Internally contracted multireference coupled-cluster methods

Dissertation

zur Erlangung des Grades
„Doktor der Naturwissenschaften“
im Promotionsfach Chemie

am Fachbereich Chemie, Pharmazie und Geowissenschaften
der Johannes Gutenberg-Universität
in Mainz

von

Matthias Hanauer
geboren in Mainz

Mainz, 2013

Dekan:

1. Berichterstatter:

2. Berichterstatter:

3. Berichterstatter:

Datum der mündlichen Prüfung: 24. April 2013

D77 (Dissertation Universität Mainz)

Acknowledgments

Contents

1 Introduction

Chemical systems obey quantum mechanical laws [1]. Since the discovery of these laws, research in the field of quantum chemistry has provided methods for the *ab initio* calculation of molecular properties, for example energies and equilibrium structures, vibrational frequencies, excitation energies as well as electrical and magnetic properties [2–4]. Although the first principles are known, their application to systems larger than the hydrogen atom has to rely on approximations. The driving force behind developments in quantum chemistry is the desire to improve its predictive power for small systems and to extend it to ever larger systems. Progress in computer science greatly assists this effort through a rising computing capacity [5,6].

For systems with a large energy difference between the ground and excited states, including most closed-shell molecules, the Hartree–Fock approximation [2] offers a qualitatively correct theoretical description. It offers a simple picture of the electronic structure by modeling it as a single Slater determinant, in which the electrons occupy a set of molecular orbitals. This concept has been important in framing the modern understanding of chemistry, as it provides a useful model for chemical reactivity based on the occupied and the energetically lowest virtual (unoccupied) orbitals [1,7,8]. The Hartree–Fock approximation has its limits. Being a mean-field approximation, it neglects the correlated motion of the electrons due to their electrostatic repulsion. Nevertheless, it serves as a convenient starting point for more accurate theories that account for electron correlation by incorporating contributions of excited electronic configurations into the system's wave function. Mathematically, these excited configurations can be obtained from the Hartree–Fock determinant by excitation operators that move electrons from occupied into virtual orbitals [3]. When aiming at predictions with chemical accuracy (i.e., errors in energy differences below 1 kcal/mol $\approx$ 4 kJ/mol), coupled-cluster theory [9–11] with single and double excitations and an approximate inclusion of triple excitations, CCSD(T), is the well-established method of choice [12–15].

However, the Hartree–Fock approximation breaks down for systems with ground states that are either degenerate or energetically close to excited states (quasi-degenerate) [16]. A qualitative description of the wave function requires more than one electronic configuration in this case. As a consequence, coupled-cluster methods using a reference function from Hartree–Fock theory fail, unless the maximum excitation level is increased accordingly at the expense of computational efficiency [3]. Chemistry abounds with systems having electronic structures of such a complicated nature, as it occurs in many transition states of chemical reactions [17] and during bond breaking processes [3] as well as in open-shell states (e.g., diradicalic singlet states [18]) and a multitude of transition metal compounds [19–21]. Multiconfigurational methods achieve a qualitatively correct description of these systems by allowing several configurations in the wave function that differ in the distribution of a number of electrons among a selected number of valence orbitals, the so-called active orbitals [22–25].

The goal of this work is the development of a coupled-cluster method that can use a multiconfigurational wave function as a reference and thereby carries over the accuracy of CCSD(T) to systems with complicated electronic structure.

The past decades have witnessed the formulation of plenty of multireference coupled-cluster (MRCC) approaches, but none has yet evolved into a standard method [26–29]. Most of the existing MRCC approaches fall short of at least one of three important

requirements: To reliably provide accurate results, the energy should scale correctly with a system's size (size-extensivity). Like the underlying multiconfigurational description, it should treat all active orbitals on equal footing and be invariant with respect to a unitary transformation of these orbitals (orbital invariance). At the same time, the ansatz for the wave function should allow an efficient treatment of electron correlation, using a compact parameterization that helps to keep the computational cost low.

The internally contracted ansatz is one of the few MRCC approaches that promise to fulfill all of these requirements [30–33]. It results from the ansatz of single-reference coupled-cluster theory by substituting a multiconfigurational wave function for the reference and by allowing excitation operators that can move electrons both into and out of the active orbitals. However, these operators do not commute with each other and cause a large number of terms in the working equations. This increased complexity has so far prevented an implementation of internally contracted multireference coupled-cluster (ic-MRCC) theory, despite its early formulation in the pioneering work by Banerjee and Simons [34–37] thirty years ago. It is possible to simulate the method by expanding the ic-MRCC wave function in a complete basis of determinants. Evangelista and Gauss [38] have employed this strategy to study the method's performance for small model systems, confirm its orbital invariance and test approximations that restrict the working equations to a finite power in the wave function parameters (cluster amplitudes).[1] However, because of its factorial computational scaling, this implementation cannot offer a computational advantage over an exact treatment of electron correlation and is limited to very small systems.

In order to unfold the full potential of ic-MRCC theory, one has to derive and implement the explicit expressions as a sum of tensor products that can be evaluated at a reduced cost, an intractable task without help from automated implementation techniques [32]. Automated approaches are not unprecedented in the implementation of related theories. For instance, Neuscamman, Yanai and Chan [39] use this strategy in the so-called canonical transformation theory (CT) [40–42]. In contrast to ic-MRCC theory, CT theory involves a unitary transformation of the Hamiltonian. Because this would lead to even more complex equations in a full realization of the method, the present CT approach invokes several approximations that may seriously affect the accuracy of the method [42, 43].

Therefore, the subject of the present work is the development of ic-MRCC methods with a favorable computational scaling while using only approximations that preserve the methods' inherent accuracy. To this end, ic-MRCC theory is implemented in a program that allows an automated derivation and evaluation of the required expressions and has been successfully used before in the context of other complicated electronic structure theories [44, 45]. This automated approach enables the efficient realization of different variants of the theory. Testing the performance of these variants on simple model systems helps us not only to forge an ic-MRCC method of choice but also to gain a deep understanding of many theoretical aspects of ic-MRCC theory. In particular, the question how to best deal with linear dependencies between the working equations, a common problem in internally contracted multireference approaches [31, 42, 46–52], has not yet been answered in a satisfactory way and forms one of the focal points of this thesis.

---

[1]A similar implementation of ic-MRCC theory exists due to J. Olsen (unpublished).

The theoretical introduction (chap. 2) provides the foundation of this work, putting an emphasis on electron correlation, and introduces the parent electronic structure theories of ic-MRCC before ending with an overview of the existing MRCC approaches. Chapter 3 then lays out the details of ic-MRCC theory. It contains new insights on several aspects of ic-MRCC theory, including its formal structure and complexity, its computational scaling, and the impact of the treatment of linear dependencies on the method's size-extensivity. In the context of the latter aspect, two improved procedures are presented. Benchmark studies of ic-MRCC theory on the potential energy surfaces of six model systems (BeH_2, HF, LiF, H_2O, N_2 and Be_3) are discussed in chap. 4. These studies also compare the performance of ic-MRCC theory with that of other multireference methods and analyse the strengths and weaknesses of different flavors and approximations of ic-MRCC theory. In chap. 5, the reader may find a perturbative analysis of ic-MRCC theory and the formulation of various methods to approximately include triple excitations based on perturbative arguments. These methods are tested on selected model systems and applied to the calculation of the equilibrium structure and the harmonic vibrational frequencies of the ozone molecule. The method ic-MRCCSD(T), which efficiently includes the effect of triple excitations in a non-iterative manner on top of an ic-MRCC singles-and-doubles (ic-MRCCSD) calculation, is further applied to the singlet-triplet splittings of benzynes (C_6H_4), the structure and vibrational frequencies of Ni_2O_2 and the ring-opening reaction of an azirine compound with the molecular formula C_6H_7NO. Finally, chap. 6 aims to extend the applicability of ic-MRCC theory to large numbers of active orbitals by truncating the cumulant expansion of the reduced density matrices of the reference function. An application of ic-MRCCSD(T) to the potential energy surface of the formally hextuply bonded chromium dimer closes this chapter and is followed by a summary of this thesis and an outlook on the ic-MRCC method's future.

2 Theoretical fundamentals

2.1 The second-quantized electronic Schrödinger equation

Atoms and molecules are many-body systems consisting of positively charged nuclei and negatively charged electrons, held together by the Coulomb interaction. In case of the relatively light elements (up to the fourth row of the periodic table) and the molecular properties studied in this work, the nuclei are to a very good approximation treated as point particles, and non-relativistic quantum mechanics is employed with success. As the current work focuses on time-independent properties of stationary states, the time-independent Schrödinger equation serves as starting point. In addition, two further common approximations are applied [2]:

The first is the Born–Oppenheimer approximation, which exploits that the nuclear and electronic motion is correlated in a very one-sided manner: The nuclei are thousands of times heavier than the electrons and therefore generally move much slower than electrons. They essentially experience only the internuclear repulsion and an effective potential created by the much faster moving electrons. The electrons, on the other hand, are assumed to follow the nuclear motion instantaneously. From their perspective, the nuclei appear to be 'clamped' in space, such that the nuclear kinetic energy can be neglected to yield a Schrödinger equation for the electrons only. Thus, each electronic state gives rise to a multidimensional potential energy surface which governs the nuclear motion for this state.

The electronic wave function for a given molecular structure is a complex-valued function of the N electrons' spatial coordinates $\mathbf{r}_i$ and spin coordinates σ_i (with $i = 1, 2, \ldots, N$), to be combined into $\mathbf{x}_i$ in the following. According to Pauli's antisymmetry principle, it must change its sign upon exchange of two electrons' coordinates. The simplest N-electron function meeting this requirement is a Slater determinant,

$$\Phi_{p_1 p_2 \cdots p_N}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{p_1}(\mathbf{x}_1) & \varphi_{p_2}(\mathbf{x}_1) & \cdots & \varphi_{p_N}(\mathbf{x}_1) \\ \varphi_{p_1}(\mathbf{x}_2) & \varphi_{p_2}(\mathbf{x}_2) & \cdots & \varphi_{p_N}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{p_1}(\mathbf{x}_N) & \varphi_{p_2}(\mathbf{x}_N) & \cdots & \varphi_{p_N}(\mathbf{x}_N) \end{vmatrix}. \tag{2.1}$$

It is composed of spin orbitals (one-electron functions) $\varphi_p(\mathbf{x})$, which are products of normalized and mutually orthogonal spatial orbitals $\phi_p(\mathbf{r})$ and either of the two one-electron spin functions $\alpha(\sigma)$ (spin-up) or $\beta(\sigma)$ (spin-down). If the spatial orbitals form a complete one-particle basis, the set of all possible Slater determinants $\{\Phi_{p_1 \cdots p_N}(\mathbf{x}_1, \ldots, \mathbf{x}_N)\}$ spans the complete space of N-electron functions and thus allows an exact representation of the electronic wave function.

In practice, a theoretical description is only possible for a finite one-particle basis set. This restriction of the one-particle basis set is the second common approximation and introduces the basis set error. To keep this error small despite a finite basis set size, an appropriate selection of basis functions is mandatory. An efficient choice for molecular electronic structure calculations is an atomic orbital basis set constructed from linear combinations of spherical-harmonic Gaussian-type orbitals, centered at the positions of the nuclei.

The dependence of the wave function on the orbitals can be made implicit by turning to the formalism of second quantization [3]. Each Slater determinant, featuring a number

of occupied spin orbitals $\varphi_{p_1}, \varphi_{p_2}, \ldots, \varphi_{p_N}$, is then solely characterized by the occupancy of the spin orbitals. It is constructed by creating N electrons in N spin orbitals on top of a vacuum state $|\rangle$, using the creation operators $\hat{a}_p^\dagger$:

$$|\Phi_{p_1 p_2 \cdots p_N}\rangle = \hat{a}_{p_1}^\dagger \hat{a}_{p_2}^\dagger \cdots \hat{a}_{p_N}^\dagger |\rangle. \tag{2.2}$$

Operators acting on n electrons are then conveniently expressed in terms of their matrix elements with respect to the spin orbitals, each of them associated with an excitation operator $\hat{a}_{p_1 \cdots p_n}^{q_1 \cdots q_n} = \hat{a}_{q_1}^\dagger \cdots \hat{a}_{q_n}^\dagger \hat{a}_{p_n} \cdots \hat{a}_{p_1}$ that first annihilates and then creates n electrons. In the language of second quantization, the electronic Schrödinger equation takes the form

$$\hat{H}|\Psi\rangle = E|\Psi\rangle, \tag{2.3}$$

with the electronic wave function $|\Psi\rangle$ and the second-quantized electronic Hamiltonian

$$\hat{H} = h_{\text{nuc}} + \sum_{pq} h_q^p \hat{a}_p^q + \tfrac{1}{4} \sum_{pqrs} g_{rs}^{pq} \hat{a}_{pq}^{rs}. \tag{2.4}$$

The internuclear repulsion energy

$$h_{\text{nuc}} = \sum_{k<l} \frac{Z_k Z_l}{R_{kl}}, \tag{2.5}$$

with nuclear charges Z_k and internuclear distances R_{kl}, is included in the electronic Hamiltonian such that the energy E coincides with a Born–Oppenheimer potential energy surface. The one-electron operator in $\hat{H}$ incorporates the kinetic energy of the electrons and the attraction between electrons and nuclei at positions $\mathbf{R}_k$,

$$h_q^p = \int d\mathbf{x} \, \varphi_q^*(\mathbf{x}) \left(-\tfrac{1}{2} \boldsymbol{\nabla}^2 - \sum_k \frac{Z_k}{|\mathbf{r} - \mathbf{R}_k|} \right) \varphi_p(\mathbf{x}), \tag{2.6}$$

while the two-electron part involves the repulsion between electrons separated by r_{12},

$$g_{rs}^{pq} = \int \int d\mathbf{x}_1 \, d\mathbf{x}_2 \left(\frac{\varphi_r^*(\mathbf{x}_1)\varphi_s^*(\mathbf{x}_2)\varphi_p(\mathbf{x}_1)\varphi_q(\mathbf{x}_2)}{r_{12}} - \frac{\varphi_r^*(\mathbf{x}_1)\varphi_s^*(\mathbf{x}_2)\varphi_q(\mathbf{x}_1)\varphi_p(\mathbf{x}_2)}{r_{12}} \right). \tag{2.7}$$

Equations (2.5) to (2.7) make use of atomic units (1 Hartree $= 1 \, E_h = 2\,625.5 \, \text{kJ/mol}$ [2]).

2.2   Electron correlation

2.2.1   Fermi correlation & the Hartree–Fock method

Electrons interact with each other and hence their motion is correlated. The underlying statistical distribution corresponds to the absolute square of the wave function, $|\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)|^2$, being the joint probability density of the electrons' spatial positions and spin orientations. For statistically independent particles, this joint probability density would be a product of probability densities for the individual particles, e.g.,

$$|\Psi_{\text{indep.}}|^2 = |\varphi_{p_1}(\mathbf{x}_1)|^2 |\varphi_{p_2}(\mathbf{x}_2)|^2 \cdots |\varphi_{p_N}(\mathbf{x}_N)|^2. \tag{2.8}$$

However, this form is only realized for a simple product of spin orbitals (Hartree product) which lacks the proper antisymmetry for exchange of Fermions. Antisymmetrization thus

introduces some correlation between the Fermions, commonly called *Fermi correlation*. In particular, an antisymmetric wave function ensures that two electrons cannot have the same values for the spatial and spin coordinates and thereby enforces Pauli's exclusion principle. In general, Fermi correlation reduces the probability of finding two electrons with their spins coupled to a triplet close to each other (*Fermi hole*). However, for two electrons with spins coupled to a singlet, Fermi correlation may have the opposite effect and increase the probability of finding them close to each other [53].

In second quantization, Fermionic wave functions are implicitly antisymmetrized and operators acting on Fermions are represented by antisymmetric tensors. This intrinsically accounts for Fermi correlation. Also, the focus is shifted away from the correlation between electronic coordinates towards the correlation between occupancies of spin orbitals. Since the occupancy of a spin orbital φ_p is probed by the number operator $\hat{a}_p^p$, the probability of finding certain spin orbitals occupied is given by the diagonal elements of the reduced one-particle density matrix $\boldsymbol{\gamma}_1$, which is defined by

$$\gamma_p^q = \langle \Psi | \hat{a}_p^q | \Psi \rangle. \tag{2.9}$$

The joint probability of finding two spin orbitals that are simultaneously occupied is on the other hand given by the diagonal elements of the two-particle density matrix $\boldsymbol{\gamma}_2$, with

$$\gamma_{pq}^{rs} = \langle \Psi | \hat{a}_{pq}^{rs} | \Psi \rangle. \tag{2.10}$$

In this framework, the occupancies of the spin orbitals are statistically independent when $\boldsymbol{\gamma}_2$ is simply an antisymmetrized product of one-particle density matrices [53,54]:

$$\gamma_{pq}^{rs} = \gamma_p^r \gamma_q^s - \gamma_p^s \gamma_q^r. \tag{2.11}$$

Statistical independence of spin orbital occupancies means that the electrons are uncorrelated except for Fermi correlation. From the trace relationship between $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_1$,

$$\sum_r \gamma_{pr}^{qr} = (N-1)\gamma_p^q, \tag{2.12}$$

one can easily deduce that the condition (2.11) implies idempotency of $\boldsymbol{\gamma}_1$,

$$\sum_r \gamma_p^r \gamma_r^q = \gamma_p^q. \tag{2.13}$$

Idempotency in turn restricts the possible eigenvalues of $\boldsymbol{\gamma}_1$ to zero and one. As the spin orbital basis can always be rotated by a unitary transformation such that $\boldsymbol{\gamma}_1$ becomes diagonal, it then falls into a set of occupied and unoccupied spin orbitals, a situation only encountered for a single Slater determinant.

The essence of Hartree–Fock (HF) theory is to use a single determinant $|0\rangle$ as an ansatz for the ground state wave function[2] and variationally optimize the orbitals. In unrestricted Hartree–Fock (UHF) theory the spin orbitals are merely constrained to remain mutually orthogonal, whereas in restricted Hartree–Fock (RHF) theory a common set of spatial orbitals is employed for both spin orientations [2].

---

[2]An exception is restricted open-shell Hartree–Fock (ROHF) theory, which uses a spin-adapted linear combination of determinants for low-spin states. Here, we will consider ROHF theory as a special case of MCSCF theory.

The orbital optimization requires a unitary transformation of the spin orbitals [3],

$$\tilde{\varphi}_p = \sum_q \varphi_q U_{qp}. \tag{2.14}$$

Since any unitary matrix $\mathbf{U}$ can be written as the exponential of an anti-Hermitian matrix $\boldsymbol{\kappa}$ (with $\boldsymbol{\kappa}^\dagger = -\boldsymbol{\kappa}$),

$$\mathbf{U} = e^{-\boldsymbol{\kappa}}, \tag{2.15}$$

it is possible to parametrize a (trial) determinant for a rotated orbital basis in the form

$$|\tilde{0}\rangle = e^{-\hat{\kappa}}|0\rangle, \tag{2.16}$$

with the anti-Hermitian operator

$$\hat{\kappa} = \sum_{pq} \kappa_{pq}(\hat{a}_p^q - \hat{a}_q^p). \tag{2.17}$$

Expanding the energy expectation value

$$\tilde{E} = \langle 0|e^{\hat{\kappa}}\hat{H}e^{-\hat{\kappa}}|0\rangle \tag{2.18}$$

in terms of $\hat{\kappa}$ with help of the Baker–Campbell–Hausdorff (BCH) expansion,

$$e^{-\hat{A}}\hat{H}e^{\hat{A}} = \hat{H} + [\hat{H}, \hat{A}] + \tfrac{1}{2}[[\hat{H}, \hat{A}], \hat{A}] + \tfrac{1}{3!}[[[\hat{H}, \hat{A}], \hat{A}], \hat{A}] + \ldots, \tag{2.19}$$

we obtain the variational condition for the orbital rotations by setting the first derivative of $\tilde{E}$ with respect to κ_{pq} to zero. Since it is always possible to translate the parameters κ_{pq} into orbital rotations according to Eq. (2.14) and set $\kappa_{pq} = 0$ afterwards without affecting the trial determinant, it is sufficient to consider the variational condition in the limit $\boldsymbol{\kappa} = \mathbf{0}$, yielding

$$0 = \langle 0|[\hat{H}, \hat{a}_p^q]|0\rangle. \tag{2.20}$$

Equation (2.20) determines the wave function $|0\rangle$. In principle, its right-hand side can be conveniently evaluated by choosing $|0\rangle$ as Fermi vacuum and representing excitation operators in their normal ordered form with respect to that new vacuum. The relationships between the excitation operators $\hat{a}_p^q$, $\hat{a}_{pq}^{rs}$ and their normal ordered counterparts $\{\hat{a}_p^q\}$, $\{\hat{a}_{pq}^{rs}\}$ are [55]

$$\hat{a}_p^q = \{\hat{a}_p^q\} + \gamma_p^q, \tag{2.21}$$

$$\hat{a}_{pq}^{rs} = \{\hat{a}_{pq}^{rs}\} + \gamma_p^r\{\hat{a}_q^s\} - \gamma_p^s\{\hat{a}_q^r\} - \gamma_q^r\{\hat{a}_p^s\} + \gamma_q^s\{\hat{a}_p^r\} + \gamma_p^r\gamma_q^s - \gamma_p^s\gamma_q^r. \tag{2.22}$$

Equations (2.21) and (2.22) can be thought of as applications of Wick's theorem [56], according to which an operator string is equal to a sum of products of normal ordered operator strings and contracted operators. Each (binary) contraction here gives rise to an element of $\boldsymbol{\gamma}_1$, which is zero for unoccupied orbitals (indexed by $a, b, \ldots$) and Kronecker's delta for occupied orbitals (indexed by $i, j, \ldots$), $\gamma_i^j = \delta_i^j$. The electronic Hamiltonian, Eq. (2.4), then takes the form

$$\hat{H} = \underbrace{h_{\text{nuc}} + \sum_i h_i^i + \tfrac{1}{2}\sum_{ij} g_{ij}^{ij}}_{E_{\text{HF}}} + \sum_{pq}(\underbrace{h_q^p + \sum_i g_{iq}^{ip}}_{f_q^p})\{\hat{a}_p^q\} + \tfrac{1}{4}\sum_{pqrs} g_{rs}^{pq}\{\hat{a}_{pq}^{rs}\}. \tag{2.23}$$

Because normal ordered operators are defined such that their expectation values with respect to the Fermi vacuum vanish, the scalar part of Eq. (2.23) is the energy expectation value $E_{\mathrm{HF}} = \langle 0|\hat{H}|0\rangle$. It contains the internuclear repulsion energy, the kinetic energy of the electrons, the electron-nuclear interaction, and a mean-field contribution from the electron-electron interaction. The last two terms in Eq. (2.23) now describe only one-particle and two-particle fluctuations around the mean field.

Applying Wick's theorem to Eq. (2.20) leads to the sufficient condition

$$0 = f_i^a = f_a^i, \tag{2.24}$$

which is known as the Brillouin theorem [2].

In canonical Hartree–Fock theory, the orbitals are chosen such that they diagonalize the Fock matrix f_q^p. This fulfills Eq. (2.24) and yields the orbital energies ε_p as diagonal elements. The ground state wave function is then given by the determinant in which the orbitals with the N lowest orbital energies are occupied. However, as f_q^p depends on the occupied orbitals, the evaluation and diagonalization of the Fock matrix must be iteratively repeated until self-consistency. Upon convergence, $|0\rangle$ is the Hartree–Fock determinant and E_{HF} the Hartree–Fock energy. It is worth noting that the Hartree–Fock approximation is equivalent to neglecting the effective two-particle interaction from the Hamiltonian, Eq. (2.23), emphasizing the mean-field nature of the approach.

2.2.2 Coulomb correlation & full configuration interaction

Correlation beyond a single determinant picture is due to the Coulomb interaction between the electrons and hence called *Coulomb correlation*. It implies that statistical independence of orbital occupancies and idempotency of $\boldsymbol{\gamma}_1$ are no longer fulfilled. Instead of Eq. (2.11), one may now write

$$\gamma_{pq}^{rs} = \gamma_p^r \gamma_q^s - \gamma_p^s \gamma_q^r + \lambda_{pq}^{rs}, \tag{2.25}$$

where the *cumulant* λ_{pq}^{rs} is a measure for the correlation between the occupancies of two spin orbitals [54,57], see also sec. 6.2 and Ref. [P2][3] for details.

Within the given orbital basis, Coulomb correlation can be exactly recovered by solving the Schrödinger equation in the full space of N-electron wave functions. This is done in full configuration interaction (for short *full CI*, or FCI), where the wave function is expanded as a linear combination of all N-electron Slater determinants:

$$|\Psi_{\mathrm{FCI}}\rangle = \frac{1}{N!} \sum_{p_1 \cdots p_N} |\Phi_{p_1 \cdots p_N}\rangle c_{p_1 \cdots p_N} = \sum_I |\Phi_I\rangle c_I. \tag{2.26}$$

In Eq. (2.26) we have merged all nonredundant strings of spin orbital indices into a single index I for labeling the determinants.

Insertion of this ansatz into the Schrödinger equation, Eq. (2.3), and projection onto the complete space of determinants yields an eigenvalue equation for the CI coefficients c_I,

$$\sum_J \langle \Phi_I|\hat{H}|\Phi_J\rangle c_J = E c_I. \tag{2.27}$$

---

[3]Throughout this thesis, citations marked with a capital 'P' refer to the list of publications in app. A.3.

FCI offers the exact solution of the electronic Schrödinger equation in a given one-electron basis. However, due to the exponential scaling of the number of determinants with the number of electrons and a high polynomial scaling with respect to the number of basis functions, FCI calculations are only affordable for very small systems and basis sets. This shortcoming, which might only be cured by the advent of the quantum computer [58], has for decades motivated the search for and development of approximate methods.

2.2.3   Static vs. dynamic correlation

For most stable molecules with a closed-shell ground state, a FCI wave function based on Hartree–Fock orbitals shows a large weight for the Hartree–Fock determinant ($c_0^2 \lessapprox 1$). Small contributions from a large number of excited determinants serve to describe the so-called *dynamic correlation*, which is due to the electrons avoiding each other instantaneously when moving through space. The reduced one-particle density matrix with respect to $|\Psi_{\text{FCI}}\rangle$ is then nearly idempotent and has eigenvalues close to one or zero.

A different situation is encountered when several energetically low-lying determinants are nearly degenerate and significantly contribute to the FCI wave function. The type of correlation involved is called *nondynamic* or *static correlation*[4] and is best explained by an example: Figure 1 shows the energetics of the automerization reaction of cyclobutadiene for an idealized reaction path and selected model wave functions. Reactant and product, both being rectangular cyclobutadiene molecules, are connected by a square transition state. Near the equilibrium structures, the Hartree–Fock determinant should
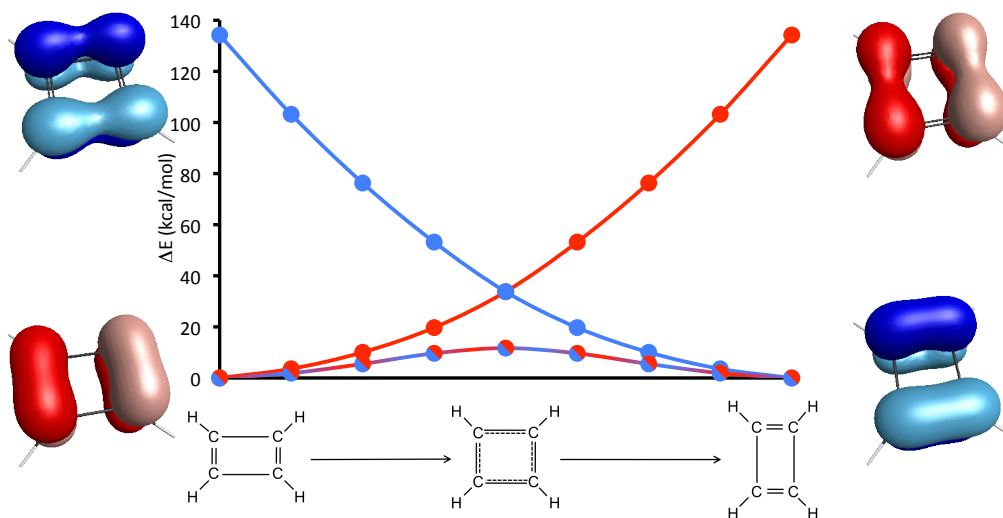
Figure 1: Illustration of static correlation in the automerization of cyclobutadiene.

