. Schütz, G. Hetzer, and H.-J. Werner. *J. Chem. Phys.*, **111**, 5692, (1999).

[65] J. Daďová, S. Kümmel, C. Feldmeier, J. Cibulková, R. Pažout, J. Maixner, R. M. Gschwind, B. König, and R. Cibulka. *Chem. Eur. J.*, **19**, 1066, (2013).