---

[4]Like most researchers in the field, I will use the terms 'nondynamic' and 'static' interchangeably. Some authors (e.g., Refs. [59,60]) draw a further distinction according to which nondynamic correlation refers to genuine multireference character due to near-degeneracy of several configuration state functions [16], whereas the term static correlation is reserved for open-shell states whose spin-adapted description requires a multideterminantal configuration.

obviously contain the doubly occupied π orbital that stabilizes the shorter C–C double bond. However, as C–C single and double bonds are exchanged in the course of the automerization, this stabilizing (or bonding) orbital correlates to a destabilizing orbital on the other side of the reaction channel (making this reaction *thermally forbidden* by the Woodward–Hoffmann rules [8]). Hartree–Fock theory yields two separate potential energy curves (red and blue curves in Fig. 1), depending on which π orbital is included in the determinant, and is thus not able to describe a smooth transition between reactant and product along the reaction path. Since the two corresponding determinants become degenerate at the square transition state, at least a linear combination of both is required in the wave function ansatz for a qualitatively correct description (indicated in Fig. 1 by a red and blue striped curve).

Before turning to the generalization of this concept, it is worth discussing a possible way to detect the presence of (strong) static correlation in systems less accessible to chemical intuition than the automerization of cyclobutadiene. A general strategy is to perform a quantum chemical calculation capable of at least partially recovering static correlation and then inspect the eigenvalues of the reduced one-particle density matrix. Interestingly, even the UHF method is flexible enough for this purpose due to its ability to break spin symmetry by using different sets of spatial orbitals for α and β spin functions [61–63]. The spin symmetry can be restored by spin projection, turning the UHF determinant into a multideterminantal wave function with a non-idempotent $\boldsymbol{\gamma}_1$. Another way to obtain a spin-adapted characterization of the UHF solution is the definition of the spin-free reduced one-particle density matrix $\boldsymbol{\Gamma}_1$, whose elements refer to a particular set of spatial orbitals (indexed by P, Q),

$$\Gamma_P^Q = \gamma_{P_\alpha}^{Q_\alpha} + \gamma_{P_\beta}^{Q_\beta}. \tag{2.28}$$

The spatial orbitals diagonalizing $\boldsymbol{\Gamma}_1$ are called *natural orbitals* (NOs) and the eigenvalues of $\boldsymbol{\Gamma}_1$ are the *natural orbital occupation numbers*. UHF NO occupation numbers deviating significantly from integer values indicate that the corresponding NOs are involved in static electron correlation. In cases without an appreciable amount of static correlation and an even number of electrons, all UHF NO occupation numbers generally become equal to zero or two, and the UHF and RHF solutions become identical.[5]

2.3 The multiconfigurational self-consistent field method

Multiconfigurational self-consistent field (MCSCF) theory offers a genuine treatment of static correlation based on multideterminantal spin eigenfunctions [3]. In essence, the wave function is written as a CI-like linear combination of (selected) determinants and one variationally optimizes both the CI coefficients and the orbitals. It is convenient to use the same set of spatial orbitals for both α and β spin functions and to subdivide the orbital space into inactive orbitals that are doubly occupied in all determinants (indexed by $i, j, \ldots$), so-called active orbitals with varying occupancies ($u, v, \ldots$), and

---

[5]Although fractional UHF NO occupation numbers are a useful criterion for the occurrence of static correlation, they might not always be reliable. For example, Hollett and Gill [64] have pointed out that UHF is insensitive to near-degeneracies featuring an orbital energy difference that is small relative to the orbital energies but large in absolute terms, as may happen for cationic systems. For alternative indicators of static correlation, the reader is referred to Refs. [65–67].

inactive virtual (or unoccupied) orbitals $(a, b, \ldots)$. The wave function thus consists of determinants formed by creating a number of N_{act} active electrons on top of a core determinant $|0\rangle$ containing all doubly occupied orbitals,

$$|\Psi_0\rangle = \frac{1}{N_{\text{act}}!} \sum_{u_1 \cdots u_{N_{\text{act}}}} \hat{a}_{u_1}^{\dagger} \cdots \hat{a}_{u_{N_{\text{act}}}}^{\dagger} |0\rangle c_{u_1 \cdots u_{N_{\text{act}}}} = \sum_{\mu} |\Phi_\mu\rangle c_\mu. \tag{2.29}$$

If we impose no constraints on the CI coefficients c_μ except the normalization condition $\sum_\mu c_\mu^2 = 1$, the determinants $\{|\Phi_\mu\rangle\}$ cover all possibilities to distribute N_{act} electrons into M_{act} active spin orbitals and hence form a *complete active space* (CAS), also denoted as $\text{CAS}(N_{\text{act}}, \frac{1}{2}M_{\text{act}})$. This special case of MCSCF theory based on a CAS is called complete active space self-consistent field (CASSCF) theory and will be used throughout this work. This choice has the advantage that it keeps the amount of arbitrariness involved in the active space design to a minimum and preserves the formally appealing property of orbital invariance with respect to separate unitary transformations of doubly occupied, active and virtual orbitals, respectively.

Variational optimization of the CI coefficients leads to an eigenvalue equation analogous to Eq. (2.27),

$$\sum_\nu \langle \Phi_\mu | \hat{H} | \Phi_\nu \rangle c_\nu = E_{\text{CASSCF}} \, c_\mu. \tag{2.30}$$

CASSCF may therefore be viewed as a FCI within a complete active space (CAS-CI), yet equipped with an additional optimization of the orbitals.

The latter, as in Hartree–Fock theory, can be parametrized in terms of an anti-Hermitian single excitation operator:

$$|\tilde{\Psi}_0\rangle = e^{-\hat{\kappa}} \sum_\mu |\Phi_\mu\rangle c_\mu = e^{-\hat{\kappa}} |\Psi_0\rangle, \tag{2.31}$$

with

$$\hat{\kappa} = \sum_{iu} \kappa_{iu}(\hat{a}_i^u - \hat{a}_u^i) + \sum_{ua} \kappa_{ua}(\hat{a}_u^a - \hat{a}_a^u) + \sum_{ia} \kappa_{ia}(\hat{a}_i^a - \hat{a}_a^i). \tag{2.32}$$

Due to the invariance of the wave function, Eq. (2.29), with respect to orbital rotations within the inactive orbital spaces (up to a phase factor), single excitations of the types $\hat{a}_i^j$ and $\hat{a}_a^b$ are excluded from $\hat{\kappa}$. Further, orbital rotations within the active space are redundant in CASSCF as they can be fully modeled by the CI coefficients.

In analogy to Hartree–Fock theory, expansion of the energy expectation value

$$\tilde{E} = \langle \Psi_0 | e^{\hat{\kappa}} \hat{H} e^{-\hat{\kappa}} | \Psi_0 \rangle \tag{2.33}$$

around $\boldsymbol{\kappa} = \mathbf{0}$ and differentiation with respect to κ_{pq} yields a variational condition for the orbital rotations,

$$0 = \langle \Psi_0 | [\hat{H}, \hat{a}_p^q] | \Psi_0 \rangle, \tag{2.34}$$

which is known as the *generalized Brillouin theorem* for MCSCF wave functions [68–70].

It should be mentioned that efficiently solving Eqs. (2.30) and (2.34) simultaneously is far from trivial. A plethora of optimization algorithms has been devised for this task, including second-order techniques based on the Newton–Raphson method. For detailed accounts on the optimization procedures in use, we refer to surveys from the literature [3, 22, 24, 25].

2.4 The single-reference coupled-cluster method

In single-reference coupled-cluster (SRCC) theory [9–11, 15], the wave function is represented by an exponential expansion,

$$|\Psi_{\mathrm{SRCC}}\rangle = e^{\hat{T}}|0\rangle. \tag{2.35}$$

The reference function $|0\rangle$ is a single determinant, usually the Hartree–Fock determinant, and the cluster operator

$$\hat{T} = \underbrace{\sum_{ia} t_a^i \hat{a}_i^a}_{\hat{T}_1} + \underbrace{\frac{1}{4}\sum_{ijab} t_{ab}^{ij}\hat{a}_{ij}^{ab}}_{\hat{T}_2} + \ldots + \underbrace{\frac{1}{N_{\mathrm{rank}}!^2}\sum_{i_1\cdots a_1\cdots} t_{a_1\cdots a_{N_{\mathrm{rank}}}}^{i_1\cdots i_{N_{\mathrm{rank}}}}\hat{a}_{i_1\cdots i_{N_{\mathrm{rank}}}}^{a_1\cdots a_{N_{\mathrm{rank}}}}}_{\hat{T}_{N_{\mathrm{rank}}}} = \sum_\rho \hat{\tau}^\rho t_\rho = \hat{\boldsymbol{\tau}}\mathbf{t} \tag{2.36}$$

contains cluster amplitudes t_ρ associated with excitations from M_{occ} occupied orbitals to M_{virt} virtual orbitals. Equation (2.36) makes use of a shorthand notation in which $\hat{\boldsymbol{\tau}}$ is a row vector of excitation operators and $\mathbf{t}$ is a column vector of cluster amplitudes. If up to N-body excitations are included in $\hat{T}$ (i.e., $N_{\mathrm{rank}} = N$), the excited functions $\hat{\tau}^\rho|0\rangle$ span the complete space of N-electron wave functions (the FCI space) and the SRCC method becomes exact. In practice, if $|0\rangle$ is qualitatively close to the exact wave function, it suffices to truncate $\hat{T}$ after double or triple excitations for an accurate description of dynamic correlation. The corresponding methods are called coupled-cluster with singles and doubles (CCSD) and coupled-cluster with singles, doubles and triples (CCSDT), respectively.

Coupled-cluster theory is traditionally formulated by means of a non-variational projection technique. In order to derive equations for the cluster amplitudes, one first inserts the ansatz (2.35) into the Schrödinger equation and pre-multiplies it with the inverse of the wave operator, $e^{-\hat{T}}$, yielding

$$\underbrace{e^{-\hat{T}}\hat{H}e^{\hat{T}}}_{\bar{H}}|0\rangle = E_{\mathrm{SRCC}}|0\rangle, \tag{2.37}$$

where we have defined the similarity transformed Hamiltonian $\bar{H}$. The energy expression then follows from projecting Eq. (2.37) onto the reference determinant,

$$E_{\mathrm{SRCC}} = \langle 0|\bar{H}|0\rangle, \tag{2.38}$$

whereas equations for the cluster amplitudes result from projection onto the set of excited determinants $\hat{\tau}^\rho|0\rangle$,

$$\mathbf{0} = \langle 0|\hat{\boldsymbol{\tau}}^\dagger \bar{H}|0\rangle. \tag{2.39}$$

The computational expense of SRCC methods is dominated by the evaluation of the term $\frac{1}{2}\sum_{ab} g_{cd}^{ab} t_{ab\cdots}^{ij\cdots}$, which scales like $M_{\mathrm{occ}}^{N_{\mathrm{rank}}} M_{\mathrm{virt}}^{N_{\mathrm{rank}}+2}$. The same scaling behavior is true for the related configuration interaction based methods which use a linear wave function ansatz. However, due to the exponential wave operator, truncated coupled-cluster approaches have the advantage to yield size-extensive energies and to ensure the correct separability of both energy and wave function for noninteracting systems (size-consistency). We refer to sec. 3.3.3 for a more refined discussion of these properties. Coupled-cluster calculations generally provide significantly higher accuracy than corresponding CI calculations, rendering coupled-cluster theory the method of choice.

2.5   Overview of multireference coupled-cluster theories

SRCC theory offers a systematic hierarchy of approximations converging towards FCI. It excels at capturing dynamic correlation even when including only low excitation ranks, but in systems with strong static correlation much higher clusters are needed for an accurate description [3]. CASSCF has the converse strengths and weaknesses: While it may provide a reasonable treatment of static correlation even with few active orbitals, an accurate account of dynamic correlation within CASSCF requires exceedingly large active spaces [23]. Figure 2 illustrates how one may regard both methods as possible routes towards the exact solution with differing emphasis on the type of electron correlation recovered. As both methods involve a rapidly rising computational scaling along their hierarchies, neither allows an efficient and accurate description of systems with both dynamic and strong static correlation. It is therefore desirable to find a 'shortcut' towards the exact solution by merging both approaches into a multireference coupled-cluster (MRCC) theory.

The related multireference configuration interaction (MRCI) method is straightfor-wardly defined by selecting a set of active orbitals and employing all N-electron deter-minants in the CI that have up to N_{rank} *holes* (electrons missing) in the doubly occupied orbital space and up to N_{rank} *particles* (electrons) in the virtual orbital space. Espe-cially MRCI methods with approximate corrections for size-inextensivity are accurate for small systems and frequently in use [25]. However, despite decades of large efforts in the development of MRCC methods [P6,27–29], the 'perfect marriage' of the exponential cluster ansatz and a multideterminantal reference function has not been achieved yet. Most of the existing MRCC methods to be discussed in the following therefore present a compromise between formal rigorosity and practical efficiency.

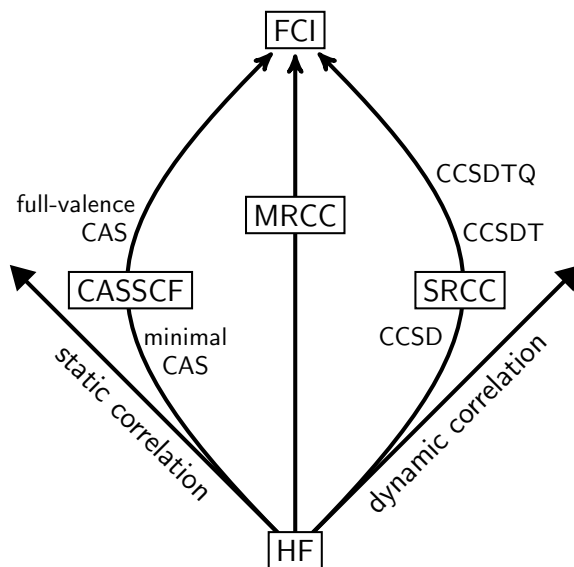Numerous methods attempt to describe static correlation within a single-reference

Figure 2: Motivation of MRCC theory as a 'shortcut' towards FCI.

framework by using multireference concepts such as subdividing the occupied and virtual orbital spaces into an inactive and an active part. The most simplistic approach is embodied by the tailored coupled-cluster method [71], which fixes certain cluster amplitudes to values derived from a CAS-CI wave function and solves only for the remaining amplitudes. The related reduced multireference CCSD (RMRCCSD) method [72] goes further by admitting in the SRCC equations additional triples and quadruples amplitudes that are predefined by an MRCI calculation. Active-space coupled-cluster methods [73] offer an increased flexibility by determining these additional amplitudes through the coupled-cluster formalism. The CASCC method [74, 75], for instance, extends the cluster operator to selected higher excitations in such a way that the resulting space of excited determinants is the same as in MRCI. The maximum excitation level in the CASCC wave function then grows with the size of the active space, leading to an increasing complexity of the method that can only be dealt with using general-order implementations of SRCC theory [76]. Instead of basing the selection of included higher excitations on a formal active space, it can also be guided by a special importance-selection function to achieve an automatic adaptation of the theoretical model to the studied problem [77]. As an alternative to including additional parameters, it is possible to estimate the effect of higher excitations from the so-called moments of the coupled-cluster equations, i.e., the projections of Eq. (2.37) onto higher excited determinants. This is the basis for the renormalized and completely renormalized coupled-cluster approaches [78, 79], which can also be combined with active-space coupled-cluster theory [80].

One disadvantage of the methods described so far is that they introduce a bias towards the reference function and therefore lack orbital invariance. In addition, an accurate treatment of static correlation often goes along with a large number of parameters in these models. Equation-of-motion (EOM) coupled-cluster theory offers a way to circumvent these problems without leaving the single-reference framework: Starting from a correlated description of a related state that may have a different spin symmetry or a different number of electrons but is well represented by a single determinant, one can target the state of interest via excitation operators involving spin-flip [81, 82] or multi-ionization [83]. However, the need to choose orbitals suited for capturing both the dynamic correlation of the initial system and the differential dynamic and static correlation of the target system limits the accuracy of EOM-based approaches. These limitations may be expected to become even more severe in future extensions of these methods with spin-flip or ionization operators acting on more than two electrons.

Most MRCC methods built upon a multideterminantal reference function rely on either of two strategies: The first one employs a separate cluster expansion for each reference determinant as initially proposed by Jeziorski and Monkhorst [84],

$$|\Psi_{\text{JM}}\rangle = \sum_{\mu} e^{\hat{T}^{\mu}} |\Phi_{\mu}\rangle c_{\mu}. \qquad (2.40)$$

The original state-universal formulation [27, 84–86] simultaneously targets all the states that can be modeled by the active space. As this may lead to numerical problems whenever a low excited configuration involving virtual orbitals becomes near-degenerate with a configuration from the active space (*intruder state problem*), various state-specific MRCC methods targeting only one state have been devised. Most of these approaches [87], comprising the so-called Brillouin–Wigner coupled-cluster (BW-MRCC) theory [88,

89], the state-specific Mk-MRCC method introduced by the Mukherjee group [90–99], and the single-root MRCC theory [100], impose additional restrictions (*sufficiency conditions*) on the coupled-cluster equations in order to match the number of equations with the number of parameters. A way to instead reduce the number of amplitudes to the number of equations is explored in the MRexpT method [101, 102].

Among these methods, Mk-MRCC theory has become the most popular approach, owing to its size-extensivity and a relatively simple structure that can be implemented by extending existing SRCC codes. However, the lack of orbital invariance of the Jeziorski–Monkhorst ansatz [33] introduces an undesired arbitrariness in these approaches and Mk-MRCC results may be very sensitive to the choice of orbitals [91, 103]. Further, the computational cost of an Mk-MRCC calculation scales factorially with the size of the active space, as each reference determinant has its own cluster operator. In order to alleviate the latter shortcoming, one may treat inactive excitations as independent from the reference functions (*anonymous parentage approximation*) and move the exponential of this part of the cluster operator in front of the summation in Eq. (2.40) [104–106]. This corresponds to an *internal contraction* of the inactive excitations and is inspired by the highly successful internally contracted approximation of MRCI theory [48, 107].

The second strategy is a fully internally contracted ansatz, in which a single cluster expansion acts onto the entire multideterminantal reference,

$$|\Psi_{\mathrm{ic}}\rangle = e^{\hat{T}}|\Psi_0\rangle = e^{\hat{T}} \sum_\mu |\Phi_\mu\rangle c_\mu. \tag{2.41}$$

This ansatz is realized in the valence-universal Fock-space coupled-cluster theory, which targets all states accessible by distributing any number of electrons in the active space, including ionized and electron-attached states [108–110]. The desire to avoid numerical instabilities from intruder states and the interest in accurate descriptions of individual states has motivated the formulation of various state-specific internally contracted multireference approaches. The internally contracted MRCC (ic-MRCC) theory studied in this thesis is rooted in the work of Banerjee and Simons [34–37] and has recently been picked up by Evangelista and Gauss [38]. Its efficient implementation has been hindered for decades by its complicated formal structure due to noncommuting excitation operators in $\hat{T}$. The same is true for a variant using a *generalized normal ordering* [111–113] with respect to the multideterminantal reference function, as proposed by Mukherjee [30, 31], for which not even pilot implementations have emerged prior to the present work.

A number of alternative internally contracted approaches exist: Nooijen and co-workers [32, 51, 114, 115] have put forth hybrid methods, designed to simultaneously treat a target state and a number of excited states by diagonalizing a similarity transformed Hamiltonian in an MRCI function space, restricted to singly excited configurations. The cluster amplitudes needed for the similarity transformation are either obtained from *projected equations* as in ic-MRCC theory, or from the simpler *many-body equations*, which consist in requiring certain elements of the similarity transformed Hamiltonian (i.e., those corresponding to excitation operators) to vanish. The canonical transformation theory by Yanai and Chan [40–42] is closely related to ic-MRCC theory, being a unitary variant, equipped with approximations of intermediate operators with ranks beyond two (cf. Ref. [43]). The unitary wave operator may also be combined with a variational approach [46, 47, 116]. Yet another alternative internally contracted scheme is block

correlated coupled-cluster theory [117, 118], where the purely active parts of excitations are represented in terms of transitions from the ground state within the active space into states with the same or a different number of active electrons.

All of these existing internally contracted schemes have important advantages in common, including orbital invariance and a compact parameterization of the wave function. However, most of them are not fully size-extensive in their present form,[6] and in view of the compromises made in their formulations, they do not seem to tap the full potential of ic-MRCC theory. The following chapters aim to close this gap by offering a rigorous account of ic-MRCC theory, shedding light on unsettled issues such as its formal complexity, computational scaling and size-extensivity, and by unveiling its predictive power through benchmark applications.

---

[6]Although CT theory is expected to be size-extensive, a statement in Ref. [42] indicates that at least its present implementation is not.

3 Internally contracted multireference coupled-cluster (ic-MRCC) theory

3.1 Basic ic-MRCC theory

3.1.1 The ic-MRCC ansatz and equations

In the ic-MRCC ansatz, Eq. (2.41), the exponential wave operator acts onto a CASSCF-like reference function, Eq. (2.29). The cluster operator therefore should cover excitations from the combined set of occupied and active orbitals (indexed by $I, J, \ldots$) to the combined set of active and virtual orbitals ($A, B, \ldots$),

$$\hat{T} = \sum_{IA} t_A^I \hat{a}_I^A + \tfrac{1}{4} \sum_{IJAB} t_{AB}^{IJ} \hat{a}_{IJ}^{AB} + \ldots = \sum_\rho \hat{\tau}^\rho t_\rho = \hat{\boldsymbol{\tau}} \mathbf{t}. \tag{3.1}$$

In analogy to single-reference coupled-cluster theory, we first insert the ic-MRCC ansatz into the Schrödinger equation and pre-multiply with $e^{-\hat{T}}$ to obtain an equation in terms of a similarity transformed Hamiltonian ($\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$),

$$\bar{H} |\Psi_0\rangle = E |\Psi_0\rangle. \tag{3.2}$$

Projection onto the reference function straightforwardly yields an expression for the ic-MRCC energy,

$$E = \langle \Psi_0 | \bar{H} | \Psi_0 \rangle, \tag{3.3}$$

while projection onto the excited functions $\hat{\tau}^\rho |\Psi_0\rangle$ initially results in

$$\langle \Psi_0 | \hat{\boldsymbol{\tau}}^\dagger \bar{H} | \Psi_0 \rangle = E \langle \Psi_0 | \hat{\boldsymbol{\tau}}^\dagger | \Psi_0 \rangle. \tag{3.4}$$

The right-hand side of Eq. (3.4) may be nonzero for excitations within the active space, i.e., for projections onto $\hat{a}_{wx\cdots}^{uv\cdots} |\Psi_0\rangle$. However, instead of considering purely active excitations, it is easier to directly project Eq. (3.2) onto the reference determinants. The resulting equations take the form of an eigensystem for the CI coefficients in the reference function (for short *reference coefficients*),

$$\sum_\nu \langle \Phi_\mu | \bar{H} | \Phi_\nu \rangle c_\nu = E c_\mu. \tag{3.5}$$

It becomes obvious that projections onto the reference space do not furnish enough equations to uniquely determine both the purely active amplitudes $t_{uv\cdots}^{wx\cdots}$ and the reference coefficients c_μ. We therefore consider the former as redundant parameters and set them to zero. Equation (3.5) is then used to relax the reference function.[7]

Projections onto reference determinants are not the only redundant part of Eq. (3.4). More redundancies arise due to linear dependencies between the excited functions $\hat{\tau}^\rho |\Psi_0\rangle$. Section 3.2 is devoted to a detailed discussion of that issue. For the present context, it suffices to point out that one can define a new set of excitation operators $\hat{\tau}'^\rho$ as linear combinations according to

$$\hat{\boldsymbol{\tau}}' = \hat{\boldsymbol{\tau}} \mathbf{X}, \tag{3.6}$$

---

[7]Alternatively we could fix the CI coefficients to their CASSCF values (which corresponds to solving Eq. (3.5) for $\hat{T} = 0$), and account for the relaxation of the reference function by a nonredundant set of purely active excitations. A comparison of these choices is presented in sec. 4.5.

such that the (rectangular) matrix $\mathbf{X}$ transforms the excited functions $\hat{\tau}^\rho|\Psi_0\rangle$ into a set of linearly independent functions $\hat{\tau}'^\rho|\Psi_0\rangle$. The cluster operator is then restricted to these new excitations,

$$\hat{T} = \hat{\boldsymbol{\tau}}'\mathbf{t}', \tag{3.7}$$

and Eq. (3.4) becomes

$$\mathbf{0} = \langle\Psi_0|\hat{\boldsymbol{\tau}}'^\dagger\bar{H}|\Psi_0\rangle. \tag{3.8}$$

The ic-MRCC solution simultaneously fulfills both the eigenvalue equations (3.5) and the amplitude equations (3.8) [35, 38].

Although different ways of removing redundant parameters from the original cluster operator given in Eq. (3.1) generally lead to different ic-MRCC wave functions and energies, all of these ic-MRCC methods properly converge to FCI when the maximum excitation rank N_{rank} in the cluster operator reaches the number of (correlated) electrons. A rationale for this property is provided in sec. 3.3.2 (see Ref. [P5] for a proof).

The discussion of the ic-MRCC theory's structure in the following sections is largely based on Ref. [P1].

3.1.2   Formal truncation & commutator approximations

Applying the BCH expansion, Eq. (2.19), to the similarity transformed Hamiltonian,

$$\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}} = \hat{H} + [\hat{H}, \hat{T}] + \tfrac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \dots, \tag{3.9}$$

one may write the ic-MRCC equations (3.5) and (3.8) as series of terms with increasing number of cluster operators. In SRCC theory the excitations in $\hat{T}$ mutually commute and hence each cluster operator in the nested commutator $[\cdots[\hat{H}, \hat{T}], \hat{T}\cdots]$ must be contracted with the Hamiltonian. Since $\hat{H}$ contains only up to twofold excitations, the expansion in Eq. (3.9) truncates at the quartic power in $\hat{T}$. Unfortunately, this is no longer the case in ic-MRCC theory where excitation operators involving creation and annihilation of active electrons do not commute with each other.

Figure 3a displays all possible cluster excitations for ic-MRCC theory with singles and doubles (ic-MRCCSD) as antisymmetrized Brandow-type diagrams [119]. The operators are divided into sets of excitations with common number of inactive hole (n_h) and
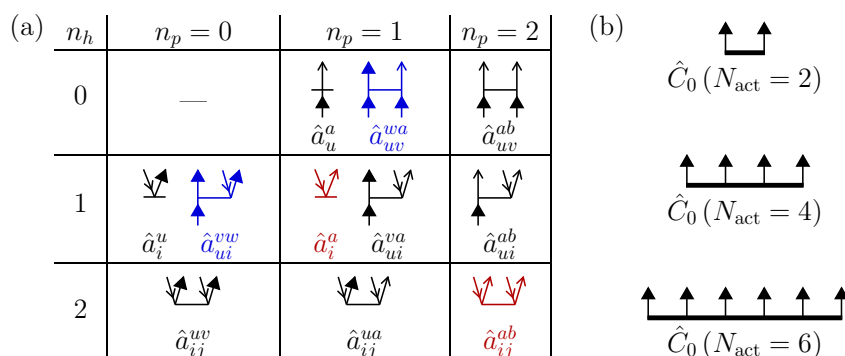

Figure 3: Diagrammatic representation of (a) the excitations in $\hat{T}$ for $N_{\text{rank}} = 2$, with excitations present in SRCCSD drawn in red, and (b) the reference function for $N_{\text{act}} = \{2, 4, 6\}$.

particle (n_p) indices, to be denoted as *excitation classes.* Since we choose the core determinant $|0\rangle$ as Fermi vacuum, indices labeling active orbitals (*valence* lines) are depicted as arrows pointing upwards and distinguished from genuine particle lines by the use of filled triangular arrow heads.

The expansion in Eq. (3.9) does not naturally truncate in ic-MRCC theory, because in principle it may contain infinite chains of contracted excitation operators with two or more valence lines each. However, due to the exclusion of purely active cluster amplitudes, each cluster operator contributes at least one hole or particle creation line to $\bar{H}$. Since in the ic-MRCC equations all of these inactive lines must be connected with hole or particle annihilation lines, coming either from the Hamiltonian or from the projection onto excited functions, the number of cluster operators in a term cannot exceed the total number of annihilation and creation operators in $\hat{H}$ and $(\hat{\tau}'^{\rho})^{\dagger}$. Therefore the eigenvalue equations (3.5) truncate at quartic power in $\hat{T}$, while the number of cluster operators N_{com} in the amplitude equations (3.8) is bounded by the maximum excitation rank according to

$$N_{\text{com}} \leq 4 + 2N_{\text{rank}}. \tag{3.10}$$

The full ic-MRCCSD equations hence involve up to eightfold commutators, see Fig. 4b for an example of such a term (in diagrammatic representation). In a full implementation of ic-MRCC theory, terms with such a large number of cluster operators would have detrimental impact on the efficiency of the method, both due to their sheer number (see also Tab. 1 and 2) and their computational expense (see sec. 3.1.3). Fortunately, the use of a qualitatively correct reference function guarantees that the cluster amplitudes need to describe dynamic correlation only and consequently will stay small [35]. Terms of high power in $\hat{T}$ may then be neglected, rendering *commutator approximations* (truncations of the BCH expansion) efficient approximations in ic-MRCC theory [38], as will be confirmed in sec. 4.2.

Another important aspect is the formal truncation with the number of active electrons. Due to our choice of Fermi vacuum, the reference function explicitly appears in the ic-MRCC equations in form of an additional operator,

$$\hat{C}_0 = \frac{1}{N_{\text{act}}!} \sum_{u_1 \cdots u_{N_{\text{act}}}} \hat{a}_{u_1}^{\dagger} \cdots \hat{a}_{u_{N_{\text{act}}}}^{\dagger} c_{u_1 \cdots u_{N_{\text{act}}}}, \tag{3.11}$$

which distributes N_{act} electrons into the active orbitals, $|\Psi_0\rangle = \hat{C}_0|0\rangle$, see Fig. 3b for a diagrammatic representation. In order to identify similar terms for different numbers of
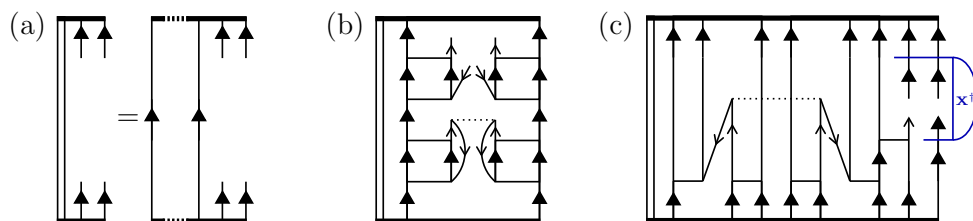

Figure 4: Diagrams showing (a) the reduced two-particle density matrix, (b) a sample eightfold commutator term and (c) a term involving the eight-particle density matrix.

active electrons, we may substitute reduced density matrices $\boldsymbol{\gamma}_n$ for (contracted) products of $\hat{C}_0$ and its adjoint. Diagrammatically these appear as operators with two *vertices* (horizontal lines) as shown in Fig. 4a, one at the bottom and one at the top of a diagram (see Figs. 4b and 4c for examples).

The rank of the reduced density matrices is bounded by the number of active electrons,

$$n \leq N_{\text{act}}. \tag{3.12}$$

Irrespective of the number of active electrons, n is further restricted by the truncation scheme applied to the cluster operator and the commutator series. This comes from the fact that a given number of cluster operators can only yield intermediates with a finite number of uncontracted valence lines. Considering that these lines are eventually connected to the $2n$ indices of a $\boldsymbol{\gamma}_n$ leads to the following upper bound in the ic-MRCC energy expression:[8]

$$n \leq N_{\text{com}}(N_{\text{rank}} - 1) + 2. \tag{3.13}$$

In the amplitude equations, the operator $(\hat{\tau}'^\rho)^\dagger$ provides additional lines, denoted as *open* lines, and allows a larger number of $\hat{T}$ operators in a term. However, even cluster operators with open inactive lines must be connected to the Hamiltonian either directly or indirectly through other cluster operators. From these considerations, finally the following upper bounds result:

$$n \leq \begin{cases} (N_{\text{com}} + 1)(N_{\text{rank}} - 1) + 2, & \text{if } N_{\text{com}} \leq 5, \\ (N_{\text{com}} + 1)(N_{\text{rank}} - 2) + 8, & \text{if } N_{\text{com}} \geq 5. \end{cases} \tag{3.14}$$

As an example, when the cluster operator is truncated at double excitations, terms including up to $\boldsymbol{\gamma}_8$ may arise, such as the diagram shown in Fig. 4c.[9]

Table 1 lists the numbers of topologically unique terms arising in ic-MRCCSD for a selected range of powers in $\hat{T}$ and ranks of reduced density matrices (where $n = 0$ characterizes terms without a reduced density matrix), based on an automated derivation of the equations (see app. A.1.4 for details). Although the drastic increase in the number of terms compared to single-reference CCSD can be partially attributed to the mere existence of an additional active orbital space (as may be seen by comparing the number of terms with $N_{\text{com}} \leq 1$), the major amount of terms is due to the variety of possible connections among cluster operators.

Regarding the large number of terms generally arising in ic-MRCC theory, it is worth discussing two special cases which result in major simplifications of the working equations: The first is the absence of correlated doubly occupied orbitals. When all occupied orbitals are either included in the active space or treated as an uncorrelated frozen core, the cluster operator for ic-MRCCSD reduces to the three excitations shown in Fig. 3a for $n_h = 0$. The number of unique terms for this scheme is collected in Tab. 2. Indeed, this strategy was followed in the original work of Banerjee and Simons [34].[10] However, this simplification is only practical for the smallest systems, as systems with many electrons would either require very large active spaces or drastic frozen core approximations.

---

[8]The eigenvalue equations, on the other hand, do not contain reduced density matrices since the adjoint of $\hat{C}_0$ does not occur.

[9]The operator drawn in blue in Fig. 4c indicates that the ic-MRCC equations are projections onto $\langle\Psi_0|\hat{\tau}'^\dagger = \mathbf{X}^\dagger\langle\Psi_0|\hat{\tau}$, see also app. A.1.4.

[10]These authors went even further by additionally neglecting the semi-internal excitations associated with $\hat{a}_{uv}^{wa}$, which is (in my experience) a very crude approximation for active spaces beyond CAS(2,2).

Table 1: Number of topologically unique terms in energy expression and amplitude equations for CCSD and ic-MRCCSD, grouped by the number of cluster operators per term (N_{com}) and the rank of the reduced density matrix (n).

N_{com}	CCSD	ic-MRCCSD (n)							
		0	1	2	3	4	5	6	...
Energy expression / Eigenvalue equations									
0	1	1	1	1	0	0	0	0	0
1	2	5	11	9	2	0	0	0	0
2	1	11	39	44	16	3	0	0	0
3	0	10	58	96	55	15	2	0	0
4	0	4	32	79	67	28	6	1	0
Amplitude equations									
0	2	5	11	9	2	0	0	0	0
1	13	73	255	255	95	12	0	0	0
2	21	542	2 664	3 452	1 713	367	28	0	0
3	8	2 518	16 822	27 932	17 899	5 119	673	34	0
4	1	7 845	70 155	148 521	120 507	44 384	7 763	664	...
...	0	...	...	...	...	...	...	...	...

Table 2: Number of topologically unique terms in energy expression and amplitude equations for special cases of ic-MRCCSD, grouped by the number of cluster operators per term (N_{com}) and the rank of the reduced density matrix (n).

N_{com}	ic-MRCCSD with $M_{\mathrm{occ}} = 0$ (n)							ic-MRCCSD without $\hat{a}^{wa}_{uv}$ and $\hat{a}^{vw}_{ui}$ (n)				
	0	1	2	3	4	5	6	0	1	2	3	4
Energy expression / Eigenvalue equations												
0	1	1	1	0	0	0	0	**1**	**1**	**1**	0	0
1	0	1	4	1	0	0	0	**5**	**9**	**5**	0	0
2	0	0	2	2	1	0	0	**9**	**18**	**9**	0	0
3	0	0	0	0	0	0	0	**5**	**9**	**5**	0	0
4	0	0	0	0	0	0	0	**1**	**1**	**1**	0	0
Amplitude equations												
0	0	1	4	1	0	0	0	**5**	**9**	**5**	0	0
1	0	2	28	24	4	0	0	**71**	**191**	**122**	20	0
2	0	1	44	100	56	6	0	**442**	**1 420**	**997**	200	10
3	0	0	22	102	141	49	3	**1 313**	**4 420**	**3 029**	593	32
4	0	0	4	30	75	57	13	**1 825**	**5 689**	**3 401**	536	21
5	0	0	0	0	0	0	0	**1 120**	**2 795**	**1 268**	120	0
6	0	0	0	0	0	0	0	**277**	**468**	**132**	0	0
7	0	0	0	0	0	0	0	**25**	**22**	**0**	0	0
8	0	0	0	0	0	0	0	**1**	**0**	**0**	0	0

The second notable special case is ic-MRCC theory based on a CAS(2,2). Here, single excitations $\hat{a}^a_u$ are sufficient for generating the complete set of excited determinants that have exactly one electron in the virtual orbital space. Double excitations of the type $\hat{a}^{wa}_{uv}$ acting on $|\Psi_0\rangle$ can only create determinants from the very same set of functions. One may therefore consider the amplitudes t^{uv}_{wa} as redundant and exclude them. The same argument justifies the exclusion of t^{ui}_{vw} in favor of t^i_u. Section 3.2.1 describes a general procedure that naturally takes advantage of this type of redundancy. As the excitations $\hat{a}^{wa}_{uv}$ and $\hat{a}^{vw}_{ui}$ (drawn in blue in Fig. 3a) trigger the largest number of terms in ic-MRCCSD (e.g., 143 430 eightfold commutator terms containing exclusively these two excitations), considerable simplifications result from their absence. Table 2 lists the total number of terms for this case, with terms arising for a CAS(2,2) printed in boldface.

3.1.3 Computational scaling

The algorithm to solve the ic-MRCC equations, detailed in app. A.1.4, involves as rate determining step the evaluation of the residual

$$\boldsymbol{\Omega} = \langle \Psi_0 | \hat{\boldsymbol{\tau}}^\dagger \bar{H} | \Psi_0 \rangle. \tag{3.15}$$

In SRCC theory, it is straightforward to characterize the scaling of the computational cost for this step with increasing system size, because typically $M_{\mathrm{virt}} \gg M_{\mathrm{occ}}$ and the dominating term with the highest scaling with respect to M_{virt} also has the highest overall computational scaling (with respect to both M_{virt} and M_{occ}). Matters are more complicated in ic-MRCC theory, as terms scaling polynomially with a high power in M_{act} may eventually become rate determining for large active spaces. Whereas the number of inactive orbitals always grows with a system's size or the basis set size, the number of active orbitals depends on the multireference character of the studied problem and does not generally increase with a system's size. Therefore, the following analysis starts by considering the scenario $M_{\mathrm{virt}} \gg M_{\mathrm{act}}$, such that terms with the highest scaling with respect to M_{virt} are dominating, and then focuses on the impact of the active space's size on the scaling behavior.

Restricting ourselves to the case $N_{\mathrm{rank}} = 2$, the most expensive terms in Eq. (3.15) scale like M^4_{virt}. The only terms for which a contraction scaling like M^4_{virt} is mandatory contain the Hamiltonian vertex $g^{ab}_{cd} \hat{a}^{cd}_{ab}$. This is because in ic-MRCCSD there can be at most four particle lines per term, two of them open and two connecting the Hamiltonian with $\hat{T}$ operators. Terms of this structure that contain Hamiltonian elements with less than four particle indices can always be evaluated starting with an M^3_{virt} contraction through a particle line. The three possible structures of terms scaling like M^4_{virt} are depicted in Fig. 5. The first of these terms (Fig. 5a) is known from single-reference CCSD, where it is equal to $\sum_{cd} g^{cd}_{ab}(\frac{1}{2} t^{ij}_{cd} + t^i_c t^j_d)$. In ic-MRCCSD, a much larger number of terms has this structure. However, all of these may be merged by defining an intermediate $\mathcal{I}^{IJ}_{ab}$ (depicted as a box-shaped vertex), which is then contracted with the Hamiltonian in a single $M^2_{\mathrm{occ}} M^4_{\mathrm{virt}}$ step. The other two terms shown in Figs. 5b and 5c are analogs arising in ic-MRCCSD theory and scale like $M_{\mathrm{occ}} M_{\mathrm{act}} M^4_{\mathrm{virt}}$ and $M^2_{\mathrm{act}} M^4_{\mathrm{virt}}$, respectively. Therefore, each evaluation of the coupled-cluster residual involves only three M^4_{virt} contractions with an approximate combined scaling of $(M_{\mathrm{occ}} + M_{\mathrm{act}})^2 M^4_{\mathrm{virt}}$. This is comparable to single-reference CCSD. However, it stands in contrast to MRCC approaches that are built on

(a) $\frac{1}{2}\sum_{cd}\mathcal{I}_{cd}^{ij}g_{ab}^{cd}$ (b) $\frac{1}{2}\sum_{cd}\mathcal{I}_{cd}^{u,i}g_{ab}^{cd}$ (c) $\frac{1}{2}\sum_{cd}\mathcal{I}_{cd}^{uv}g_{ab}^{cd}$

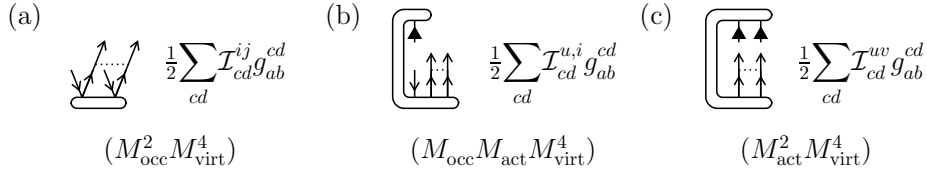$(M_{\text{occ}}^2 M_{\text{virt}}^4)$ $(M_{\text{occ}} M_{\text{act}} M_{\text{virt}}^4)$ $(M_{\text{act}}^2 M_{\text{virt}}^4)$

Figure 5: Terms in ic-MRCCSD scaling formally like M_{virt}^4.

the Jeziorski–Monkhorst ansatz, where the number of M_{virt}^4 contractions scales linearly with the number of reference determinants, d. The latter grows exponentially with the active space's size, as the number of possibilities to distribute N_{act} electrons into M_{act} spin orbitals is given by

$$\binom{M_{\text{act}}}{N_{\text{act}}}. \tag{3.16}$$

The binomial coefficient in Eq. (3.16) may serve as an upper bound for d, which is generally somewhat smaller due to spin and spatial symmetry constraints.

While the computational cost of the M_{virt}^4 terms in ic-MRCCSD theory is almost independent of the active space's size, the scaling with respect to the active space tends to become steeper when moving to terms that scale with lower powers in M_{virt}. Even when the theory is truncated at $N_{\text{com}} = 2$, two terms arise with contractions scaling like $M_{\text{occ}}^3 M_{\text{act}} M_{\text{virt}}^3$, which is higher than the $M_{\text{occ}}^3 M_{\text{virt}}^3$ scaling of the second most expensive contractions from SRCC theory. Figure 6 illustrates these two terms and how they can be merged by factoring out a suitable intermediate involving $\boldsymbol{\gamma}_1$ and a unit tensor, leading to a total of two contractions scaling like $M_{\text{occ}}^3 M_{\text{act}} M_{\text{virt}}^3$. Summation over repeated indices is implied in Fig. 6 (Einstein summation convention). It should be noted that an alternative contraction sequence (namely ①③②) exists with a scaling like $M_{\text{occ}} M_{\text{act}} M_{\text{virt}}^4$. Whether this route is more efficient depends on the actual sizes of the orbital spaces.

The number of active electrons impacts the formal scaling of ic-MRCC theory in two ways: First, an increase in the number of active electrons can lead to a rising number of terms. For example, the term in Fig. 7a only appears for $N_{\text{act}} \geq 4$. Second, even terms already present for a lower number of active electrons may require higher scaling contractions. The diagram in Fig. 7b is for instance identical to the one shown in Fig. 7a except for the number of active electrons involved. While the term in Fig. 7b is formally
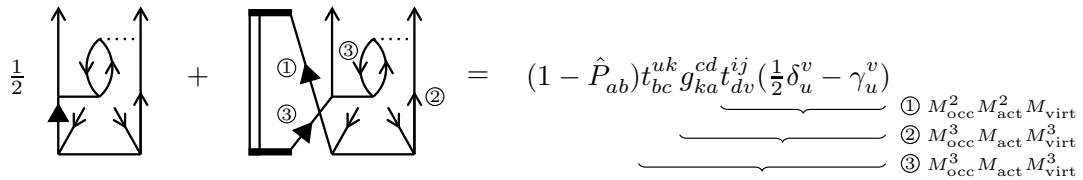

$$\frac{1}{2} \quad + \quad = \quad (1-\hat{P}_{ab})t_{bc}^{uk}g_{ka}^{cd}t_{dv}^{ij}(\tfrac{1}{2}\delta_u^v - \gamma_u^v)$$

① $M_{\text{occ}}^2 M_{\text{act}}^2 M_{\text{virt}}$
② $M_{\text{occ}}^3 M_{\text{act}} M_{\text{virt}}^3$
③ $M_{\text{occ}}^3 M_{\text{act}} M_{\text{virt}}^3$

Figure 6: Terms in ic-MRCCSD scaling formally like $M_{\text{occ}}^3 M_{\text{act}} M_{\text{virt}}^3$, shown as diagrams and in a tensor notation. The permutation operator $\hat{P}_{ab}$ interchanges indices a and b. Numbers in circles indicate the contraction sequence, braces highlight the tensors involved in each contraction.

25

best dealt with starting from γ_4, in the evaluation of the term in Fig. 7a it is advantageous to contract the operators $\hat{C}_0$ and $\hat{C}_0^\dagger$ with other operators first. This allows steps with a reduced scaling ($M_{\mathrm{act}}^4 M_{\mathrm{virt}}^2$ instead of $M_{\mathrm{act}}^8 M_{\mathrm{virt}}^2$) for the smaller active space.[11]

The scaling with respect to M_{act} generally rises with the number of cluster operators in a term, as complicated networks of connections between operators may prevent a reduction of the operator rank for several consecutive contractions. Figure 7c shows a sample fourfold commutator term from the energy expression. Its evaluation involves contractions with sums running over up to five active and three inactive indices. On the other hand, the scaling with respect to the number of inactive orbitals generally decreases with the number of cluster operators, because connections between $\hat{T}$ operators are only possible through active indices. For instance, one may easily deduce from the number of inactive lines in the intermediates $\mathcal{I}_{ab}^{IJ}$ (Fig. 5) that terms with more than four $\hat{T}$ operators must scale lower with M_{virt} than M_{virt}^4.

Especially when moving towards large active spaces, commutator approximations become indispensable for a practically useful implementation of ic-MRCC theory, in order to avoid huge numbers of terms scaling steeply with M_{act}. Further, at some point it will become advantageous to factor out the reduced density matrices from the ic-MRCC amplitude equations, thereby making the residual expression independent of N_{act}.

Figure 8 gives an overview of the computational scaling for $N_{\mathrm{com}} \leq 2$ with respect to all orbital spaces. The presented information is based on an automated factorization of the residual expression (see app. A.1.2), followed by a manual analysis of those terms that scale highest with M_{act} for a given scaling behavior with respect to inactive orbitals. The steepest scaling with M_{act}, namely $M_{\mathrm{act}}^{10} M_{\mathrm{virt}}$, is found for contractions of γ_5 (the highest reduced density matrix arising in that case) with either semi-internal double excitations or the Hamiltonian vertex $g_{vw}^{au} \hat{a}_{au}^{vw}$. Taking the actual subdivision of active indices into
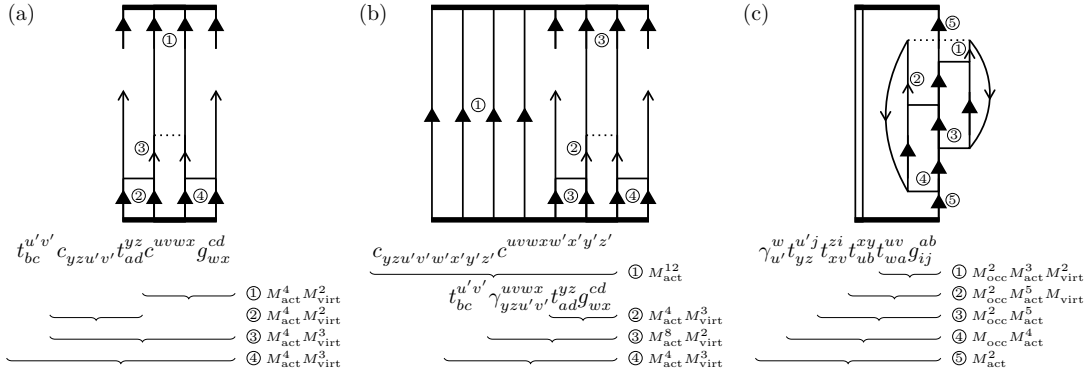
Figure 7: Sample terms to demonstrate the effect of a larger number of active electrons and higher commutators on the formal scaling of individual ic-MRCCSD terms.

---

[11]Note that the characterization of the formal scaling in terms of polynomials in M_{act} is only strictly valid in the limit $M_{\mathrm{act}} \to \infty$. If M_{act} is small, a proper assignment of the formal scaling should also take into account the actual subdivision of the active indices involved in a contraction into groups of indices with permutational antisymmetry. For simplicity, we retain the notation using polynomials in M_{act}, although the reader should keep in mind that the factual scaling with respect to M_{act} is usually somewhat lower than indicated.
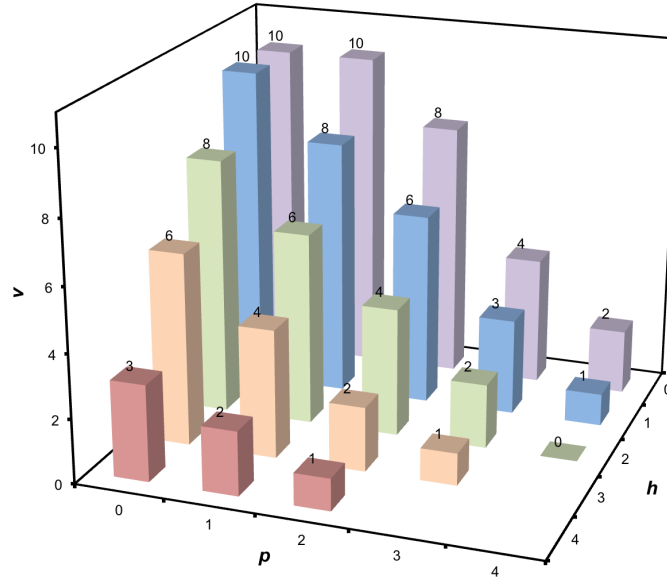
Figure 8: Diagram showing the maximum value of v for terms scaling like $M_{\text{occ}}^h M_{\text{act}}^v M_{\text{virt}}^p$ in case of ic-MRCCSD truncated at $N_{\text{com}} \leq 2$ (excluding terms needed for evaluating reduced density matrices).
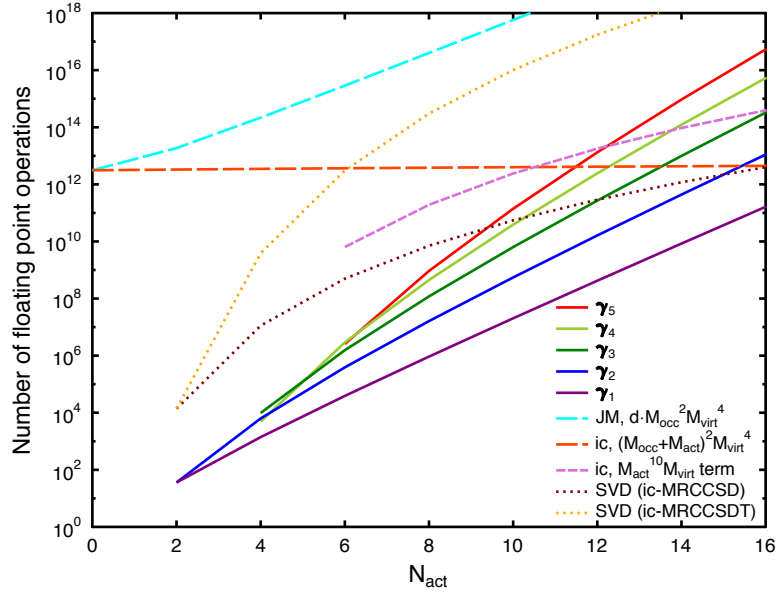

Figure 9: Dependence of the estimated number of floating point operations (log. scale) on the active space size for the following computational steps: the evaluation of $\boldsymbol{\gamma}_n$ [see Eq. (3.19)], the M_{virt}^4 terms in MRCCSD methods based on the Jeziorski–Monkhorst (JM) and the internally contracted (ic) ansatz, a sample term from ic-MRCCSD scaling like $M_{\text{act}}^{10} M_{\text{virt}}$ [see Eq. (3.17)], and the singular value decomposition (SVD) step [see Eq. (3.35)]. The plot assumes a CAS(N_{act},N_{act}), $N = 40$ and $M_{\text{occ}} + M_{\text{act}} + M_{\text{virt}} = 250$.

account, this scaling is more accurately characterized as

$$M_{\text{act}} \binom{M_{\text{act}}}{2} \binom{M_{\text{act}} - 1}{4} \binom{M_{\text{act}} - 2}{3} M_{\text{virt}}. \tag{3.17}$$

Hence, in contrast to MRCC methods based on the Jeziorski–Monkhorst ansatz, in which all terms inherit a factorial scaling behavior from the active space, the expense for evaluating the ic-MRCC residual is limited by a polynomial in M_{act}.

Therefore, in case of very large active spaces, the formation of the reduced density matrices $\boldsymbol{\gamma}_n$ according to

$$\gamma_{uv\cdots}^{wx\cdots} = \frac{1}{(N_{\text{act}} - n)!} \sum_{yz\cdots} c^{wx\cdots yz\cdots} c \underbrace{_{uv\cdots}}_{n} \underbrace{_{yz\cdots}}_{N_{\text{act}} - n} \tag{3.18}$$

will eventually become the rate determining step, with

$$\binom{M_{\text{act}}}{N_{\text{act}} - n} \binom{M_{\text{act}} - N_{\text{act}} + n}{n}^2 \tag{3.19}$$

as an upper bound for the computational scaling. Figure 9 visualizes this issue and summarizes the current discussion by comparing the scaling behaviour of selected computational steps in an ic-MRCCSD calculation with each other and with the most expensive term in a competing Jeziorski–Monkhorst-based theory.

It should be mentioned that the relaxation of the reference function involves the evaluation of $\langle \Phi_\mu | \bar{H} | \Psi_0 \rangle$ (see app. A.1.4) and also contributes to an exponential scaling with respect to the active space. For ic-MRCCSD with $N_{\text{com}} \leq 2$, active parts of $\bar{H}$ with ranks up to four can contribute, leading to a computational scaling similar to the one for computing $\boldsymbol{\gamma}_4$.

3.2 The redundancy problem in ic-MRCC theory

Unlike in SRCC theory, the excited functions $\hat{\tau}^\rho | \Psi_0 \rangle$ are not mutually orthogonal, leading to a nontrivial overlap (or *metric*) matrix,

$$\mathbf{S} = \langle \Psi_0 | \hat{\boldsymbol{\tau}}^\dagger \hat{\boldsymbol{\tau}} | \Psi_0 \rangle. \tag{3.20}$$

Because S_ρ^σ is zero for different inactive indices in $(\hat{\tau}^\rho)^\dagger$ and $\hat{\tau}^\sigma$, $\mathbf{S}$ naturally falls into a direct sum over individual excitation classes. In each excitation class, labeled by the number of inactive hole (n_h) and particle (n_p) indices, the metric can further be decomposed into a tensor $\mathbf{S}_{\text{act}}^{n_h, n_p}$ having solely active indices, and an antisymmetric unit tensor with inactive indices only:

$$\mathbf{S} = \bigoplus_{n_h, n_p} \left(\mathbf{S}_{\text{act}}^{n_h, n_p} \otimes \mathbf{1}_{\text{inact}}^{n_h, n_p} \right). \tag{3.21}$$

The antisymmetric unit tensor $\mathbf{1}$ is a generalization of Kronecker's delta and its elements may be defined through determinants,

$$1_{q_1 q_2 \cdots q_n}^{p_1 p_2 \cdots p_n} = \begin{vmatrix} \delta_{q_1}^{p_1} & \delta_{q_2}^{p_1} & \cdots & \delta_{q_n}^{p_1} \\ \delta_{q_1}^{p_2} & \delta_{q_2}^{p_2} & \cdots & \delta_{q_n}^{p_2} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{q_1}^{p_n} & \delta_{q_2}^{p_n} & \cdots & \delta_{q_n}^{p_n} \end{vmatrix}. \tag{3.22}$$

The explicit expressions of the blocks of $\mathbf{S}$ in ic-MRCCSD, collected in Tab. 3, consist of unit tensors and up to three-particle reduced density matrices in the active space. Table 3 also makes use of the one- and two-hole density matrices [P2, 54, 120], defined by

$$\eta_u^v = \delta_u^v - \gamma_u^v, \tag{3.23}$$

$$\eta_{uv}^{wx} = \gamma_{uv}^{wx} - \delta_u^w \gamma_v^x + \delta_u^x \gamma_v^w + \delta_v^w \gamma_u^x - \delta_v^x \gamma_u^w + 1_{uv}^{wx}. \tag{3.24}$$

Linear dependencies in $\{\hat{\tau}^\rho |\Psi_0\rangle\}$ occur whenever more than one distinct linear combination of the functions $\hat{\tau}^\rho |\Psi_0\rangle$ leads to the same configuration. Linear dependencies are always present in ic-MRCC theory for $N_{\text{rank}} \geq 2$ and we may divide them into three classes: (i) those between functions created by excitation operators of different rank, (ii) remaining linear dependencies due to overparameterization, and (iii) near-linear dependencies originating from small reference coefficients.

Linear dependencies of class (i) arise because summations over *spectator excitations* (that excite an electron from an active orbital into the same orbital) involve the number operator for active electrons, $\hat{n}_{\text{act}} = \sum_u \hat{a}_u^u$. For the simplest case of single and double excitations, these linear dependencies take the form

$$\sum_u \hat{a}_{ui}^{uA} |\Psi_0\rangle = N_{\text{act}} \hat{a}_i^A |\Psi_0\rangle, \tag{3.25}$$

and

$$\sum_u \hat{a}_{uv}^{ua} |\Psi_0\rangle = (N_{\text{act}} - 1)\hat{a}_v^a |\Psi_0\rangle. \tag{3.26}$$

Jeziorski and Monkhorst have early pointed out that such linear dependencies cause infinitely many solutions to the amplitude equations [84]. In order to see this, first realize that Eqs. (3.25), (3.26) and (3.15) imply the elements of the singles residual $\mathbf{\Omega}_1$ to be linear combinations of elements from $\mathbf{\Omega}_2$, e.g.,

$$\Omega_A^i = \frac{1}{N_{\text{act}}} \sum_u \Omega_{uA}^{ui}. \tag{3.27}$$

This in principle allows us to arbitrarily choose the singles amplitudes in $\hat{T}_1$ and only solve for $\hat{T}_2$ by fulfilling $\mathbf{\Omega}_2 = \mathbf{0}$. By virtue of Eq. (3.27) (and its analog for Ω_a^u), $\mathbf{\Omega}_1 = \mathbf{0}$ is then also satisfied. In an internally contracted MRCI method (or in the FCI limit, see sec. 3.3.2) this does not pose a problem since the redundant components in the wave operator simply destroy the reference function and therefore do not change the wave function. However, in ic-MRCC theory, components in $\hat{T}$ that destroy $|\Psi_0\rangle$ may affect the wave function (and hence the energy) through terms nonlinear in $\hat{T}$.

Spectator excitations generally give rise to additional, class (ii), linear dependencies, because there are different linear combinations of them that leave the reference function unchanged. Next to the number operator $\hat{n}_{\text{act}}$, they can also represent the spin operator $\hat{S}_z$, see Eq. (A.8) in the appendix for a definition. The relation

$$(M_s \hat{n}_{\text{act}} - N_{\text{act}} \hat{S}_z)|\Psi_0\rangle = 0 \tag{3.28}$$

therefore always implies a linear dependency for excitations containing both active annihilators and creators. These linear dependencies stem from an overparameterization

Table 3: Nonzero blocks of the metric matrix for ic-MRCCSD, grouped by excitation classes. In the compact notation used for $n_h = 0$, $n_p = 0$, omission of the subscript i of an index v_i implies a summation, e.g., $\delta_v^x \gamma_{wv}^{u_1 u_2} = \sum_{i \neq j} (-1)^i \delta_{v_i}^x \gamma_{wv_j}^{u_1 u_2}$.

n_h	$n_p = 0$	$n_p = 1$	$n_p = 2$
0	$S_{v,w}^{u,x} = -\gamma_{wv}^{ux} + \delta_v^x \gamma_w^u$ $S_{v_1 v_2, w}^{u_1 u_2, x} = \gamma_{wv_1 v_2}^{u_1 u_2 x} + \delta_v^x \gamma_{wv}^{u_1 u_2}$ $S_{v, w_1 w_2}^{u, x_1 x_2} = \gamma_{w_1 w_2 v}^{u x_1 x_2} + \delta_v^x \gamma_{w_1 w_2}^{ux}$ $S_{v_1 v_2, w_1 w_2}^{u_1 u_2, x_1 x_2} = \gamma_{w_1 w_2 v_1 v_2}^{u_1 u_2 x_1 x_2}$ $- \delta_v^x \gamma_{w_1 w_2 v}^{u_1 u_2 x} + 1_{v_1 v_2}^{x_1 x_2} \gamma_{w_1 w_2}^{u_1 u_2}$	$S_{a,v}^{u,b} = \delta_a^b \gamma_v^u$ $S_{wa,x}^{uv,b} = \delta_a^b \gamma_{wx}^{uv}$ $S_{a,vw}^{u,xb} = \delta_a^b \gamma_{vw}^{xu}$ $S_{wa,xy}^{uv,zb} = \delta_a^b(-\gamma_{wxy}^{zuv} + \delta_w^z \gamma_{xy}^{uv})$	$S_{ab,wx}^{uv,cd} = 1_{ab}^{cd} \gamma_{wx}^{uv}$
1	$S_{u,j}^{ai,v} = \delta_j^i \eta_u^v$ $S_{vw,j}^{ui,x} = \delta_j^i(-\gamma_{vw}^{ux} - \delta_v^u \gamma_w^x + \delta_w^u \gamma_v^x)$ $S_{u,vj}^{ai,wx} = \delta_j^i(-\gamma_{vu}^{wx} - \delta_u^w \gamma_v^x + \delta_u^x \gamma_v^w)$ $S_{vw,xj}^{ui,yz} = \delta_j^i(\gamma_{xvw}^{uyz} - \delta_v^y \gamma_{xw}^{uz} + \delta_v^z \gamma_{xw}^{uy} + \delta_w^y \gamma_{xv}^{uz} - \delta_w^z \gamma_{xv}^{uy} + 1_{vw}^{yz} \gamma_x^u)$	$S_{a,j}^{i,b} = 1_{ja}^{ib}$ $S_{va,j}^{ui,b} = 1_{ja}^{ib} \gamma_v^u$ $S_{a,uj}^{i,vb} = 1_{ja}^{ib} \gamma_u^v$ $S_{va,wj}^{ui,xb} = 1_{ja}^{ib}(-\gamma_{wv}^{ux} + \delta_v^x \gamma_w^u)$	$S_{ab,vj}^{ui,cd} = 1_{jab}^{icd} \gamma_v^u$
2	$S_{uv,kl}^{ij,wx} = 1_{kl}^{ij} \eta_{uv}^{wx}$	$S_{ua,kl}^{ij,vb} = 1_{kla}^{ijb} \eta_u^v$	$S_{ab,kl}^{ij,cd} = 1_{klab}^{ijcd}$

Table 4: Nonzero blocks of the GNO-based metric matrix for ic-MRCC (see sec. 3.2.2), grouped by excitation classes. For $n_h = 2$ and $n_p = 2$, the metric elements are identical to the ones shown in Tab. 3. See also the caption of Tab. 3.

n_h	$n_p = 0$	$n_p = 1$
0	$\tilde{S}_{v,w}^{u,x} = -\lambda_{wv}^{ux} + \eta_v^x \gamma_w^u$ $\tilde{S}_{v_1 v_2, w}^{u_1 u_2, x} = \lambda_{wv_1 v_2}^{u_1 u_2 x}$ $+ \eta_v^x \lambda_{wv_1 v_2}^{u_1 u_2} + \gamma_w^u \lambda_{v_1 v_2}^{u_1 u_2}$ $\tilde{S}_{v,w_1 w_2}^{u,x_1 x_2} = \lambda_{w_1 w_2 v}^{u x_1 x_2} + \eta_v^x \lambda_{w_1 w_2}^{ux} + \gamma_w^u \lambda_{w_1 w_2}^{x_1 x_2}$ $\tilde{S}_{v_1 v_2, w_1 w_2}^{u_1 u_2, x_1 x_2} = \lambda_{w_1 w_2 v_1 v_2}^{u_1 u_2 x_1 x_2} + \eta_v^u \lambda_{w_1 w_2 v_1 v_2}^{u_1 u_2} + \eta_v^{x_1 x_2} \gamma_{w_1 w_2}^{u_1 u_2}$ $+ \gamma_w^u \lambda_{wv_1 v_2}^{u_1 u_2 x} + \eta_{v_1 v_2}^{u_1 x_2} \gamma_{w_1 w_2}^{u_1 u_2} + \frac{1}{2} \lambda_{wv}^{ux} \lambda_{w_1 w_2}^{ux}$ $+ \gamma_w^u \lambda_{wv_1 v_2}^{ux} - \lambda_{wv}^{u_1 u_2} \lambda_{w_1 w_2}^{ux} - \eta_v^u \gamma_v^u \lambda_{wv}^{ux}$	$\tilde{S}_{a,v}^{u,b} = \delta_a^b \gamma_v^u$ $\tilde{S}_{wa,x}^{uv,b} = \delta_a^b \lambda_{wx}^{uv}$ $\tilde{S}_{a,vw}^{u,xb} = \delta_a^b \lambda_{vw}^{xu}$ $\tilde{S}_{wa,xy}^{uv,zb} = \delta_a^b(-\lambda_{wxy}^{zuv} - \gamma_x^u \lambda_{wy}^{zv} + \gamma_x^v \lambda_{wy}^{zu}$ $+ \gamma_y^u \lambda_{wx}^{zv} - \gamma_y^v \lambda_{wx}^{zu} + \gamma_{xy}^{uv} \eta_w^z)$
1	$\tilde{S}_{u,j}^{ai,v} = \delta_j^i \eta_u^v$ $\tilde{S}_{vw,j}^{ui,x} = -\delta_j^i \lambda_{vw}^{ux}$ $\tilde{S}_{u,vj}^{ai,wx} = -\delta_j^i \lambda_{vu}^{wx}$ $\tilde{S}_{vw,xj}^{ui,yz} = \delta_j^i(\lambda_{xvw}^{uyz} - \eta_v^y \lambda_{xw}^{uz} + \eta_v^z \lambda_{xw}^{uy}$ $+ \eta_w^y \lambda_{xv}^{uz} - \eta_w^z \lambda_{xv}^{uy} + \eta_{vw}^{yz} \gamma_x^u)$	$\tilde{S}_{a,j}^{i,b} = 1_{ja}^{ib}$ $\tilde{S}_{va,j}^{ui,b} = 0$ $\tilde{S}_{a,uj}^{i,vb} = 0$ $\tilde{S}_{va,wj}^{ui,xb} = 1_{ja}^{ib}(-\lambda_{wv}^{ux} + \eta_v^x \gamma_w^u)$

30

related to admitting excitations both into and out of the active orbitals. Another instance of this feature occurs whenever the number of active indices becomes large enough to 'exhaust' the reference space. This is easiest explained by an example: For $N_{\text{act}} = 2$, all the excitations of the type $\hat{a}_{uv}^{ab}$ create the same excited function when acting on the reference function,

$$\frac{1}{c_{uv}} \hat{a}_{uv}^{ab} |\Psi_0\rangle = \hat{a}_a^\dagger \hat{a}_b^\dagger |0\rangle, \tag{3.29}$$

hence they yield only a single linearly independent component. Likewise, as mentioned in sec. 3.1.2, for $N_{\text{act}} = 2$ the single excitations $\hat{a}_u^a$ exhaust the reference space within their excitation class, such that the double excitations $\hat{a}_{uv}^{wa}$ are completely redundant.

Class (i) and (ii) linear dependencies will be denoted as *exact* linear dependencies, as one may write them in the form $\hat{\tau}'^\rho |\Psi_0\rangle = 0$. Near-linear dependencies of class (iii), on the other hand, arise whenever a configuration in the reference function assumes a small weight ($\ll 1$). In the limit that a configuration vanishes (which may happen along a reaction path on a potential energy surface), no excitation from this configuration is possible and the linear dependency becomes exact.

Since there is no unique way to remove the redundant parameters in methods using an internally contracted ansatz, various possibilities have been considered in the literature [31, 42, 46–52]. Most of these restrict the cluster operator to a nonredundant set of excitations according to Eqs. (3.6) and (3.7), and are based on a diagonalization (or singular value decomposition, SVD) of the metric matrix,

$$\mathbf{s} = \mathbf{U}^\dagger \mathbf{S} \mathbf{U}. \tag{3.30}$$

In Eq. (3.30), $\mathbf{s}$ is a diagonal matrix containing the eigenvalues (or singular values) of $\mathbf{S}$. A vanishing eigenvalue $s_\rho^\rho = 0$ corresponds to an exact linear dependency. In the present work, the transformation matrix $\mathbf{X}$ in Eq. (3.6) is chosen such that it performs a canonical orthogonalization of the excited functions $\hat{\tau}^\rho |\Psi_0\rangle$, also known as Löwdin orthogonalization [38, 50, 121]:

$$\mathbf{X} = \mathbf{U} \mathbf{s}_\eta^{-\frac{1}{2}}. \tag{3.31}$$

The subscript η in $\mathbf{s}_\eta^{-1/2}$ indicates that those matrix elements are set to zero for which the corresponding eigenvalues are below a certain threshold η in order to exclude near-linear dependencies. The metric matrix with respect to the orthogonalized functions $\hat{\tau}'^\rho |\Psi_0\rangle$ hence becomes a truncated unit matrix with all elements belonging to discarded excitations equal to zero,

$$\mathbf{S}' = \mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbf{1}_\eta. \tag{3.32}$$

To ensure that the cluster operator is restricted to nonredundant amplitudes according to Eq. (3.7) during the iterative solution of the ic-MRCC equations, it is sufficient to project out the redundant components from the unconstrained amplitude vector $\mathbf{t}_{\text{raw}}$ in each iteration:

$$\mathbf{t} = \mathbf{P} \mathbf{t}_{\text{raw}}, \tag{3.33}$$

with

$$\mathbf{P} = \mathbf{X} \mathbf{X}^{-1} = \mathbf{U} \mathbf{1}_\eta \mathbf{U}^\dagger. \tag{3.34}$$

The computational cost associated with the removal of redundancies is dominated by the matrix diagonalization, Eq. (3.30), which scales cubically with the dimension of the

matrix. It does not depend on the number of inactive orbitals, as only the purely active blocks $\mathbf{S}_{\text{act}}^{n_h,n_p}$ need to be diagonalized. The largest number of active indices arises for $\mathbf{S}_{\text{act}}^{1,0}$ and $\mathbf{S}_{\text{act}}^{0,1}$, leading to an upper bound for the computational scaling of

$$\binom{M_{\text{act}}}{N_{\text{rank}}}^3 \binom{M_{\text{act}}}{N_{\text{rank}} - 1}^3 \overset{M_{\text{act}} \gg N_{\text{rank}}}{\approx} \frac{1}{N_{\text{rank}}!^3 (N_{\text{rank}} - 1)!^3} M_{\text{act}}^{6N_{\text{rank}}-3}. \tag{3.35}$$

The comparison with the scaling of other computational steps presented in Fig. 9 shows that the SVD is unlikely to become a computational bottleneck in case of ic-MRCCSD. Upon inclusion of triple excitations, however, a steep polynomial scaling like M_{act}^{15} results for large active spaces, which may well become the rate determining step.

3.2.1 A sequential orthogonalization technique

Equations (3.25) and (3.26) exemplify that the metric matrix couples excitations of different rank, i.e., $\hat{\tau}_n^\rho$ and $\hat{\tau}_m^\sigma$ with $n \neq m$. A *full orthogonalization* according to Eqs. (3.30), (3.31) and (3.6) therefore yields excitation operators $\hat{\tau}'^\rho$ that are linear combinations of excitations with different ranks. This is an undesirable feature of the theory, because (i) excitations of different rank usually arise in different orders of the corresponding perturbation theory (see sec. 5.3) and should therefore be distinguished, (ii) it makes all excitation operators dependent on N_{rank}, thus the definition of a perturbative triples correction (chap. 5) becomes less straightforward, and (iii) it causes a lack of size-extensivity, to be discussed in more detail in sec. 3.3.3.

This mixing of excitations with different rank can be avoided by a stepwise procedure starting from the lowest rank in each excitation class. Fink and Staemmler [50] have reported such a procedure, however, without defining it in detail. The following outlines the technique published in Ref. [P1]. The aim is to construct an operator basis $\{\hat{\tau}_n'^\rho\}$ that yields an orthonormal set of functions $\hat{\tau}_n'^\rho|\Psi_0\rangle$ while the individual operators $\hat{\tau}_n'^\rho$ should be linear combinations of excitation operators from the same rank, i.e., $\boldsymbol{\tau}_n' = \boldsymbol{\tau}_n \mathbf{X}_n$.

In order to project out the excited functions $\hat{\tau}_n^\rho|\Psi_0\rangle$ from higher excited functions, one may use the projector

$$\hat{p}_n = 1 - \sum_\rho \hat{\tau}_n'^\rho|\Psi_0\rangle\langle\Psi_0|(\hat{\tau}_n'^\rho)^\dagger = 1 - \boldsymbol{\tau}_n|\Psi_0\rangle\mathbf{X}_n\mathbf{X}_n^\dagger\langle\Psi_0|\boldsymbol{\tau}_n^\dagger. \tag{3.36}$$

Because the matrices $\mathbf{X}_n$ should individually transform the corresponding blocks of $\mathbf{S}$ according to $\mathbf{X}_n^\dagger\mathbf{S}_n\mathbf{X}_n = \mathbf{1}_{n,\eta}$, the definition of $\mathbf{X}_n$ has to be based on the diagonalization of a projected metric matrix that does not couple excitations of different rank. In case of single excitations, $n = 1$, no projection is needed and the transformation matrix $\mathbf{X}_1 = \mathbf{U}_1\mathbf{s}_{1,\eta}^{-1/2}$ follows directly from the diagonalization of $\mathbf{S}_1$.[12] For higher ranks $n > 1$, the contributions from effective lower-rank excitations can be projected out according to

$$\mathbf{S}_{n,\text{proj}} = \langle\Psi_0|\boldsymbol{\tau}_n^\dagger\hat{P}_n\boldsymbol{\tau}_n|\Psi_0\rangle = \mathbf{S}_n - \sum_{m=1}^{n-1} \mathbf{S}_{mn}^\dagger\mathbf{X}_m\mathbf{X}_m^\dagger\mathbf{S}_{mn}, \tag{3.37}$$

where

$$\hat{P}_n = \prod_{m=1}^{n-1} \hat{p}_m \tag{3.38}$$

---

[12]The product $\mathbf{X}_1\mathbf{X}_1^\dagger$ in Eq. (3.36) is then given by the pseudoinverse $\mathbf{S}_1^{-1}$, cf. Ref. [122].

and

$$\mathbf{S}_{mn} = \langle \Psi_0 | \boldsymbol{\tau}_m^\dagger \boldsymbol{\tau}_n | \Psi_0 \rangle. \tag{3.39}$$

In order to derive an expression for $\mathbf{X}_n$, it is advantageous to reformulate Eq. (3.37), using instead of $\hat{P}_n$ a projector $\mathbf{Q}_n$ that can be applied to $\boldsymbol{\tau}_n | \Psi_0 \rangle$ from the right, satisfying

$$\boldsymbol{\tau}_n | \Psi_0 \rangle \mathbf{Q}_n = \hat{P}_n \boldsymbol{\tau}_n | \Psi_0 \rangle. \tag{3.40}$$

It then follows immediately from the diagonalization

$$\mathbf{s}_n = \mathbf{U}_n^\dagger \mathbf{S}_{n,\text{proj}} \mathbf{U}_n = \mathbf{U}_n^\dagger \mathbf{Q}_n^\dagger \mathbf{S}_n \mathbf{Q}_n \mathbf{U}_n \tag{3.41}$$

that the transformation matrices must be defined as

$$\mathbf{X}_n = \mathbf{Q}_n \mathbf{U}_n \mathbf{s}_{n,\eta}^{-\frac{1}{2}} \tag{3.42}$$

in order to fulfill $\mathbf{X}_n^\dagger \mathbf{S}_n \mathbf{X}_n = \mathbf{1}_{n,\eta}$.

The construction of $\mathbf{Q}_n$ requires knowledge of the explicit dependencies between excitations of different rank. If these dependencies are known for all $m < n$ in the form

$$\boldsymbol{\tau}_m | \Psi_0 \rangle = \boldsymbol{\tau}_n | \Psi_0 \rangle \mathbf{Z}_{nm}, \tag{3.43}$$

the projectors $\hat{p}_m$ can be reformulated in terms of n-body excitations:

$$\hat{p}_m = 1 - \boldsymbol{\tau}_n | \Psi_0 \rangle \mathbf{Z}_{nm} \mathbf{X}_m \mathbf{X}_m^\dagger \mathbf{Z}_{nm}^\dagger \langle \Psi_0 | \boldsymbol{\tau}_n^\dagger. \tag{3.44}$$

Inserting Eq. (3.44) into Eq. (3.40), one may identify the desired projector with

$$\mathbf{Q}_n = \mathbf{1}_n - \left(\sum_{m=1}^{n-1} \mathbf{Z}_{nm} \mathbf{X}_m \mathbf{X}_m^\dagger \mathbf{Z}_{nm}^\dagger \right) \mathbf{S}_n. \tag{3.45}$$

In general, the relations (3.43) are not unique, and different choices of $\mathbf{Z}_{nm}$ lead to different variants of ic-MRCC theory. The current work uses a definition based on the following generalization of Eqs. (3.25) and (3.26) [P1]:

$$\sum_{u_1 \cdots u_g} \hat{a}_{u_1 \cdots u_g v_1 \cdots v_h ij \cdots}^{u_1 \cdots u_g AB \cdots} | \Psi_0 \rangle = \frac{(N_{\text{act}} - h)!}{(N_{\text{act}} - h - g)!} \hat{a}_{v_1 \cdots v_h ij \cdots}^{AB \cdots} | \Psi_0 \rangle. \tag{3.46}$$

Restriction of the sum in Eq. (3.46) to nonredundant index tuples and rearrangement yields

$$\hat{a}_{v_1 \cdots v_h ij \cdots}^{AB \cdots} | \Psi_0 \rangle = \frac{g!(N_{\text{act}} - h - g)!}{(N_{\text{act}} - h)!} \sum_{u_1 < \cdots < u_g} \hat{a}_{u_1 \cdots u_g v_1 \cdots v_h ij \cdots}^{u_1 \cdots u_g AB \cdots} | \Psi_0 \rangle. \tag{3.47}$$

By comparison with Eq. (3.43) one can identify the elements of $\mathbf{Z}_{nm}$ with

$$(Z_{nm})_{x_1 \cdots x_g CD \cdots, v_1 \cdots v_h ij \cdots}^{u_1 \cdots u_g y_1 \cdots y_h kl \cdots, AB \cdots} = \binom{N_{\text{act}} - h}{g}^{-1} 1_{x_1 \cdots x_g v_1 \cdots v_h CD \cdots ij \cdots}^{u_1 \cdots u_g y_1 \cdots y_h AB \cdots kl \cdots}. \tag{3.48}$$

Equations (3.41), (3.42), (3.45) and (3.48) are the working equations of this *sequential orthogonalization* technique. It is a recursive procedure, as the projectors $\mathbf{Q}_n$ depend on the transformation matrices of all lower ranks. Figure 10 illustrates how the recursive

	$\hat{\tau}_1|\Psi_0\rangle$	$\hat{\tau}_2|\Psi_0\rangle$	$\hat{\tau}_3|\Psi_0\rangle$
$\langle\Psi_0|\hat{\tau}_1^\dagger$	$\mathbf{S}_1$	$\neq 0$	$\neq 0$
$\langle\Psi_0|\hat{\tau}_2^\dagger$	$\neq 0$	$\mathbf{S}_2$	$\neq 0$
$\langle\Psi_0|\hat{\tau}_3^\dagger$	$\neq 0$	$\neq 0$	$\mathbf{S}_3$

$\xrightarrow{\text{Projection}}$

	$\hat{\tau}_1|\Psi_0\rangle$	$\hat{\tau}_2|\Psi_0\rangle\mathbf{Q}_2$	$\hat{\tau}_3|\Psi_0\rangle\mathbf{Q}_3$
$\langle\Psi_0|\hat{\tau}_1^\dagger$	$\mathbf{S}_1$	$= 0$	$= 0$
$\mathbf{Q}_2^\dagger\langle\Psi_0|\hat{\tau}_2^\dagger$	$= 0$	$\mathbf{Q}_2^\dagger\mathbf{S}_2\mathbf{Q}_2$	$= 0$
$\mathbf{Q}_3^\dagger\langle\Psi_0|\hat{\tau}_3^\dagger$	$= 0$	$= 0$	$\mathbf{Q}_3^\dagger\mathbf{S}_3\mathbf{Q}_3$
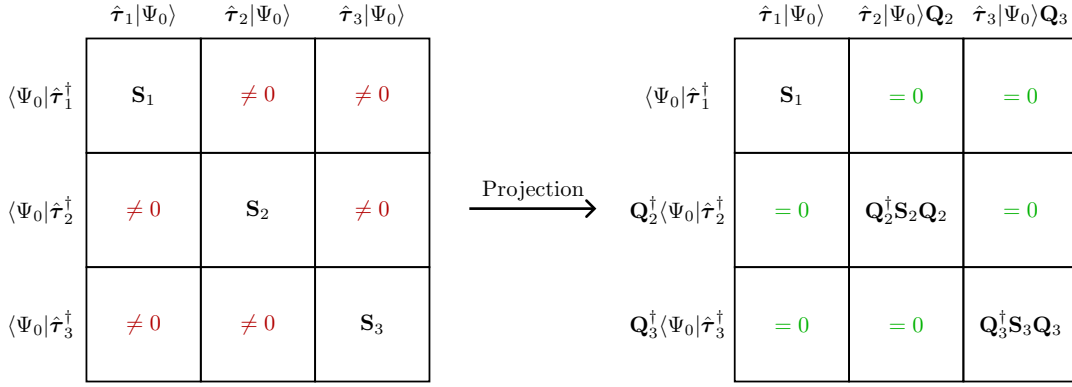
Figure 10: Illustration of the projection that decouples excitations of different rank in the metric.

projection ensures a decoupling of excitations with different ranks. This procedure also leads to a natural truncation of the operator rank in each excitation class, because, for a given active space, a finite number of active creation and annihilation operators are sufficient to create all possible distributions of active electrons.

In contrast to a full orthogonalization, the sequential orthogonalization does not affect purely inactive excitations. The resulting ic-MRCC theory thus inherits the size-extensivity of SRCC with respect to parts of a system described by inactive orbitals, a property termed *core-extensivity* [123–125]. However, the sequential orthogonalization does not guarantee size-extensivity in the active space.

3.2.2   A technique based on the generalized normal ordering

A fully size-extensive ic-MRCC theory can be achieved within a formalism that takes the multiconfigurational reference function $|\Psi_0\rangle$ as Fermi vacuum [P4, 30].

Such a generalized normal ordering (GNO) gives rise to an extended version of Wick's theorem, which is not limited to binary contractions but allows joint contractions involving any number of n annihilation operators and n creation operators [30, 111, 112]. In analogy to Eqs. (2.21) and (2.22), relations between excitation operators in the two normal orderings result from applying the extended Wick theorem to a string of annihilation and creation operators:

$$
\begin{aligned}
\hat{a}^{rst\cdots}_{opq\cdots} = {} & \{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p\hat{a}_o\}_0 + \{\overparen{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p}\hat{a}_o\}_0 + \ldots \\
& + \{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p\hat{a}_o\}_0 + \ldots + \{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p\hat{a}_o\}_0 + \ldots \\
& + \{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p\hat{a}_o\}_0 + \ldots + \{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p\hat{a}_o\}_0 + \ldots \\
& + \{\hat{a}^r\hat{a}^s\hat{a}^t\cdots\hat{a}_q\hat{a}_p\hat{a}_o\}_0 + \ldots\,.
\end{aligned}
\tag{3.49}
$$

While a binary contraction triggers an element of either the one-particle or the one-hole

density matrix (with respect to the reference function),

$$\overset{\frown}{\hat{a}^q \hat{a}_p} = \gamma_p^q, \tag{3.50}$$

$$\overset{\frown}{\hat{a}_p \hat{a}^q} = \delta_p^q - \gamma_p^q = \eta_p^q, \tag{3.51}$$

a joint contraction of $2n$ operators with $n \geq 2$ always yields an n-body cumulant:

$$\overset{\frown}{\hat{a}^r \hat{a}^s \hat{a}_q \hat{a}_p} = -\overset{\frown}{\hat{a}^r \hat{a}^s \hat{a}_p \hat{a}_q} = -\overset{\frown}{\hat{a}^r \hat{a}_q \hat{a}^s \hat{a}_p} = \ldots = \lambda_{pq}^{rs}, \tag{3.52}$$

$$\overset{\frown}{\hat{a}^r \hat{a}^s \hat{a}^t \hat{a}_q \hat{a}_p \hat{a}_o} = \ldots = \lambda_{opq}^{rst}, \tag{3.53}$$

$$\ldots \ldots$$

Equation (3.49) now takes the form

$$\hat{a}_{opq\cdots}^{rst\cdots} = \tilde{a}_{opq\cdots}^{rst\cdots} + \gamma_o^r \tilde{a}_{pq\cdots}^{st\cdots} + \ldots + \gamma_o^r \gamma_p^s \tilde{a}_{q\cdots}^{t\cdots} + \ldots + \lambda_{op}^{rs} \tilde{a}_{q\cdots}^{t\cdots} + \ldots, \tag{3.54}$$

where operators normal ordered with respect to $|\Psi_0\rangle$ are denoted as $\tilde{a}_{pq\cdots}^{rs\cdots} = \{\hat{a}_{pq\cdots}^{rs\cdots}\}_0$ and the summations cover all nonredundant permutations of upper and lower indices. Only purely active elements of the one-particle density matrix arise in Eq. (3.54), since there are no hole annihilation operators in $\hat{a}_{opq\cdots}^{rst\cdots}$ that could enable inactive elements $\gamma_i^j = \delta_i^j$. Introducing an antisymmetrized tensor product $\otimes_A$ that implicitly excludes redundant permutations among the indices within each antisymmetric tensor [P2],[13] Eq. (3.54) is conveniently expressed as

$$\boldsymbol{\tau}_n = \tilde{\boldsymbol{\tau}}_n + \boldsymbol{\gamma}_1 \otimes_A \tilde{\boldsymbol{\tau}}_{n-1} + \underbrace{(\tfrac{1}{2} \boldsymbol{\gamma}_1 \otimes_A \boldsymbol{\gamma}_1 + \boldsymbol{\lambda}_2)}_{\langle \Psi_0 | \boldsymbol{\tau}_2^{\mathrm{act}} | \Psi_0 \rangle} \otimes_A \tilde{\boldsymbol{\tau}}_{n-2} + \ldots. \tag{3.55}$$

As indicated in Eq. (3.55), each summand contains a factor identical to the fully contracted contribution (i.e., the vacuum expectation value) of an excitation operator of the same rank. This expectation value is given by an active reduced density matrix, $\langle \Psi_0 | \boldsymbol{\tau}_k^{\mathrm{act}} | \Psi_0 \rangle = \boldsymbol{\gamma}_k$, and therefore provides a straightforward definition of the cumulant tensors (see, e.g., Eq. (2.25)), which are also purely active quantities. The relation between excitation operators in the two normal orderings then becomes

$$\boldsymbol{\tau}_n = \sum_{k=0}^{n} \boldsymbol{\gamma}_k \otimes_A \tilde{\boldsymbol{\tau}}_{n-k}, \qquad \text{with} \quad \boldsymbol{\gamma}_0 = \tilde{\boldsymbol{\tau}}_0 = 1. \tag{3.56}$$

Equation (3.56) can be inverted by recursively inserting the relations for lower ranks. As there are no purely active excitations in the present ic-MRCC formalism, the sum in Eq. (3.56) truncates at $k = n-1$. While Eq. (3.56) then implies no difference for single excitations, $\tilde{\boldsymbol{\tau}}_1 = \boldsymbol{\tau}_1$, double excitations with both active annihilators and creators differ by an effective single excitation:

$$\tilde{a}_{ui}^{va} = a_{ui}^{va} - \gamma_u^v a_i^a, \tag{3.57}$$

$$\tilde{a}_{ui}^{vw} = a_{ui}^{vw} - \gamma_u^v a_i^w + \gamma_u^w a_i^v, \tag{3.58}$$

$$\tilde{a}_{uv}^{wa} = a_{uv}^{wa} - \gamma_u^w a_v^a + \gamma_v^w a_u^a. \tag{3.59}$$

---

[13]The operator $\otimes_A$ is a Grassmann (or wedge) product with a modified prefactor. An example for its action is $(\boldsymbol{\gamma}_1 \otimes_A \boldsymbol{\tau}_2)_{opq}^{rst} = \gamma_o^r \hat{a}_{pq}^{st} - \gamma_o^s \hat{a}_{pq}^{rt} - \gamma_o^t \hat{a}_{pq}^{sr} - \gamma_p^r \hat{a}_{oq}^{st} + \gamma_p^s \hat{a}_{oq}^{rt} + \gamma_p^t \hat{a}_{oq}^{sr} - \gamma_q^r \hat{a}_{po}^{st} + \gamma_q^s \hat{a}_{po}^{rt} + \gamma_q^t \hat{a}_{po}^{sr}$.

Generally, the excitations can be transformed between the two normal orderings according to

$$\tilde{\boldsymbol{\tau}} = \boldsymbol{\tau}\mathbf{Y}, \tag{3.60}$$

$$\boldsymbol{\tau} = \tilde{\boldsymbol{\tau}}\mathbf{Y}^{-1}, \tag{3.61}$$

where $\mathbf{Y}$ and its inverse are upper triangular matrices of the structure

$$\mathbf{Y} = \begin{pmatrix} \mathbf{1} & \mathbf{Y}_{12} & \mathbf{Y}_{13} & \cdots \\ \mathbf{0} & \mathbf{1} & \mathbf{Y}_{23} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{3.62}$$

$$\mathbf{Y}^{-1} = \begin{pmatrix} \mathbf{1} & -\mathbf{Y}_{12} & (\mathbf{Y}_{12}\mathbf{Y}_{23} - \mathbf{Y}_{13}) & \cdots \\ \mathbf{0} & \mathbf{1} & -\mathbf{Y}_{23} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.63}$$

With the cluster operator being invariant with respect to the normal ordering, $\hat{T} = \boldsymbol{\tau}\mathbf{t} = \tilde{\boldsymbol{\tau}}\tilde{\mathbf{t}}$, the GNO-based cluster amplitudes may be transformed by

$$\mathbf{t} = \mathbf{Y}\tilde{\mathbf{t}}. \tag{3.64}$$

Since $\tilde{\mathbf{t}}_n$ only contributes to $\mathbf{t}_k$ with $k \leq n$ and vice versa, a cluster operator truncated at rank N_{rank} can be translated from one normal ordering to the other without including additional excitation ranks. Likewise, it is possible to obtain the GNO-based residual from the residual evaluated in the conventional particle-hole formalism by a simple transformation:

$$\tilde{\boldsymbol{\Omega}} = \langle \Psi_0 | \tilde{\boldsymbol{\tau}}^{\dagger} \bar{H} | \Psi_0 \rangle = \mathbf{Y}^{\dagger}\boldsymbol{\Omega}. \tag{3.65}$$

Equations (3.65) and (3.64) allow the implementation of a fully GNO-based ic-MRCC theory while circumventing the application of the extended Wick theorem to the residual expression [P4].

The main difference to the approach detailed in the previous chapters is that the removal of linear dependencies may now be conveniently based on the metric matrix in the generalized normal ordering,

$$\tilde{\mathbf{S}} = \langle \Psi_0 | \tilde{\boldsymbol{\tau}}^{\dagger} \tilde{\boldsymbol{\tau}} | \Psi_0 \rangle, \tag{3.66}$$

see Tab. 4 for explicit expressions. $\tilde{\mathbf{S}}$ does not couple purely inactive excitations with excitations involving additional active indices, because the generalized normal ordering prevents contractions among active indices of the same operator. While this is sufficient to ensure core-extensivity, $\tilde{\mathbf{S}}$ further restricts the coupling between excitations with active indices. Contractions between active indices on $\tilde{\boldsymbol{\tau}}^{\dagger}$ and $\tilde{\boldsymbol{\tau}}$ can only take place through elements of $\boldsymbol{\delta}$, $\boldsymbol{\gamma}_1$, or through cumulants. As all of these quantities are extensive [P2,111], GNO-based ic-MRCC theory is fully size-extensive. A more elaborate explanation of the size-extensivity in ic-MRCC theory is given in sec. 3.3.3.

3.3 Formal aspects of ic-MRCC theory

Key features of ic-MRCC theory are the invariance with respect to unitary transformations within the three orbital spaces, the convergence towards FCI upon increasing N_{rank} and its size-extensivity properties. Mathematical proof of these properties was published in Refs. [P1,P4,P5,38] and will not be repeated here. Nevertheless, the following sections shall provide a conceptual understanding of these key features.

3.3.1 Orbital invariance

Orbital rotations imply a unitary transformation of all strings of creation and annihilation operators. In principle, all operators can be expressed in the rotated orbital basis by a corresponding counter-rotation of the associated tensors. While operators involving summations over the whole orbital space such as $\hat{H}$ are thus trivially orbital invariant, summations over partial orbital spaces in the operator $\hat{T}$ and the reference function $|\Psi_0\rangle$ restrict the invariance of these quantities to orbital rotations within the hole, valence and particle orbital spaces. The ic-MRCC method is called orbital invariant if the fulfillment of the ic-MRCC equations and the ic-MRCC energy are independent of orbital rotations within these three orbital spaces.

The ic-MRCC energy expression, Eq. (3.3), is an expectation value of $\hat{H}$ and $\hat{T}$ with respect to $|\Psi_0\rangle$ and therefore orbital invariant. The eigenvalue equations (3.5) consist of the same operators and merely translate an active orbital rotation into a corresponding rotation of the reference coefficients. The residual $\boldsymbol{\Omega}$ defined in Eq. (3.15) is also rotated. If only exact linear dependencies are excluded in the construction of the excitation basis $\{\hat{\tau}'\}$, fulfillment of $\boldsymbol{\Omega}' = \mathbf{0}$ implies $\boldsymbol{\Omega} = \mathbf{0}$ and is independent of the orbital rotation. However, if near-linear dependencies are excluded as well, $\boldsymbol{\Omega}$ is generally nonzero upon fulfillment of the ic-MRCC equations and orbital invariance hinges on the procedure used for removing redundant excitations. Since the diagonalization of $\mathbf{S}$ is a unitary transformation, the canonical orthogonalization naturally compensates for any active orbital rotation and always leads to the same nonredundant excitation basis [38]. The sequential orthogonalization also preserves orbital invariance, as it is based on Eq. (3.46), which is an orbital invariant relationship between excitations of different rank [P1].

3.3.2 Convergence to the full configuration interaction limit

The equivalence of ic-MRCC theory and FCI in the limit $N_{\mathrm{rank}} = N$ is guaranteed by the possibility to represent the FCI wave function with the ic-MRCC ansatz, as the ic-MRCC equations are then projections of the Schrödinger equation onto the full space of excited functions. Let us first recognize that the FCI wave function can be written in terms of an ic-MRCI ansatz of the form

$$|\Psi_{\mathrm{ic\text{-}MRCI}}\rangle = (1 + \hat{C})|\Psi_0\rangle, \tag{3.67}$$

where $\hat{C}$ is isomorphic to the cluster operator. While the coefficients in $|\Psi_0\rangle$ are easily obtained from projecting the FCI wave function onto the reference determinants, the nonredundant amplitudes in $\hat{C}$ follow directly from projection upon the functions $\hat{\tau}'^{\rho}|\Psi_0\rangle$. The redundant components of $\hat{C}$ are arbitrary, as they give no contribution to the wave function when acting on $|\Psi_0\rangle$.

In order to obtain the corresponding cluster amplitudes, we can equate the ic-MRCI wave function from Eq. (3.67) with the ic-MRCC ansatz

$$|\Psi_{\text{ic-MRCC}}\rangle = \left(1 + \hat{T} + \mathcal{F}(\hat{T})\right)|\Psi_0\rangle. \tag{3.68}$$

In Eq. (3.68), $\mathcal{F}(\hat{T})$ is a function at least quadratic in $\hat{T}$ in accordance with the exponential wave operator. Equations (3.67) and (3.68) lead to the following condition:

$$\hat{T}|\Psi_0\rangle = \left(\hat{C} - \mathcal{F}(\hat{T})\right)|\Psi_0\rangle, \tag{3.69}$$

from which we can determine the values of the nonredundant cluster amplitudes by projecting onto the functions $\hat{\tau}'^\rho|\Psi_0\rangle$. Because the right-hand side of Eq. (3.69) depends quadratically on $\hat{T}$, this projection is done successively for excitation operators with increasing number of inactive indices. Equation (3.69) allows one to freely choose the redundant components of $\hat{T}$, and different choices may impact cluster amplitudes with more inactive indices through $\mathcal{F}(\hat{T})$. However, as long as the cluster operator is not truncated, each of these choices is a valid parameterization of the FCI wave function [P5].

3.3.3   Size-extensivity & size-consistency

A many-body theory is size-extensive [126–128] if the energy expression is a sum of connected terms. For two noninteracting subsystems A and B with orbitals localized on either subsystem, this means that each diagram consists of a connected product of additively separable operators (i.e., $\hat{X} = \hat{X}_A + \hat{X}_B$). Since operators working on different subsystems have no index in common, the only surviving contributions from connected terms are the ones to the subsystems' energies E_A and E_B, rendering the total energy additively separable or size-consistent [129]. Additive separability for arbitrary partitionings of an interacting system into artificial noninteracting fragments is a sufficient criterion for size-extensivity [125]. Due to the orbital invariance of ic-MRCC theory, size-consistency then also holds irrespective of the choice of orbitals.

In SRCC theory, connectedness of the residual $\boldsymbol{\Omega}$ implies additive separability of the cluster operator for noninteracting subsystems and thereby ensures size-extensivity. In addition to this criterion, size-extensivity in ic-MRCC theory also requires additive separability of the projection $\mathbf{Pt}$ from Eq. (3.33) and separability of the eigenvalue equations (3.5). An important prerequisite for the connectedness of $\boldsymbol{\Omega}$ is the connectedness of $\bar{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$. This is fulfilled because the BCH expansion consists of nested commutators only.

While connectedness of $\bar{H}$ directly ensures the separability of the eigenvalue equations [P4], the residual expression may contain disconnected terms due to the occurrence of reduced density matrices, which are not additively separable in contrast to the density matrix cumulants [P2, 111]. The disconnected terms become apparent when inserting Eq. (3.56) into the residual expression:

$$\boldsymbol{\Omega}_n = \langle\Psi_0|\hat{\boldsymbol{\tau}}_n^\dagger\bar{H}|\Psi_0\rangle = \sum_{k=0}^{n-1}\boldsymbol{\gamma}_k\otimes_A\langle\Psi_0|\tilde{\boldsymbol{\tau}}_{n-k}^\dagger\bar{H}|\Psi_0\rangle = \sum_{k=0}^{n-1}\boldsymbol{\gamma}_k\otimes_A\tilde{\boldsymbol{\Omega}}_{n-k}, \tag{3.70}$$

where the residuals $\tilde{\boldsymbol{\Omega}}_n$ in the generalized normal ordering are connected by virtue of the extended Wick theorem. Since $\tilde{\boldsymbol{\Omega}}$ is the residual in the GNO-based scheme from sec. 3.2.2,

no disconnected terms are present in this case [30]. Even when using the particle-hole normal ordering, the disconnected terms in Eq. (3.70) occur only for residual elements with active indices and thus do not affect the core-extensivity of the method. Further, they may cancel naturally upon fulfillment of the ic-MRCC equations, e.g., $\boldsymbol{\Omega}_1 = \mathbf{0}$ leads to a connected expression for $\boldsymbol{\Omega}_2$.

Hence, the residual expression does not necessarily introduce a lack of size-extensivity in ic-MRCC theory. Instead, the separability of the projection $\mathbf{Pt}$ is the critical issue. The projector $\mathbf{P}$ has the same block structure as the metric matrix $\mathbf{S}$. It mixes connected amplitudes (with indices either on A or B) and disconnected amplitudes (with indices on both A and B) whenever the corresponding excitation operators couple through the metric matrix. An analysis of the product $\mathbf{St}$ reveals that the full orthogonalization approach within the particle-hole normal ordering spoils the core-extensivity of the method. Core-extensivity is restored by the sequential orthogonalization technique, as the projected metric matrix does not couple excitations of different rank and therefore avoids a mixing of purely inactive amplitudes with those involving active indices. Finally, GNO-based ic-MRCC theory is fully size-extensive because the metric matrix $\tilde{\mathbf{S}}$ avoids any coupling between connected and disconnected excitations due to the extended Wick theorem. For further details and numerical tests of these size-extensivity properties, the reader is referred to Ref. [P4].

4 Benchmark calculations

The ic-MRCC method described in chap. 3 has been implemented in the General Contraction Code (GeCCo) [P1, 44], see app. A.1 for details.

This chapter's aim is to evaluate the accuracy of the ic-MRCC method for (ground state) energy calculations. Because the most meaningful assessment is a comparison to FCI energies, which are the exact eigenvalues of the Hamiltonian in a given basis set, the current study is limited to small model systems accessible to the FCI method. While the small basis sets of double-zeta quality considered in this chapter are not sufficient for quantitative predictions of potential energy surfaces, they nevertheless allow an assessment of the correlation energy recovered by the ic-MRCC method and a clear-cut comparison to competing electronic structure methods. An evaluation of the method's computational efficiency at the example of these model systems would, however, be pointless. The reason for this is a large number of low-scaling terms in the ic-MRCC equations that cause a significant computational overhead in small applications but have little impact on the computational cost once the system size or basis set size is increased. For benchmark applications allowing a comparison to experimental results, the reader is referred to secs. 5.6 and 6.5.

The current chapter is organized as follows: First, sec. 4.1 studies the overall performance of ic-MRCC theory for six selected model systems, using a set of standard approximations. Both the ic-MRCC equations and the energy expression are truncated at quadratic power in $\hat{T}$ and redundant amplitudes are eliminated by the sequential orthogonalization procedure described in sec. 3.2.1 (based on the core determinant $|0\rangle$ as Fermi vacuum), using the threshold $\eta = 10^{-6}$ for discarding small metric eigenvalues unless otherwise noted. These default approximations turn ic-MRCC theory into an affordable method while hardly impairing its intrinsic accuracy. A justification for these approximations is provided in the subsequent sections, each of which focuses on an individual aspect of ic-MRCC theory. The assessment of truncations in the BCH expansion (sec. 4.2, see also Ref. [38]) is followed by an analysis of the neglect of metric eigenvalues below η (sec. 4.3). Section 4.4 investigates various schemes to remove linear dependencies between excitations of different rank, before the chapter closes with a discussion of alternatives to the relaxation of the reference function (sec. 4.5).

All calculations employ CASSCF orbitals constructed from spherical Gaussian basis functions. The active space sizes range from two electrons in two orbitals to six electrons in twelve orbitals. The numerical data underlying the presented figures is collected in app. A.2.

4.1 Performance of ic-MRCC for model systems

4.1.1 Insertion of a beryllium atom into hydrogen

The reaction $BeH_2 \rightarrow Be + H_2$, illustrated in Fig. 11, offers a simple model process featuring a transition state with strong multireference character [130]. At one side of the reaction ($z = 4\ a_0$), the separated beryllium atom and hydrogen molecule are well described by a closed-shell configuration $1a_1^2 2a_1^2 3a_1^2$ (Φ_1), with $3a_1$ being the bonding σ orbital of H_2. The linear BeH_2 molecule at the other side ($z = 0\ a_0$) is dominated by the configuration $1\sigma_g^2 2\sigma_g^2 1\sigma_u^2$. Assuming a C_{2v} reaction channel in which Be is placed
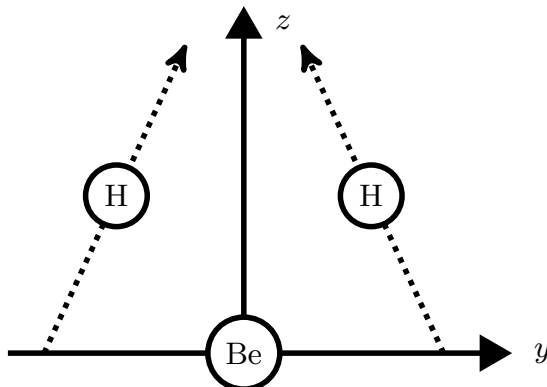
Figure 11: The C_{2v} model path for the reaction $BeH_2 \rightarrow Be + H_2$.

at the origin and the positions of the hydrogen nuclei are varied according to $y(z) = \pm(2.54\ a_0 - 0.46z)$ [131], this corresponds to $1a_1^2 2a_1^2 1b_2^2$ (Φ_2). The orbital $1b_2$ is bonding, but along the reaction path it evolves into the unoccupied p_y orbital of Be.

Figure 12a shows the potential energy of the $^1A_1$ ground state along the model reaction path, computed using the CASSCF and FCI methods and the cc-pVDZ basis set [132]. The CASSCF wave function is based on the active orbitals $3a_1$ and $1b_2$ and therefore consists of the two determinants Φ_1 and Φ_2. The energetically lowest CASSCF orbital, mainly a Be $1s$ orbital, is not correlated in all subsequent calculations.

CASSCF accommodates the multireference character around the transition state at $z \approx 2.8\ a_0$ by gradually switching the weights of Φ_1 and Φ_2 in the wave function and thus yields a qualitatively correct potential energy curve. However, as shown in the inset of Fig. 12b, the errors of the CASSCF energy range between 60 and 80 mE_h, which implies a predicted energy barrier that is too high by about 10 kcal/mol. The desire to obtain *chemical accuracy*, i.e., relative energies correct up to 1 kcal/mol (1.6 mE_h), calls for an accurate treatment of dynamic electron correlation.

The BeH_2 model has been frequently used to benchmark multireference methods [76, 90, 101, 133–137], including ic-MRCCSD [38]. Figure 12b compares the errors in the potential energy curves of MRCI singles-and-doubles (MRCISD) with those of CASCCSD and ic-MRCC, which fare among the best MRCC methods for this model system [P6, 38, 101]. As the CASCC results available in the literature [101] apply to a different basis set, the present values were recomputed using Olsen's LUCIA code [138]. Both MRCC methods clearly outperform MRCISD and lead to energy errors changing by less than 2 mE_h throughout the reaction path. Although CASCCSD can compete with ic-MRCCSD in terms of accuracy, ic-MRCC theory is conceptually more appealing as it is based on a true multiconfigurational reference function and treats both reference determinants on an equal footing. When using CASCC, on the other hand, one is forced to choose either Φ_1 or Φ_2 as reference function. Single and double excitations from the second determinant are then accounted for by incorporating selected triple and quadruple excitations in the cluster operator. Since it is not an internally contracted approach, the number of cluster amplitudes is the same as in the corresponding MRCI method
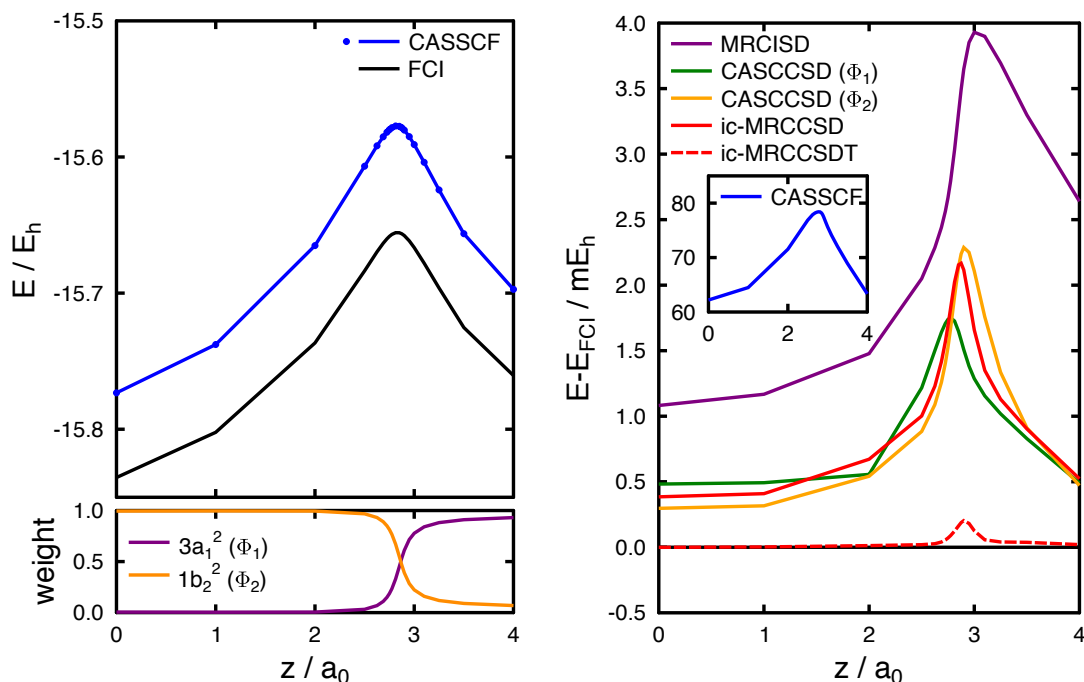
(a) Potential energy curve (top) and weights of the CASSCF configurations (bottom).

(b) Errors with respect to FCI.

Figure 12: Potential energy along the model reaction path of the $^1A_1$ ground state of BeH_2. Note the different energy scales.

and therefore usually larger than in ic-MRCC or in ic-MRCI.[14,15] The single-reference framework of CASCC theory introduces a bias towards the reference function. A close inspection of the ic-MRCCSD curve in Fig. 12b reveals that for $z < 2.8$ a_0 it is almost parallel to the CASCC curve that uses the dominating determinant Φ_2 as a reference, whereas for larger values of z it more and more resembles the CASCC curve based on Φ_1.

Despite giving an accurate potential energy curve, the ic-MRCCSD method still features a significant rise of the energy error near the transition state. It turns out that this shortcoming can be essentially cured by including triple excitations [P1], as shown in Fig. 12b. The dramatic improvement coming along with the incorporation of triples motivates the formulation of a perturbative triples correction, to be detailed in chapter 5.

---

[14]The present example is somewhat exceptional due to its small number of orbitals. While there are 1 500 cluster amplitudes in CASCCSD, ic-MRCCSD has 1 616 elements in $\mathbf{t}$ (as defined in Eq. (3.1)). The number of nonredundant elements in $\mathbf{t}'$ (1 043) is indeed smaller. The reader is referred to sec. 4.1.5 for a more representative example.

[15]The reader might wonder why no numerical comparison between ic-MRCC and the internally contracted variant of MRCI is presented. Using only a portion of the variational space of uncontracted MRCI, ic-MRCI is an approximation to MRCI and always yields a higher energy than MRCI. An additional comparison to ic-MRCI results is thus considered as superfluous in the present context. Further, in contrast to ic-MRCI, ic-MRCC should not be viewed as an approximation to any other MRCC method but rather as a method in its own right, as to date there is no uncontracted MRCC method meeting the same formal requirements such as invariance with respect to active orbital rotations.

4.1.2   Hydrogen fluoride bond breaking

Bond breaking processes are classical examples of multireference cases. Labeling the bonding and antibonding spatial orbitals of a system as ϕ_1 and ϕ_2, respectively, one may qualitatively characterize a single bond breaking as a change from the dominant configuration $(\phi_1)^2$ to the negative linear combination of the configurations $(\phi_1)^2$ and $(\phi_2)^2$ with equal weights. The spins of the unpaired electrons on the resulting fragments are entangled, i.e., coupled to a singlet. If the bond breaking is described with a CAS(2,2), the orbitals ϕ_1 and ϕ_2 constitute the natural CASSCF orbitals, since they diagonalize the one-particle density matrix. In the event that both orbitals belong to the same irreducible representation, as for the 3σ and 4σ orbitals of HF, they can be arbitrarily rotated in a symmetry-adapted CASSCF description.

The present data is based on the Dunning–Hay basis set DZV, which is available in the GAMESS program [139] and in Ref. [103], and has been used to describe the HF molecule with a multitude of MRCC approaches [93, 103, 140–142], including ic-MRCCSD [38]. *Pseudocanonical orbitals* are used, which diagonalize the active block of an effective Fock matrix defined as [49, 143]

$$(f_{\text{eff}})_q^p = f_q^p + \sum_{uv} g_{qv}^{pu}\gamma_u^v, \tag{4.1}$$

and no frozen core approximation is applied. In the dissociation limit, the pseudocanonical orbitals localize on the two atoms. The CASSCF wave function then becomes a spin-adapted linear combination of the two determinants that have both orbitals 3σ and 4σ singly occupied, see Fig. 13a.

Figure 13b reveals that ic-MRCCSD is much more accurate than MRCISD and ic-MRCCSDT is even competitive with MRCISDTQ. The *nonparallelity error* (NPE), which is the difference between the maximum and the minimum error along the studied section of the potential energy curve, is only 0.08 mE_h for ic-MRCCSD and just 0.04 mE_h in case of ic-MRCCSDT. While the Mk-MRCCSD and Mk-MRCCSDT results from Ref. [141] agree well with ic-MRCCSD and ic-MRCCSDT in the dissociation limit, they provide a much less accurate description of the bond breaking process, with NPEs as large as 1.9 mE_h and 1.1 mE_h, respectively. In addition, the Mk-MRCC approach suffers from a severe lack of orbital invariance in this example, as quite different potential energy curves have been obtained from Mk-MRCCSD when using natural orbitals (Ref. [103], NPE > 2.1 mE_h) or localized orbitals (Ref. [93], NPE = 1.4 mE_h). It is worth emphasizing that ic-MRCC theory is rigorously invariant with respect to active orbital rotations (see sec. 3.3.1) and thus free of such problems.

4.1.3   Lithium fluoride avoided curve crossing

The dissociation of the lithium fluoride molecule is a challenging multireference problem that exhibits a weakly avoided crossing of the $^1\Sigma^+$ ground state and the lowest excited state of same spatial and spin symmetry. Near the equilibrium bond length, the ground state is dominated by the ionic configuration $1\sigma^2 2\sigma^2 3\sigma^2 4\sigma^2 1\pi^4$, where the 4σ orbital has mainly F $2p_\sigma$ character, and the excited state is essentially covalent, with one electron in the 4σ orbital and another in the 5σ orbital of Li $2s$ type. Upon increasing the internuclear distance, the electrostatic attraction between the ions at some point falls

(a) Potential energy curve (top) and weights of the CASSCF configurations (bottom).

(b) Errors with respect to FCI.
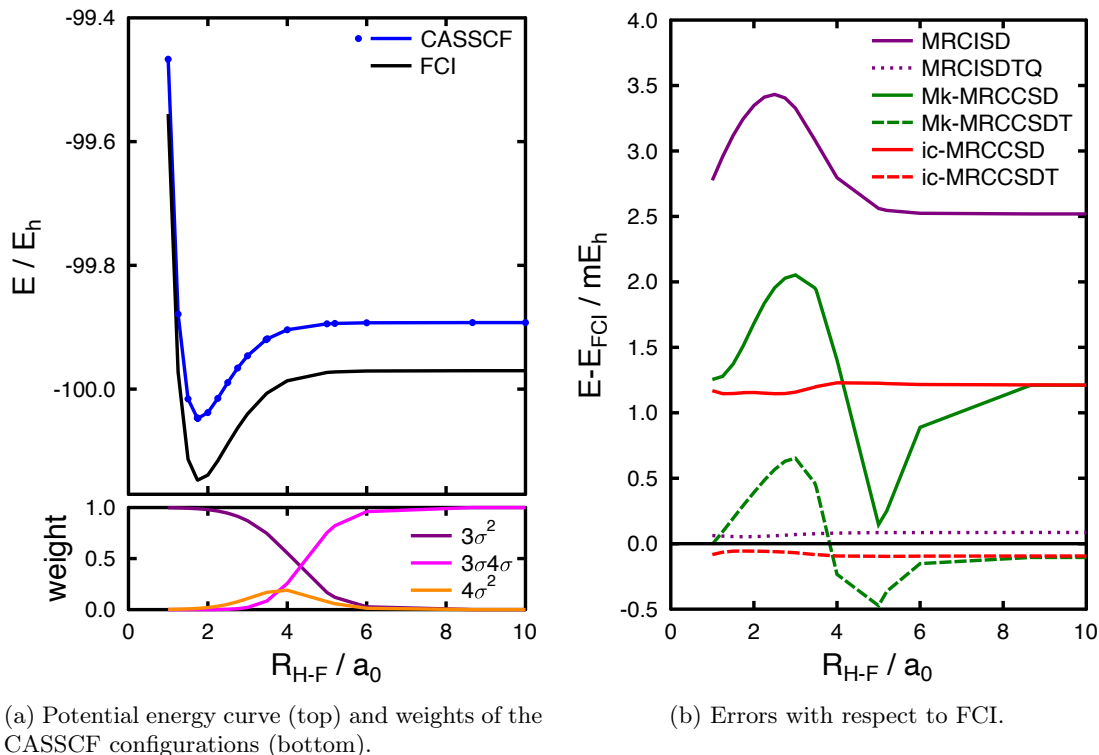
Figure 13: Potential energy surface of the $^1\Sigma^+$ ground state of hydrogen fluoride.

below the difference between the ionization potential of Li and the electron affinity of F, causing one electron to move (or rather 'jump') from the F atom to the Li atom [144–146].

This behavior can be qualitatively captured by a CAS(2,2). However, state-averaged CASSCF (SA-CASSCF) calculations are required in order to avoid discontinuities in the active orbitals along the dissociation process [146,147]. Following a benchmark study by Varandas [148], the results presented in the following are based on using the cc-pVDZ basis set for Li, the aug-cc-pVDZ basis set [149] for F, and a frozen core composed of the three lowest orbitals corresponding to F $1s$, Li $1s$ and F $2s$. Figure 14a reveals that SA-CASSCF underestimates the distance where the switching of the ionic and covalent configurations occurs by about 5.5 a_0 and predicts a more gradual transition than FCI. As a consequence, the SA-CASSCF wave functions have a qualitatively wrong character in the range 8 $a_0 < R_{\text{Li-F}} < 13.5$ a_0 and are thus ill-suited as reference functions in a subsequent MRCC computation.

This problem is solved in ic-MRCC theory by dynamically relaxing the reference function in the presence of the cluster operator. As shown in Fig. 14b, ic-MRCCSD is then able to accurately describe the potential energy surfaces of both states with NPEs below 1.4 mE_h. The relaxed reference functions predict the avoided crossing at virtually the same distance as FCI, namely around $R_{\text{Li-F}} = 13.5$ a_0.

It is also striking that ic-MRCCSD yields very similar energy errors for both the ionic and the covalent configuration, whereas MRCISD is much less accurate for the ionic configuration than for the covalent one (by more than 4 mE_h). One drawback of
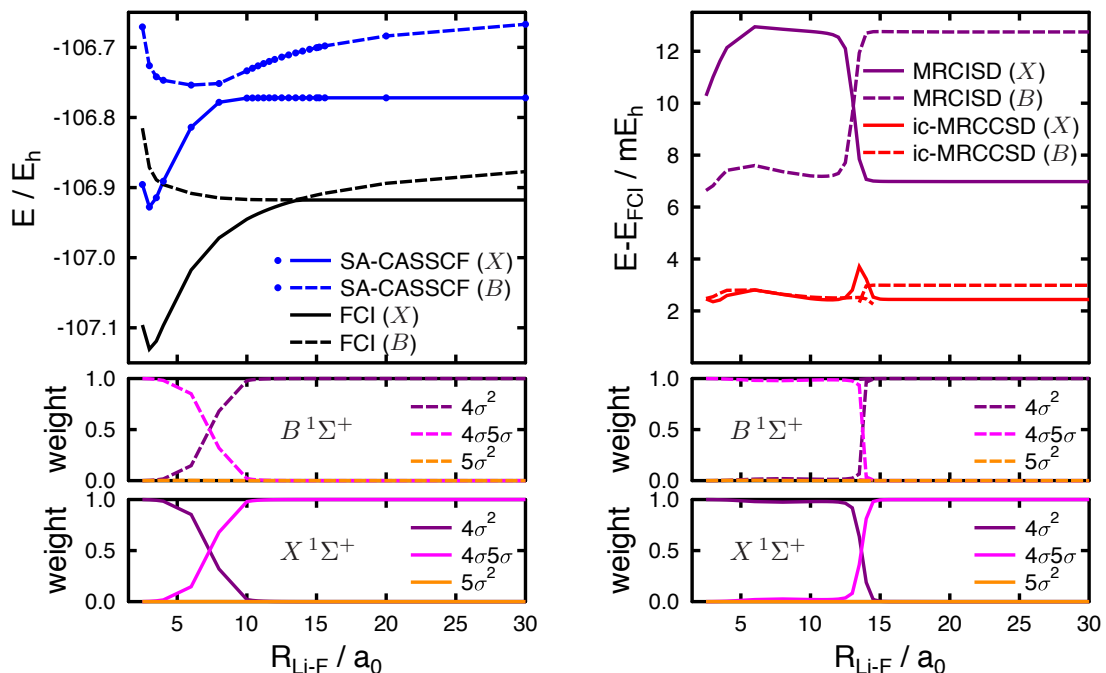
(a) Potential energy curve (top) and weights of the CASSCF configurations (bottom).

(b) Errors with respect to FCI (top) and configuration weights in the ic-MRCCSD reference function (bottom).

Figure 14: Potential energy surfaces of the $X\,^1\Sigma^+$ ground state (X) and the $B\,^1\Sigma^+$ excited state (B) of lithium fluoride.

ic-MRCC, however, is that it does not yield mutually orthogonal states as MRCI does, preventing a straightforward computation of properties depending on both states such as transition moments. The nonorthogonality of the relaxed reference functions for the two states is even visible in the plotted configuration weights at the bottom of Fig. 14b.

Considering the nontrivial task of targeting two states of same symmetry with the ic-MRCC method, a few comments of a more technical nature are in place: Since control over the target state is mainly achieved through a suitable initial guess for the reference function [P1, 37], use of the relaxed reference function from a neighboring point on the potential energy surface is often preferable to starting from the SA-CASSCF wave function. Yet, it is not possible to obtain a continuous energy surface of the excited state with the ic-MRCC method, due to the special nature of the reference function for the separated ions. In the dissociation limit, this reference function becomes exactly a single determinant, with the F $2p_\sigma$ orbital doubly occupied and the Li $2s$ orbital empty. Its one-particle density matrix then becomes singular, which in turn implies redundancy of excitations out of 5σ or into 4σ. Hence, when moving along the excited state surface beyond the point of maximum multireference character, metric eigenvalues associated with amplitudes such as $t_a^{5\sigma}$ steadily approach zero, forcing one to exclude them at some point for numerical reasons. In addition, one now has to include double excitations of the type $\hat{a}_{uv}^{wa}$, because the argument for excluding them, laid out in sec. 3.1.2, is not valid

in this special case.[16] As illustrated by two dashed ic-MRCCSD curves in Fig. 14b, the exclusion of amplitudes belonging to small metric eigenvalues causes a discontinuity (or at best a cusp) in the energy. While this problem is a rare exception for CAS(2,2) cases, small discontinuities along a potential energy surface are an undesired but inevitable feature for most larger active spaces.

4.1.4 Symmetric dissociation of the H_2O molecule

The simultaneous breaking of both O–H bonds in the water molecule in the course of a symmetric dissociation process (Fig. 15) has been extensively used as a model system for benchmarking electronic structure methods [38,41,52,76,134,141,150–153]. The minimal active space is a CAS(4,4) that comprises two sets of bonding and antibonding orbitals, namely $3a_1$, $4a_1$, $1b_2$ and $2b_2$ [76]. Figure 16a displays the CASSCF and FCI potential energy curves for the dissociation process, in which the O–H bonds are symmetrically stretched starting from $R_e = 0.9929$ Å and the H–O–H angle is held fixed at $\theta_e = 109.57°$. The results are based on the cc-pVDZ basis set and on freezing the $1a_1$ orbital (O $1s$).

As shown in Fig. 16b, ic-MRCCSD yields an accurate description of the dissociation, with energy errors below 1.5 mE_h and an NPE of 0.7 mE_h [P5,P3]. The curve is almost smooth despite a changing number of metric eigenvalues above the threshold $\eta = 10^{-6}$. The exclusion of amplitudes that are associated with small metric eigenvalues is necessary, as the active two-particle density matrix becomes singular and causes a number of metric eigenvalues that are still above η at $R_{O–H} = 2.6\ R_e$ to approach zero upon stretching the molecule further. A more detailed analysis of this effect is presented in sec. 4.3.

The open symbols in Fig. 16b represent results for selected competing multireference methods from the literature [41]. These are based on a CAS(6,5) that also includes the $2p_x$ orbital of O. Even though this ensures that all $2p$ orbitals of the isolated O atom are treated as active, the $2p_x$ orbital remains essentially doubly occupied along the reaction. Treating it as inactive therefore hardly affects the results, as Fig. 16b reveals for the case of ic-MRCCSD.

The methods MRCI+Q, MR-ACPF and MR-AQCC are modifications of MRCISD,
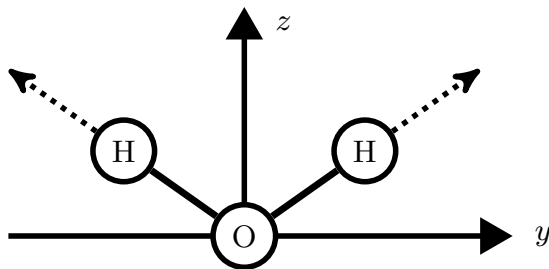
Figure 15: Symmetric dissociation of the water molecule.

---

[16]By virtue of the newly included double excitations, the ic-MRCC method based on a sequential orthogonalization turns into the SRCC method in the event of a single contributing reference determinant. This was verified numerically for $R_{Li–F} = 30\ a_0$. However, note that the double excitations in question are again redundant in the dissociation limit, since excitations from the active orbital localized on F to the active orbital on Li are ruled out by the size-consistency of SRCC theory.
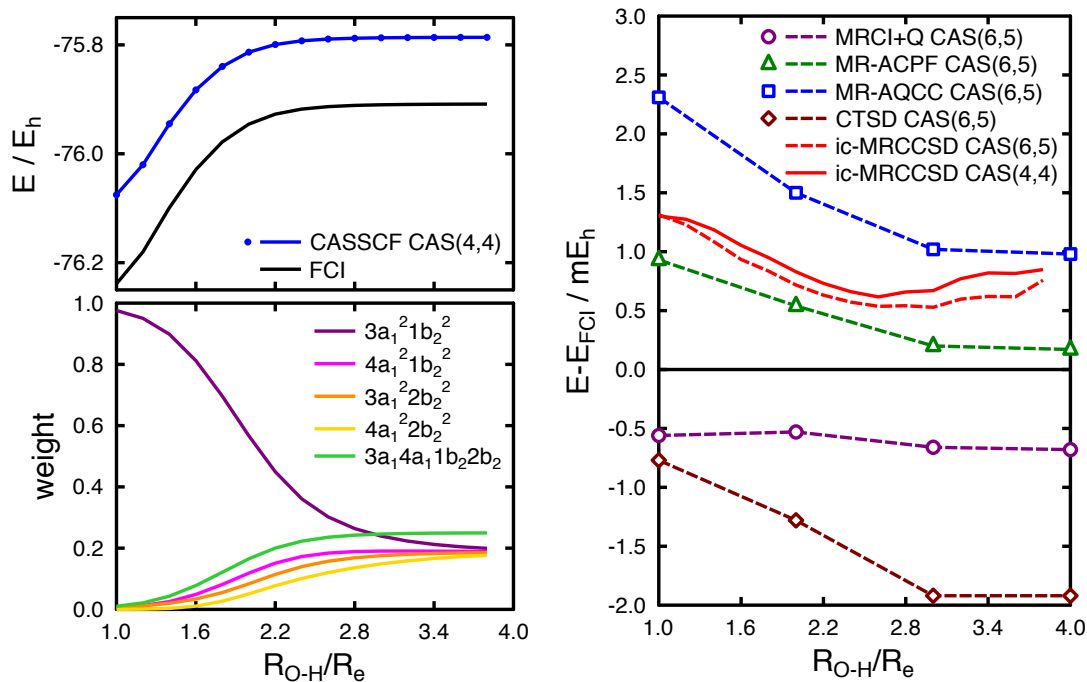
(a) Potential energy curve (top) and weights of the CASSCF configurations (bottom).

(b) Errors with respect to FCI.

Figure 16: Potential energy along the symmetric dissociation of the $^1A_1$ ground state of H_2O.

aimed at approximately restoring size-extensivity. They achieve this by estimating energy contributions from the missing disconnected higher excitations, most importantly from $\hat{T}_2^2$. While MRCI+Q accounts for these contributions through Davidson's *a posteriori* correction [154, 155], the averaged coupled-pair functional (MR-ACPF, Ref. [133]) and the so-called averaged quadratic coupled-cluster theory (MR-AQCC, Ref. [156]) rely on *a priori* modifications of the MRCISD energy functional. MR-ACPF rests upon the assumption that the system consists of $N/2$ noninteracting electron pairs and yields exact results for this special case. However, both MRCI+Q and MR-ACPF generally tend to overcorrect for the effect of disconnected excitations [134, 157] and thereby often give energies beneath or close to FCI, as in the present example (see also sec. 4.1.5). In contrast to MR-ACPF, the related MR-AQCC method considers all electrons as interacting and distributes the correlation energy in an averaged way over all $\binom{N}{2}$ electron pairs. Although not always yielding better results than MR-ACPF, MR-AQCC is known to be more reliable and is usually preferred [25]. The size-extensivity corrected MRCI methods are competitive with MRCC methods for model systems, and indeed MRCI+Q outperforms ic-MRCCSD in the present example, but the remaining lack of size-extensivity renders these methods inferior to a size-extensive MRCC theory for larger systems (e.g., Ref. [118]).

The linear canonical transformation theory with singles and doubles (L-CTSD or, for brevity, CTSD) is closely related to ic-MRCCSD. It is an internally contracted approach that uses a unitary wave operator in conjunction with the rather radical approximation to

neglect the effective three-body component of each intermediate commutator expression in the BCH expansion [40–43]. The method is denoted as linear when this approximation is applied at each commutator level, i.e., after each linear transformation of the (effective) Hamiltonian. In addition, due to numerical reasons a rather large threshold of $\eta = 10^{-2}$ for discarding small metric eigenvalues is used in CTSD. These approximations are likely the reason for CTSD's overshooting the energy by up to 1.9 mE_h and for the comparatively large NPE of 1.2 mE_h.

4.1.5 Nitrogen triple bond breaking

The triple bond breaking of the N_2 molecule is a prime example of a reaction with strong static correlation, which involves six electrons in the orbitals $3\sigma_g$, $3\sigma_u$ as well as the orbital pairs $1\pi_u$ and $1\pi_g$, hence requiring a CAS(6,6). Figure 17a shows the energy surface of the $^1\Sigma_g^+$ ground state as computed with CASSCF and FCI for the cc-pVDZ basis set and a frozen core consisting of the two orbitals $1\sigma_g$ and $1\sigma_u$ (N $1s$). The FCI calculations were performed with the LUCIA code [138]. At the equilibrium distance (around $R_{N-N} = 2.1$ a_0), the wave function is dominated by the closed-shell configuration $3\sigma_g^2 1\pi_u^4$ (with inactive orbitals omitted). Stretching the molecule first leads to the breaking of the π bonds, indicated by a large contribution from the configuration $3\sigma_g^2 1\pi_u^2 1\pi_g^2$, until the system finally dissociates into two N atoms in their ^4S ground state, with the electronic spins still coupled to a singlet.

Figure 17b compares the performance of ic-MRCCSD with MRCI-based methods
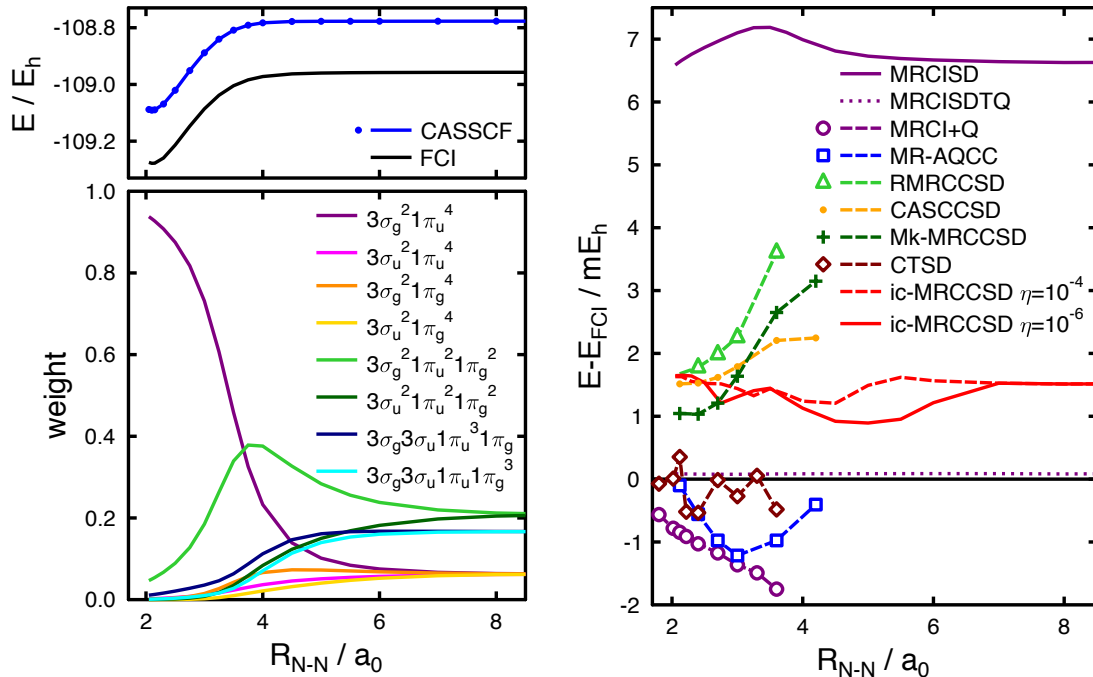
(a) Potential energy curve (top) and weights of the dominant CASSCF configurations (bottom).

(b) Errors with respect to FCI.

Figure 17: Potential energy surface of the $^1\Sigma_g^+$ ground state of nitrogen.

and several MRCC methods from the literature. The MRCISDTQ curve with an almost constant error below 0.1 mE_h is in excellent agreement with FCI and was therefore used as benchmark in Refs. [P1,P3]. With a computational scaling like N^{10}, MRCISDTQ is of course far too expensive for most larger applications. While MRCISD is substantially less accurate than most MRCC methods in absolute terms, its error is mainly systematic and remains around 7 mE_h with an NPE of only 0.6 mE_h. The values from both MRCI+Q [42] and MR-AQCC [141] overshoot the FCI energies. They exhibit a much smaller absolute error than MRCISD, but at the expense of an NPE increased to at least 1.1 mE_h.

All of the discussed MRCC methods provide a reasonable description around the equilibrium bond length, but the curves for RMRCCSD [158] (also cf. Ref. [159]), CASCCSD [158,160] and Mk-MRCCSD [93] tend to become inaccurate when stretching the bond beyond $R_{\text{N–N}} = 3 \, a_0$. It is worth emphasizing that the Mk-MRCCSD energies are extremely sensitive to active orbital rotations. The plotted results are based on localized orbitals, which is the most reliable choice to date [93], whereas use of pseudo-canonical orbitals yields energies higher by up to 15 mE_h [141]. Both RMRCCSD [72] and CASCCSD employ a single determinant with occupancy $3\sigma_g^2 1\pi_u^4$ as a reference function and model excitations from other active space configurations by higher-rank excitation operators. While RMRCCSD derives the values of the 97 440 additional amplitudes in $\hat{T}_3$ and $\hat{T}_4$ from the MRCISD wave function and subsequently solves the equations for the 2 089 singles and doubles amplitudes, CASCCSD solves the CC equations for up to selected octuple excitations and therefore uses the same number of parameters as MRCISD, namely 241 755. With this amount of flexibility in the wave function, CASCCSD is able to accurately describe the N_2 triple bond breaking with an NPE of about 0.7 mE_h up to $R_{\text{N–N}} = 4.2 \, a_0$. A similar accuracy is achieved by ic-MRCCSD, however based on a much more compact parameterization that consists of 6 714 singles and doubles amplitudes as defined by Eq. (3.1).

The results from the related linear canonical transformation theory [42] fluctuate around the exact solution, but it is unclear to which extent this accuracy is due to fortuitous compensation of errors from the various approximations used in CT theory. The large thresholds for discarding metric eigenvalues ($\eta = 10^{-1}$ for single excitations and those double excitations involving three active indices, and $\eta = 10^{-2}$ otherwise) are likely the origin for rapid changes in the energy error of up to 0.9 mE_h for a bond length increased by 0.1 a_0. The ic-MRCCSD curve is much smoother, although irregularities caused by a truncation of the metric eigenvalues are visible. The remarkable finding that the ic-MRCCSD curve improves upon *raising* the threshold to $\eta = 10^{-4}$ indicates that amplitudes that belong to very small metric eigenvalues can have a deteriorating effect on the computed energies [P1,51]. The impact of η is more extensively studied in sec. 4.3.

4.1.6 Symmetric dissociation of linear Be$_3$

The last model system presented here is a cut section of the ground state energy surface of the beryllium trimer, Be–Be–Be, which is confined to be linear with two identical internuclear distances $R_{\text{Be–Be}}$ [161]. Since the true equilibrium structure of Be$_3$ is not linear but a unilateral triangle, the studied process is not related to a physical dissociation channel. Nevertheless the model offers a critical test for electronic structure methods, as the potential wells encountered along the curve result from bonding situations quite different from the ones in the previous examples. Because the ground state of the beryllium atom
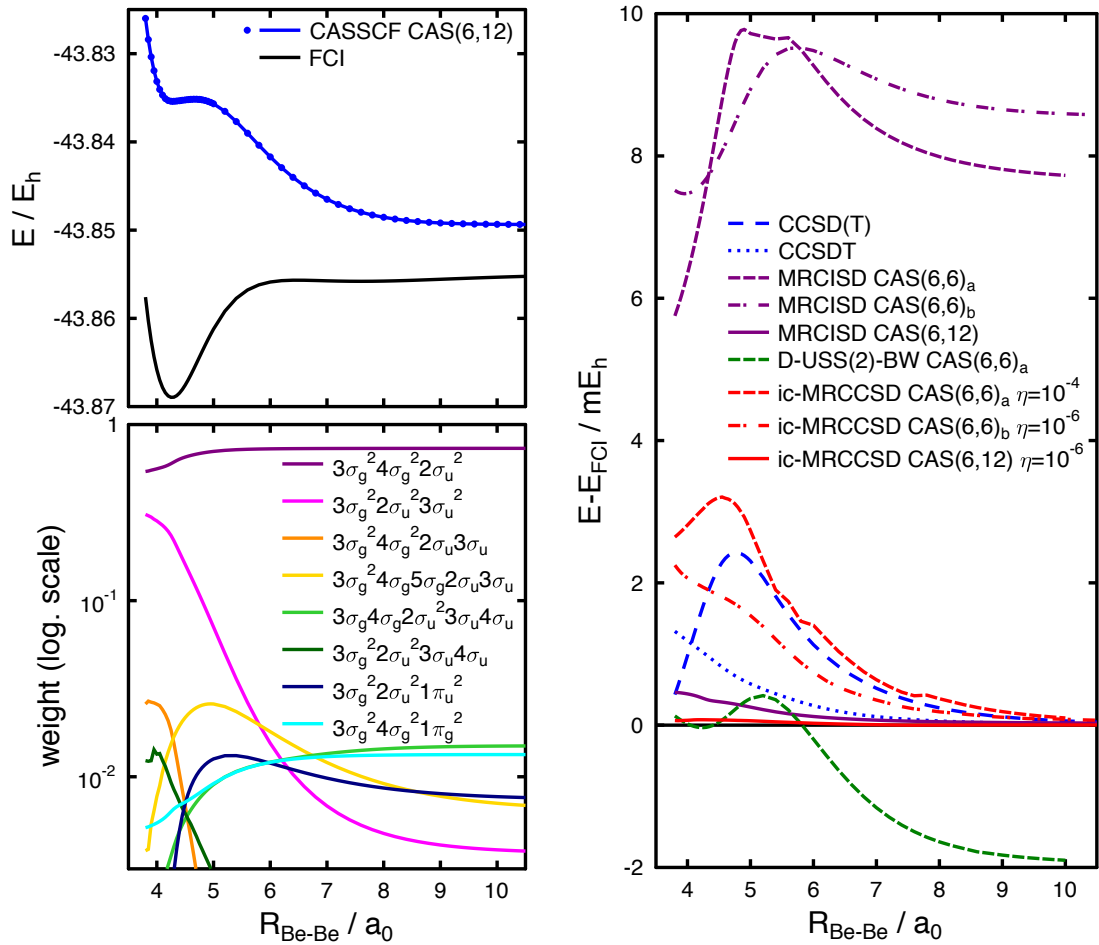
(a) Potential energy curve (top) and weights of the dominant CASSCF configurations (bottom).

(b) Errors with respect to FCI.

Figure 18: Potential energy surface of a $^1\Sigma_g^+$ dissociation path of linear Be$_3$.

is dominated by the closed-shell configuration $1s^2 2s^2$, the bonding in Be clusters cannot be described within a mean-field approach. One may understand the bonding between Be atoms as being facilitated by a hybridization of the $2s$ orbitals and the energetically low-lying $2p$ orbitals. The resulting potential energy surfaces are rather flat and to a large extent shaped by electron correlation [162, 163].

The FCI energy curve displayed in Fig. 18a [161] has been computed using a $[3s2p1d]$ ANO basis set [164], keeping the lowest three RHF orbitals $1\sigma_g$, $2\sigma_g$ and $1\sigma_u$ frozen. Notably, the ic-MRCC formalism simplifies dramatically for this choice, as there are no correlated doubly occupied orbitals left, reducing the number of terms in the ic-MRCCSD equations to a few hundred, see Tab. 2. All of the present CASSCF calculations are based on the same frozen core approximation at the RHF level, in order to allow a direct comparison of the computed energy curves with the FCI benchmark. As the $2p$ orbitals are essential for bonding and all of them are degenerate in the dissociation limit, an appropriate choice for an active space comprises all of the twelve valence orbitals, i.e.,

$3\sigma_g$, $4\sigma_g$, $5\sigma_g$, $2\sigma_u$, $3\sigma_u$, $4\sigma_u$ and the pairs $1\pi_u$, $2\pi_u$ and $1\pi_g$. An inspection of selected configurations' weights in the CASSCF wave function (Fig. 18a, note the logarithmic scale) reveals that the configuration $3\sigma_g^2 4\sigma_g^2 2\sigma_u^2$ (with core orbitals omitted) remains dominant throughout the model reaction, with most other configuration weights not exceeding a few percent.

Despite the size of the active space, the CASSCF energy curve is essentially unbound, emphasizing the importance of dynamic correlation for a bonding between the Be atoms. Besides the minimum at $R_{\text{Be–Be}} = 4.25\ a_0$, the FCI curve features a very shallow van der Waals minimum at $R_{\text{Be–Be}} = 7.6\ a_0$, the prediction of which requires an accurate treatment of dynamic correlation as well. The model also offers a simple test for size-consistency, as all of the methods discussed below are exact for two-electron systems and should agree with FCI in the dissociation limit if they were size-consistent.

The fact that a closed-shell determinant dominates the CASSCF wave function raises the question whether single-reference methods can provide a satisfactory description of the Be$_3$ model. While CCSD is certainly not sufficient [165] and gives errors larger than the total height of the potential energy well (see Tab. A11), good results are achieved with CCSD(T) (cf. Ref. [162]) and CCSDT, see Fig. 18b.

Brabec *et al.* used this model system to assess the performance of an *a posteriori* correction to Brillouin–Wigner MRCC termed D-USS(2)-BW [166]. It is a diagonal correction employing Kowalski's universal state-selective MRCC functional [167], which is applicable to all Jeziorski–Monkhorst-based MRCC methods and is related to Λ corrections in single-reference CC theory (see Ref. [168] for a review). While approaches of this kind are commonly used to incorporate the effect of triple excitations, the D-USS(2) correction is restricted to single and double excitations and aims at alleviating shortcomings caused by the introduction of sufficiency conditions, such as the lack of size-extensivity in BW-MRCC theory. Figure 18b shows the D-USS(2)-BW results from Ref. [166] that use an active space consisting of $3\sigma_g$, $4\sigma_g$, $2\sigma_u$, $3\sigma_u$ and $1\pi_u$, to be abbreviated as CAS(6,6)$_a$. The D-USS(2)-BW values rely on RHF orbitals while CASSCF orbitals were used for all other multireference approaches discussed in this section. Although being much more accurate than MRCISD based on a CAS(6,6)$_a$ and in excellent agreement with FCI at small internuclear distances, the D-USS(2)-BW results reveal a pronounced size-consistency error of -1.9 mE_h.

This problem is not shared by ic-MRCCSD, which exhibits negligible energy errors for large values of $R_{\text{Be–Be}}$. Yet the CAS(6,6)$_a$ is not large enough to provide a highly accurate description of the bonding region with ic-MRCCSD. A small improvement is obtained by switching to an alternative active space that includes all the σ orbitals, i.e., $3\sigma_g$, $4\sigma_g$, $5\sigma_g$, $2\sigma_u$, $3\sigma_u$ and $4\sigma_u$, denoted as CAS(6,6)$_b$. This active space is more balanced, as it at least preserves the degeneracy among p_σ and p_π orbitals in the dissociation limit. It was also found to facilitate the convergence of the ic-MRCCSD equations, which was so poor in the CAS(6,6)$_a$ calculations that solving the ic-MRCC equations in that case required an increased metric eigenvalue threshold of $\eta = 10^{-4}$.

The excellent results for the CAS(6,12) support that this is indeed the proper choice of an active space for this example. Even MRCISD achieves very high accuracy and an NPE below 0.5 mE_h. Its performance is, however, clearly surpassed by ic-MRCCSD, which features both absolute and nonparallelity errors below 0.08 mE_h.

4.2 Commutator approximations

As discussed in sec. 3.1.2, truncations of the commutator series in the BCH expansion, Eq. (3.9), are indispensable for an efficient realization of the ic-MRCC method. Their justification in ic-MRCC theory rests on the ability of the reference function to describe the exact wave function qualitatively well, such that cluster amplitudes can stay small ($\ll 1$) and terms with many $\hat{T}$ operators may be neglected [35]. Evangelista and Gauss have carried out a thorough investigation of such commutator approximations and came to the conclusion that terms quadratic in $\hat{T}$ are essential for preserving the inherent accuracy of ic-MRCC theory, whereas the neglect of higher than double commutators is a viable approximation [38].

This is supported by results for the BeH_2 and N_2 model systems, presented in the following. Figure 19 demonstrates that a truncation at the single commutator level greatly deteriorates the accuracy of the ic-MRCCSD method and tends to overshoot the FCI energies, in this case by up to 6 mE_h. This approximation turns ic-MRCC theory into a method of the coupled electron-pair approximations (CEPA) type, see Refs. [50, 169, 170] and Ref. [171] for a review. Figure 19a uses individual restrictions on the maximum commutator level for the amplitude equations and the eigenvalue equations yielding the energy, denoted as $N_{com}(\Omega)$ and $N_{com}(E)$, respectively. An exact evaluation of the eigensystem, i.e., by setting $N_{com}(E) = 4$, is usually affordable, as the relatively few additional terms (see Tab. 1) cause little extra effort. This leads to substantial changes close to the transition state in the BeH_2 model but fails to restore the accuracy of the method. Therefore it is crucial to consider up to double commutators in the amplitude equations. In the case of the BeH_2 model, the exact evaluation of the eigensystem then provides a moderate improvement of the ic-MRCC energy curves, lowering the NPEs by about 0.1 mE_h for both ic-MRCCSD and ic-MRCCSDT [P3].

The application to the N_2 triple bond breaking (Fig. 19b) reveals that the linearized
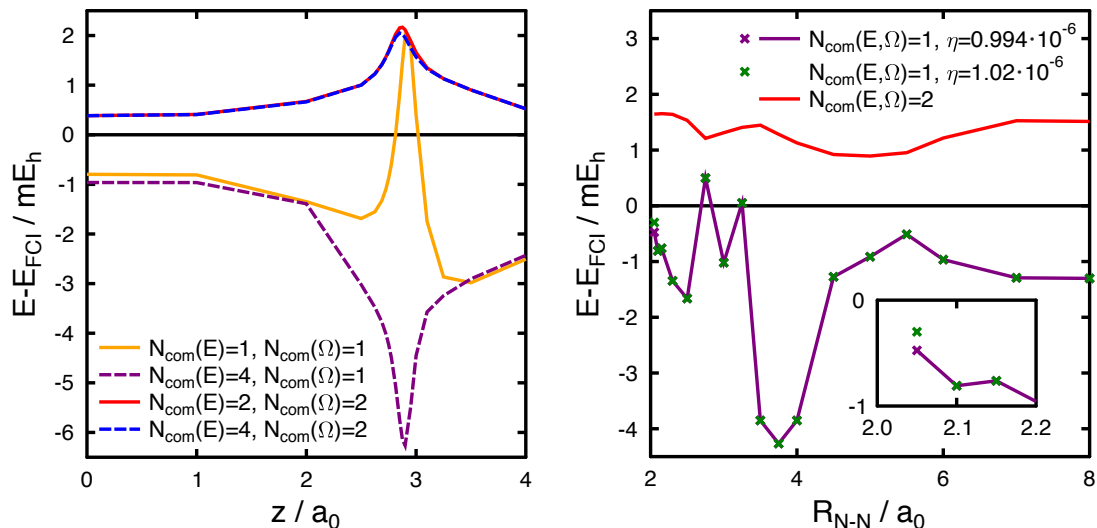
(a) Errors with respect to FCI for BeH_2.   (b) Errors with respect to FCI for N_2.

Figure 19: Impact of linear commutator approximations on ic-MRCCSD energies.
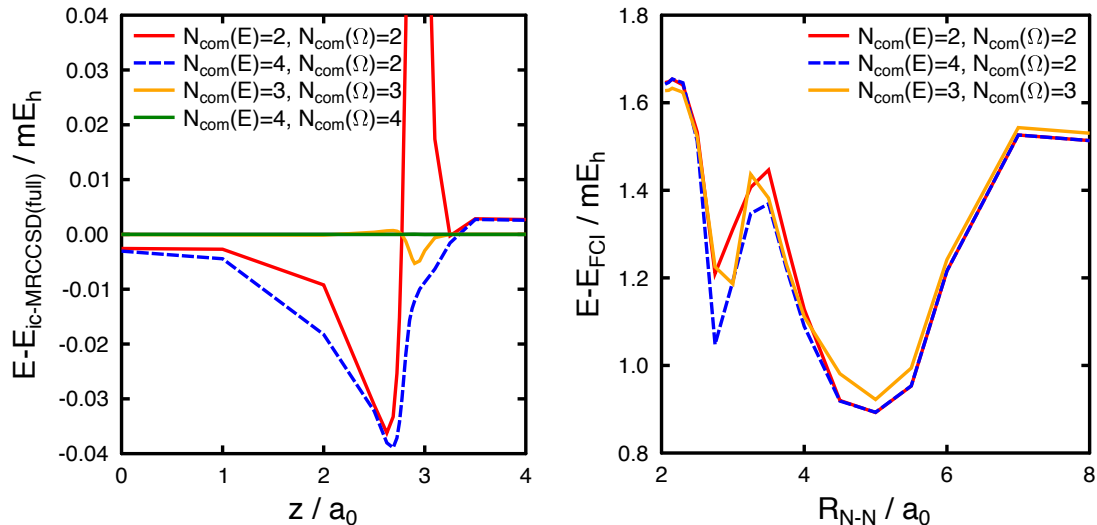
(a) Errors with respect to ic-MRCCSD with $N_{\mathrm{com}}(E){=}4$, $N_{\mathrm{com}}(\Omega){=}8$ for BeH$_2$.

(b) Errors with respect to FCI for N$_2$.

Figure 20: Impact of neglecting higher commutators on ic-MRCCSD energies.

ic-MRCC theory yields a much more irregular curve, possibly owing to an increased sensitivity to a varying number of metric eigenvalues above η. The fluctuations around the FCI curve up to $R_{\mathrm{N-N}} = 3$ a_0 are reminiscent of the performance of CTSD (see Fig. 17b). Indeed, the neglect of three-body components of commutators in CTSD induces substantial errors at the twofold commutator level [43] and could therefore be a major contributing factor to the error compensation in CTSD's total energies.

Note the occurrence of a rare technical problem in ic-MRCCSD with $N_{\mathrm{com}}(E, \Omega) = 1$ for the internuclear distance $R_{\mathrm{N-N}} = 2.05$ a_0: Setting $\eta = 10^{-6}$ leads to an oscillation between two solutions when converging the ic-MRCC equations. The two results differ in the number of included metric eigenvalues and are both displayed in the inset of Fig. 19b. The solution with $\eta = 1.02 \cdot 10^{-6}$ excludes certain metric eigenvalues that are still above 10^{-6} but move slightly below 10^{-6} upon their inclusion. As a consequence, there is no stable solution for the choice $\eta = 10^{-6}$.

Figure 20 compares ic-MRCCSD approximations that include twofold or higher commutators. For BeH$_2$, a direct comparison to the untruncated ic-MRCCSD method involving up to eightfold commutator terms is possible, as the number of terms in the amplitude equations for a CAS(2,2) is still manageable (see Tab. 2). The maximum absolute error caused by truncation at the double commutator level is about 144 μE_h and reduces to 39 μE_h when solving the eigensystem exactly. The rapid convergence of the commutator series is evident from the decrease of this error to about 5 μE_h for $N_{\mathrm{com}}(E, \Omega) = 3$ and to well below 0.05 μE_h for $N_{\mathrm{com}}(E, \Omega) = 4$.

For the N$_2$ system treated with a CAS(6,6), calculations including higher commutators quickly become intractable due to an exploding number of terms in the amplitude equations (see Tab. 1). Figure 20b shows that the neglect of threefold commutators affects the ic-MRCCSD energies by tens of μE_h in most parts of the potential energy

curve, and up to 0.2 mE_h around $R_{\text{N–N}} = 3\ a_0$ where the gradient of the total energy is largest. Truncation at twofold commutators in this example hardly impacts the accuracy of ic-MRCCSD. There is also no clear advantage of an exact treatment of the eigensystem for relaxing the reference function.

Overall, the quadratic approximation of ic-MRCC theory can be expected to preserve the inherent accuracy of the method, although an (approximate) incorporation of three-fold commutators might be needed when aiming at very high precision in the order of 0.1 kcal/mol.

4.3 Truncation of the excitation basis

In the calculations discussed so far, the threshold η for discarding small metric eigenvalues was kept at 10^{-6} in most cases. The following analysis aims at providing some idea about the impact of this threshold on the numerical results. The N_2 model system is well suited for this purpose, as it relies on a rather large active space and features metric eigenvalues that change dramatically along the dissociation process. Based on the converged ic-MRCCSD calculations with $\eta = 10^{-6}$ for this system, Fig. 21 shows the number of metric eigenvalues falling between 10^{-x} and 10^{1-x} (with an integer exponent x) for each computed point on the energy curve. Since only purely active blocks of the metric are diagonalized (see sec. 3.2), the plotted numbers of eigenvalues correspond to these active blocks. The reader may think of Fig. 21 as a sequence of histograms for each value for $R_{\text{N–N}}$, which are extracted by slicing the diagram parallel to the x-axis. A number of 133 eigenvalues are numerically zero (i.e., below 10^{-14}, not shown in Fig. 21) throughout the reaction and correspond to exact linear dependencies.

For most internuclear distances, there is a clear-cut gap between significant ($> 10^{-10}$) and numerically vanishing eigenvalues. In principle it would be desirable to place the threshold η into this gap, in order to avoid the arbitrariness involved in discarding a number of nonredundant eigenvalues by setting it somewhere above the gap. It is evident from Fig. 21 that the choice $\eta = 10^{-6}$ complies with this condition only for the points $R_{\text{N–N}} = 3\ a_0$ and between $R_{\text{N–N}} = 3.5\ a_0$ and $R_{\text{N–N}} = 4\ a_0$. The strategy to choose η unambiguously, e.g., by setting $\eta = 10^{-10}$, was tested, but faced severe limitations on both ends of the studied curve [P1].

The first limitation is immediately obvious from Fig. 21: When increasing the separation between the atoms beyond $R_{\text{N–N}} = 4\ a_0$, several eigenvalues start moving towards the region of numerical noise. This happens because γ_2 becomes singular upon dissociation. In the dissociation limit, the eigenvalues in question are exactly zero. However, the contribution of the corresponding cluster amplitudes to the energy does not necessarily vanish when approaching the dissociation limit. This is because these amplitudes might incorporate determinants into the wave function that are otherwise only accessible through higher-rank excitations. A numerically stable algorithm for solving the ic-MRCC equations must at some point exclude near-linear dependencies and will therefore yield certain energy changes in the process. The rise in the energy error for ic-MRCCSD with $\eta = 10^{-6}$ between $R_{\text{N–N}} = 5\ a_0$ and $R_{\text{N–N}} = 7\ a_0$ visible in Fig. 17b can be attributed to this effect. In the curve for $\eta = 10^{-4}$, this transition occurs before $R_{\text{N–N}} = 6\ a_0$, which helps to maintain the error with respect to FCI nearly constant.

A similar feature exists in the potential energy curves for the H_2O model system, as can be seen by comparing the results for $\eta = 10^{-6}$ and $\eta = 10^{-4}$ displayed in Fig. 22a.
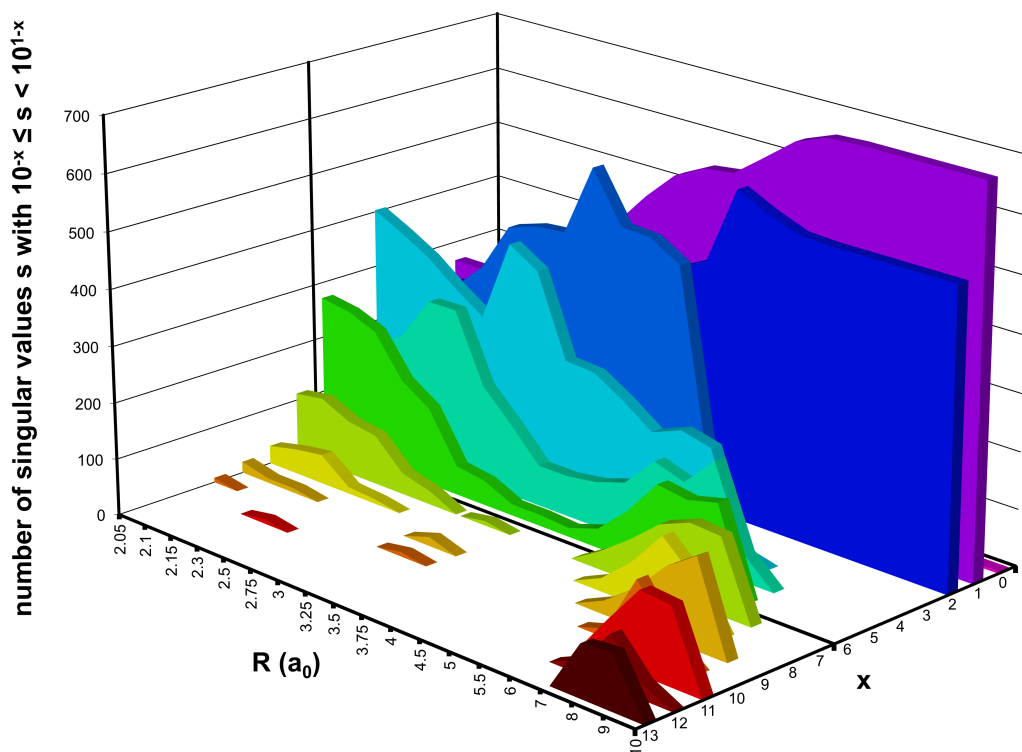
Figure 21: Histogram of eigenvalues for the active part of the metric matrix along the dissociation of N_2. The coordinate axis is not to scale.

Increasing η beyond 10^{-4} does not lead to further improvements, however. While $\eta = 10^{-3}$ still provides a reasonably accurate energy curve, setting $\eta = 10^{-2}$ deteriorates the energies by up to about 1.5 mE_h, especially around the equilibrium structure.

Figure 22b shows in more detail how the ic-MRCCSD energy depends on the threshold η for selected points on the N_2 energy surface ranging from the equilibrium structure to $R_{N-N} = 3\ a_0$. The energies remain most stable when setting η between 10^{-5} and 10^{-4}. Increasing η to 10^{-3} causes the energies to rise by up to 0.4 mE_h. Even larger thresholds yield energy changes of more than a mE_h and therefore have a detrimental impact on the accuracy of ic-MRCCSD.

Attempts to decrease η down to values that guarantee inclusion of all nonredundant metric eigenvalues work well for $R_{N-N} = 3\ a_0$ and $R_{N-N} = 2.75\ a_0$. At the latter point, however, a rather large energy shift of -0.4 mE_h occurs when going from $\eta = 10^{-5}$ to $\eta = 3 \cdot 10^{-7}$. For smaller internuclear distances, the strategy to use the smallest possible value for η encounters its second severe limitation: Small metric eigenvalues start to have enormous influence on the energy, as exemplified by sudden energy drops below $\eta = 10^{-6}$ for $R_{N-N} = 2.5\ a_0$. The metric eigenvalues that are responsible for this feature appear to decrease when approaching the equilibrium structure, as the energies for $R_{N-N} = 2.3\ a_0$ and $R_{N-N} = 2.1\ a_0$ remain stable down to $\eta = 10^{-7}$ and $\eta = 3 \cdot 10^{-9}$, respectively. Their effect on the energy becomes worse, though, as indicated by the dotted lines in Fig. 22b.
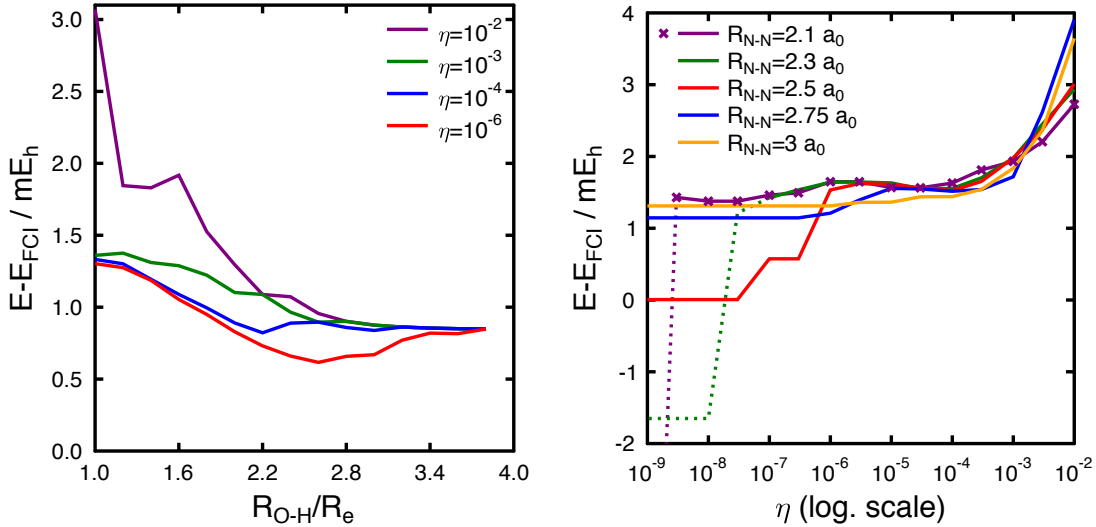
(a) Different choices of η in the H_2O model.

(b) Dependence of the energy error on η for selected points on the energy surface of N_2.

Figure 22: Impact of the threshold η on ic-MRCCSD energies.

No convergence was obtained in this region, probably due to a sensitive interdependence between amplitudes associated with small eigenvalues and the reference function. The results displayed in Fig. 22b are therefore based on fixing the reference coefficients to their values from the last converged calculation and solving the equations for the cluster amplitudes only (cf. sec. 4.5).

It is worth mentioning that such unphysical energy contributions do not arise in the internally contracted variant of MRCI. This was verified numerically for $R_{N-N} = 2.3\ a_0$. Also, the ic-MRCI energy difference between $\eta = 10^{-3}$ and $\eta = 10^{-10}$ at $R_{N-N} = 6\ a_0$ is only in the order of $1\ \mu E_h$, in sharp contrast to the sensitive behavior of ic-MRCC in this region. In order to rationalize the robustness of ic-MRCI, we recall that ic-MRCI is invariant with respect to the procedure used to eliminate redundant amplitudes. Within the sequential orthogonalization technique, exclusion of a small metric eigenvalue s_ρ at a given excitation rank gives rise to an additional nonredundant excitation at the next highest rank in the same excitation class, usually associated with a much larger metric eigenvalue. Hence there are two equivalent ways to include the corresponding excited function $\hat{\tau}'^\rho |\Psi_0\rangle$, both of them leading to the same coefficient c'_ρ due to the linear parameterization of the wave function. The only difference between the two possibilities occurs in the CI coefficients in the nonorthogonal excitation basis, since

$$\hat{\tau}'^\rho |\Psi_0\rangle c'_\rho = \sum_\sigma \hat{\tau}^\sigma |\Psi_0\rangle U^\rho_\sigma s_\rho^{-\frac{1}{2}} c'_\rho \tag{4.2}$$

involves the inverse square root of s_ρ. In this way, a small metric eigenvalue s_ρ may trigger large contributions to the CI coefficients. If we include all ranks for the considered excitation classes (until the active space is 'exhausted'), ic-MRCI becomes equivalent to uncontracted MRCI. In this limit, η has no influence on the number of included excitations but merely fine-tunes which excitation ranks are used to parametrize the variational

space. The internally contracted approximation is thus identical to a truncation of the excitation rank (within a certain set of excitation classes), and the negligibility of small metric eigenvalues is closely connected to the quality of this approximation, for which there is sufficient numerical support [48].

In ic-MRCC theory, one may also expect a decreasing importance of higher excitation ranks within the same excitation class (see sec. 5.6.1 for numerical evidence), which should go hand in hand with a low impact of excitations with small metric eigenvalues. The latter is clearly not always the case, likely because the cluster operator also appears on the left of $\hat{H}$ and in higher powers, enabling large cluster amplitudes to have significant influence on other residual elements and the energy [51]. However, the detailed mechanism behind this effect is not well understood to date.

Overall, the present results suggest that the threshold for discarding small metric eigenvalues should be chosen between 10^{-6} and 10^{-3} in order to prevent severe losses of the accuracy of ic-MRCC theory. Note that this range of acceptable values for η lies lower than the thresholds currently used in canonical transformation theory, for which a similar analysis has been performed [42].

4.4 Removal of linear dependencies between operators of different rank

As described in sec. 3.2, all the procedures to remove linearly dependent cluster amplitudes used in this work rely on canonical orthogonalizations of active blocks of the metric matrix. Within this approach, however, there is some flexibility regarding the treatment of linear dependencies between operators of differing rank. This section compares the following four selected techniques: (i) The complete omission of low-rank excitations in favor of the highest possible rank in each excitation class is allowed by virtue of Eq. (3.46) and will be denoted as 'highest ranks only' scheme. In the case of ic-MRCCSD, this method employs double excitation operators only, as is done in second-order complete active space perturbation theory (CASPT2, Ref. [49]). (ii) The 'full orthogonalization' procedure is the one adopted by Evangelista and Gauss [38]. It does not distinguish between operators of different rank and mixes all ranks within the same excitation class. (iii) The default choice from the previous sections is the 'sequential orthogonalization', which avoids a mixing of different ranks by sequentially orthonormalizing higher excitations to the ones with lower ranks. The three schemes (i) to (iii) use operators normal ordered with respect to the core determinant $|0\rangle$. They will be contrasted to (iv) the 'GNO-based' technique described in sec. 3.2.2. The current implementation of this scheme for ic-MRCCSD also makes use of a full orthogonalization of the underlying metric matrix.

Figure 23 compares the various schemes' performances for the BeH$_2$ model system. The results for procedures (i) to (iii) are from Ref. [P1] and involve no commutator approximation in the eigensystem for the c_μ coefficients. Figure 23a shows the relative energy differences between the four procedures for ic-MRCCSD, arbitrarily taking the GNO-based scheme with $N_{\text{com}}(E) = 4$ as a reference. All methods agree within 20 μE_h. An explanation for this is the fact that the main difference between the four schemes lies in their treatment of (effective) single excitations, which are less important than double excitations as they arise at higher order in perturbation theory (see sec. 5.3.2).

The situation changes dramatically when including triple excitations, as is obvious from Fig. 23b. The significant errors for schemes (i) and (ii), which even exceed those from ic-MRCCSD for large values of z, are a strong argument against these two procedures,
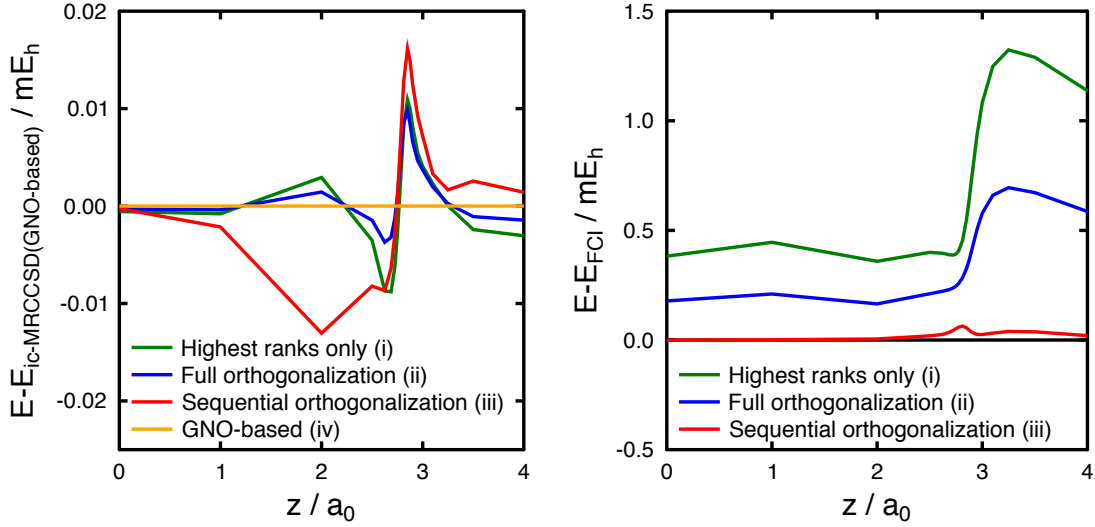
(a) Energy difference between various ic-MRCCSD schemes and the GNO-based one (reference).

(b) Energy errors for ic-MRCCSDT methods.

Figure 23: Comparison between different schemes to remove redundancies for the BeH_2 model.

whereas the results for the sequential orthogonalization are excellent (NPE = 70 μE_h). In order to understand this behavior, it is worth to have a look at the excitation operators included in the various schemes, illustrated in Fig. 24.[17] Whereas scheme (i) uses only the operator with highest rank in each occupation class, and the full orthogonalization employs all of the displayed operators, the operators drawn in blue in Fig. 24 turn out to be redundant for a CAS(2,2) and are completely projected out in a sequential orthogonalization. Of course, they can then be excluded from the start, greatly simplifying the ic-MRCC equations. Among the schemes based on a particle-hole normal ordering,

Figure 24: Excitations in the cluster operator for ic-MRCCSDT in the case of the BeH_2 model.

---

[17]Note that no operator with three hole lines is present due to the frozen core approximation that leaves only one doubly occupied orbital correlated. Operators annihilating three active electrons are not possible for a CAS(2,2) and hence also excluded.

the sequential orthogonalization is thus not only the most accurate but also the most efficient one.

The bad performance of schemes (i) and (ii) for ic-MRCCSDT is mainly due to an incorrect treatment of inactive double excitations $\hat{a}_{ij}^{ab}$, which also contributes to a lack of core-extensivity for these methods [P1, 32]. In scheme (i), such double excitations are completely omitted in favor of the excitations $\hat{a}_{uij}^{vab}$. The latter excitations are even used to mimic double excitations on a noninteracting closed-shell subsystem, but the corresponding cluster amplitudes cannot be extensive quantities anymore, as the active orbitals labeled by u and v are localized on the distant open-shell fragment. This case occurs in the dissociation limit of the BeH$_2$ model, where the active space localizes on the Be atom such that double excitations on the H$_2$ fragment are simulated by disconnected triple excitation operators. Although the inactive double excitations are present in procedure (ii), the full orthogonalization creates fixed linear combinations of the operators $\hat{a}_{ij}^{ab}$ and $\hat{a}_{uij}^{vab}$ and therefore does not provide a fully extensive description of the H$_2$ fragment, either. The core-extensive sequential orthogonalization, on the other hand, clearly distinguishes between double and triple excitations and guarantees that closed-shell fragments are treated with the same inactive double excitations as when using the ic-MRCCSD approximation. The GNO-based scheme, being fully size-extensive, is also expected to yield an accurate ic-MRCCSDT energy curve.[18] As $|\Psi_0\rangle$ is the Fermi vacuum, the normal ordered excitations $\tilde{a}_{ij}^{ab}$ and $\tilde{a}_{uij}^{vab}$ are even orthogonal to each other from the start. While it would make no difference whether one used a full or a sequential orthogonalization for these operators, this is not true for pairs of operators that both have active indices (cf. Tab. 4).

(a) H$_2$O model system.

(b) N$_2$ model system.
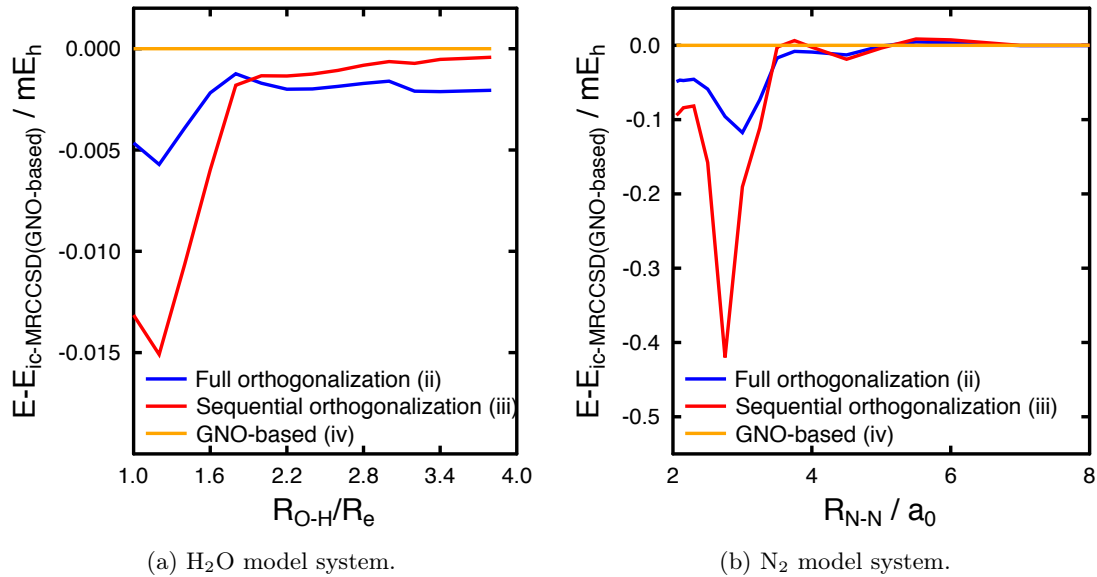
Figure 25: Comparison of ic-MRCCSD methods using the core determinant as Fermi vacuum to the GNO-based scheme (used as reference).

---

[18]At the time of writing this thesis, the GNO-based scheme is not yet available for triple excitations, as I have only implemented the additional transformations (3.64) and (3.65) in a rudimentary form [P4].

It is an interesting question whether a sequential orthogonalization could offer any advantage over a full orthogonalization within the GNO-based approach. Answering this question is beyond the scope of this work. Nevertheless, a comparison of schemes (ii) and (iii) with (iv) for the H_2O and N_2 models, shown in Fig. 25, indicates that the energy difference between the two procedures using a full orthogonalization is often smaller than the one between (iii) and (iv), despite the fact that (iii) and (iv) have core-extensivity in common. Further, procedure (iv) was found to give a potential energy curve for the ground state of LiF less accurate than method (iii) by up to 2 mE_h (see Tab. A4 in the appendix). This suggests that it might be desirable to employ a sequential orthogonalization technique even within the already size-extensive GNO-based framework.

4.5 Treatment of purely active excitations

In most ic-MRCC computations reported in this work, the reference function $|\Psi_0\rangle$ is dynamically relaxed in the presence of the cluster operator. As the metric matrix depends on the reference coefficients c_μ, it must then be recomputed and diagonalized several times during the iterative procedure, preferably in each iteration, in order to achieve a simultaneous fulfillment of Eqs. (3.5) and (3.8). This is computationally feasible in ic-MRCCSD (see sec. 3.2), but it is worth to consider the following two alternative approaches:

The first is based on the assumption that a CASSCF reference already constitutes a good approximation to the projection of the exact wave function onto the model space and simply dispenses with the relaxation of the reference function. The second approach also fixes the reference coefficients at their CASSCF values, but includes additional purely active excitations in the cluster operator in order to allow a corrected description within the model space. For the purpose of the present section, these excitations will be called *internal excitations*. While the latter scheme has less independent parameters than the

(a) BeH_2 model system.

(b) N_2 model system.
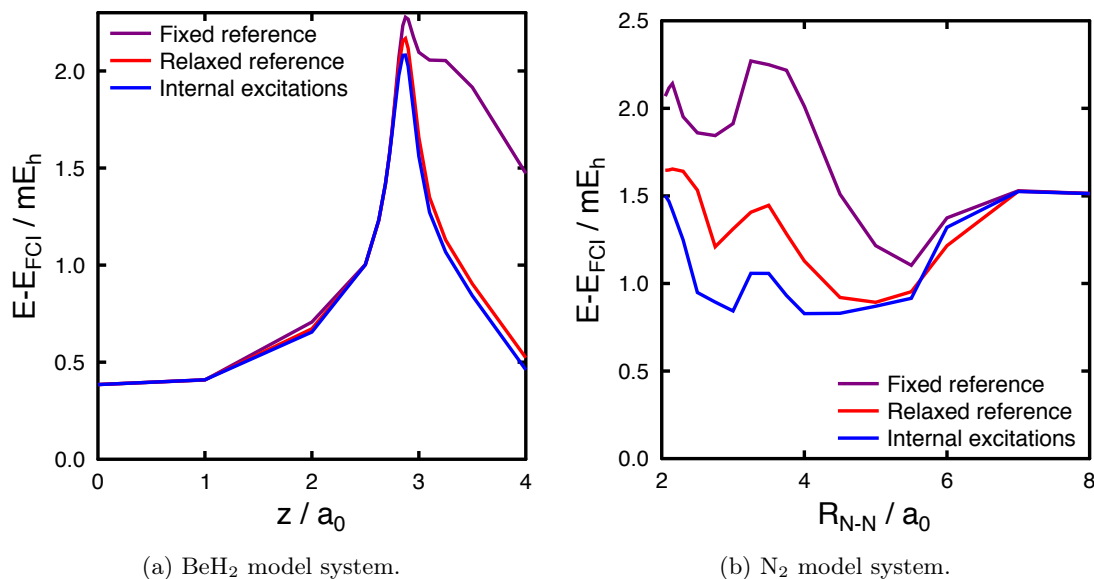
Figure 26: Comparison of alternatives to the relaxation of the reference function in ic-MRCCSD.

one based on a relaxed reference whenever $N_{\text{rank}} < N_{\text{act}}/2$, it also has the property to converge towards FCI when increasing N_{rank}. Note that internal excitations such as $\hat{a}_u^v$ and $\hat{a}_{uv}^{wx}$ contain an effective scalar component, which is redundant and has to be projected out in the course of the sequential orthogonalization procedure.

Figure 26 illuminates the performance of these alternatives for the BeH_2 and N_2 model systems. The use of a fixed reference without internal excitations leads to energies higher by up to about 1 mE_h and is therefore not recommended. This approximation is also not suited in cases such as the LiF model (sec. 4.1.3), where the CASSCF wave functions have qualitatively wrong character in large parts of the potential energy surface.

On the other hand, an inclusion of internal excitations even tends to give slightly lower energies than a relaxation of the c_μ. However, this minor improvement comes at an increased computational cost for evaluating a much larger number of terms in the ic-MRCCSD equations. In the calculations on N_2, the number of individual terms in the amplitude equations truncated at $N_{\text{com}} = 2$ rises from $9\,483$ (Tab. 1) to $17\,631$ upon inclusion of internal singles and doubles. Whether such a scheme might become a computationally attractive alternative to the one that relaxes the reference function, depends on the extent to which an efficient factorization of the ic-MRCC equations is possible.

5 Perturbative triples in ic-MRCC theory

5.1 Motivation

The benchmark applications from chap. 4 demonstrate the power of ic-MRCCSD to accurately describe systems with strong static correlation. Nevertheless, higher than double excitations may be needed for reliably achieving chemical accuracy ($\sim$1 kcal/mol). The applications of ic-MRCCSDT to BeH_2 (sec. 4.1.1) and HF (sec. 4.1.2) exemplify that inclusion of triples can lead to much increased accuracy. Unfortunately, full inclusion of triples entails a much steeper scaling of the computational expense with system size, namely N^8 instead of N^6 (assuming that $M_{\text{virt}} \gg M_{\text{act}}$), drastically restricting the applicability of ic-MRCCSDT. This calls for an approximate treatment of triple excitations that is able to improve the ic-MRCCSD model at a moderate extra cost.

There is a long history of perturbative triples models in SRCC theory. A common strategy is to start from the CCSDT equations (or its Lagrange function, cf. sec. 5.2) and neglect terms appearing at high orders in perturbation theory, including all N^8 scaling terms. This leads to the iterative models CCSDT-n with $n = \{1a, 1b, 2, 3\}$ [172, 173] and to CC3 [174]. These methods require to solve the equations for $\hat{T}_1$ and $\hat{T}_2$ in the presence of $\hat{T}_3$ and therefore involve N^7 steps in each iteration. One can also follow a noniterative approach and use the converged singles and doubles amplitudes from a CCSD calculation to compute approximate triples amplitudes and the resulting energy correction in a single N^7 step. This strategy was pioneered in the CCSD[T] model [173], where the triples amplitudes are derived from fourth-order terms containing $\hat{T}_2$. Additional inclusion of contributions from $\hat{T}_1$ by a selected fifth-order term has given rise to CCSD(T) [12, 13, 175–177], a widely used method often called the 'gold standard' of quantum chemistry [29]. There are alternative approaches aimed at curing deficiencies of CCSD(T) in bond breaking situations [79, 178–182] (see Ref. [168] for a review).

Since MRCC theories are designed to cope with static correlation, it is worthwhile to pursue their multireference analogs of CCSD(T). This has been done for both the tailored [183] and reduced [184] MRCC approaches as well as for spin-flip EOM coupled-cluster theory [185] and a number of methods based on the Jeziorski–Monkhorst ansatz [167, 186–190]. The following sections are dedicated to the development of perturbative triples models in ic-MRCC theory [P3].

5.2 The ic-MRCC Lagrangian

The formulation of SRCC theory in terms of a Lagrange function has proven to be a convenient starting point for the derivation of perturbative triples models [174]. It is straightforward to construct a Lagrangian for the ic-MRCC method by adding to the energy expression (3.3) the ic-MRCC equations (3.8) and (3.5), each equipped with a Lagrange multiplier λ'^ρ or $\bar{c}_\mu$, and the normalization constraint of the reference function (with the Lagrange multiplier α):

$$\mathcal{L} = \langle\Psi_0|\bar{H}|\Psi_0\rangle + \sum_\rho \lambda'^\rho \langle\Psi_0|(\hat{\tau}'^\rho)^\dagger \bar{H}|\Psi_0\rangle + \sum_\mu \bar{c}_\mu \langle\Phi_\mu|(\bar{H} - E)|\Psi_0\rangle - \alpha\left(\langle\Psi_0|\Psi_0\rangle - 1\right).$$

(5.1)

The solution of the ic-MRCC equations is equivalent to requiring stationarity of $\mathcal{L}$ with respect to a variation of the wave function parameters t'_ρ, c_μ and the Lagrange multipli-

ers. Since the normalization of $|\Psi_0\rangle$ has no impact on the fulfillment of the ic-MRCC equations, we can simplify this Lagrangian by setting the multiplier α to its value at the stationary point of $\mathcal{L}$. This is done by multiplying the condition $0 \overset{!}{=} \partial\mathcal{L}/\partial c_\mu$ with c_μ and summing over μ. Because the second and third term of Eq. (5.1) vanish at the stationary point, it follows that $\alpha = \langle\Psi_0|\bar{H}|\Psi_0\rangle/\langle\Psi_0|\Psi_0\rangle$, leading to

$$\mathcal{L} = \frac{\langle\Psi_0|\bar{H}|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle} + \langle\Psi_0|\hat{\Lambda}\bar{H}|\Psi_0\rangle + \langle\overline{\Psi_0}|(\bar{H}-E)|\Psi_0\rangle. \tag{5.2}$$

In Eq. (5.2), we have merged the Lagrange multipliers into the operator

$$\hat{\Lambda} = \sum_\rho \lambda'^\rho(\hat{\tau}'^\rho)^\dagger = \boldsymbol{\lambda}'\boldsymbol{\tau}'^\dagger \tag{5.3}$$

and the left-hand reference function

$$\langle\overline{\Psi_0}| = \sum_\mu \bar{c}_\mu\langle\Phi_\mu|. \tag{5.4}$$

Contrary to Eq. (5.1), the energy expression in Eq. (5.2) is explicitly normalized. A further simplification results from multiplying the condition $0 \overset{!}{=} \partial\mathcal{L}/\partial\bar{c}_\mu$ with c_μ followed by a summation over μ. This directly gives $\langle\Psi_0|\bar{H}|\Psi_0\rangle = E\langle\Psi_0|\Psi_0\rangle$ and allows us to rewrite the Lagrangian as

$$\mathcal{L} = \langle\overline{\Psi_0}|\bar{H}|\Psi_0\rangle + \langle\Psi_0|\hat{\Lambda}\bar{H}|\Psi_0\rangle - E\left(\langle\overline{\Psi_0}|\Psi_0\rangle - 1\right). \tag{5.5}$$

Equation (5.5) is a valid Lagrangian for ic-MRCC theory [P1], since the conditions $\boldsymbol{0} \overset{!}{=} \partial\mathcal{L}/\partial\boldsymbol{\lambda}'$ and $0 \overset{!}{=} \partial\mathcal{L}/\partial\bar{c}_\mu$ are identical to the ic-MRCC equations and $\mathcal{L}$ becomes the ic-MRCC energy at its stationary point. This Lagrangian is more compact than Eqs. (5.1) and (5.2) and therefore used throughout the current work.

5.3   Perturbative analysis

5.3.1   Dyall's zeroth-order Hamiltonian

Single-reference approaches such as Møller–Plesset perturbation theory [191] and perturbative approximations to SRCC methods use the Fock operator, $\hat{F} = \sum_{pq} f_q^p\hat{a}_p^q$, as zeroth-order Hamiltonian. The operator $\hat{F}_{\text{eff}}$ defined in Eq. (4.1) is the multireference analog of the Fock operator in a sense that it is the effective one-particle operator of the Hamiltonian normal ordered with respect to a CASSCF reference function [112]. Unfortunately, the CASSCF function is not an eigenfunction of $\hat{F}_{\text{eff}}$. In the widespread complete active space perturbation theory (CASPT, Ref. [49]), this is remedied by introducing a projection operator onto the reference space. Formally more appealing is the inclusion of the two-electron interactions between active electrons, which leads to Dyall's choice of a zeroth-order Hamiltonian, referred to as CAS/A [192]. In a generalization that does not presume use of pseudocanonical CASSCF orbitals, the Dyall Hamiltonian reads

$$\hat{H}_{\text{D}} = \underbrace{\langle 0|\hat{H}|0\rangle + \sum_{ij}(f_{\text{eff}})_j^i\hat{a}_i^j + \sum_{ab}(f_{\text{eff}})_b^a\hat{a}_a^b}_{\hat{H}_{\text{D,inact}}} + \underbrace{\sum_{uv}f_v^u\hat{a}_u^v + \frac{1}{4}\sum_{uvwx}g_{wx}^{uv}\hat{a}_{uv}^{wx}}_{\hat{H}_{\text{D,act}}}, \tag{5.6}$$

with the elements of the Fock matrix $\mathbf{f}$ and the effective Fock matrix $\mathbf{f}_{\text{eff}}$ as defined in Eqs. (2.23) and (4.1), respectively. $\hat{H}_{\text{D}}$ is employed in n-electron valence state perturbation theory (NEVPT, Refs. [193,194]) and provides a subset of eigenfunctions that span the complete active space,

$$\hat{H}_{\text{D}}|\Psi_m\rangle = E_m|\Psi_m\rangle, \tag{5.7}$$

with $E_0 = E_{\text{CASSCF}}$.

While the Dyall Hamiltonian contains the active two-electron integrals, it is still a sum of purely inactive and purely active contributions. This feature is important for ic-MRCC theory, as replacement of $\hat{H}$ by Dyall's Hamiltonian gives rise to a computationally feasible approximate Jacobian (cf. Refs. [193,195]),

$$A_\rho^\sigma = \left.\frac{\partial^2 \mathcal{L}(\hat{H} \rightarrow \hat{H}_{\text{D}})}{\partial \lambda'^\rho \partial t'_\sigma}\right|_{\mathbf{t}'=\mathbf{0},\boldsymbol{\lambda}'=\mathbf{0}} = \langle\Psi_0|(\hat{\tau}'^\rho)^\dagger[\hat{H}_{\text{D}},\hat{\tau}'^\sigma]|\Psi_0\rangle, \tag{5.8}$$

with a simple block structure similar to the metric matrix from Eq. (3.21),

$$\mathbf{A} = \bigoplus_{n_h,n_p} \left(\mathbf{A}_{\text{inact}}^{n_h,n_p} \otimes \mathbf{1}_{\text{act},\eta}^{n_h,n_p} + \mathbf{1}_{\text{inact}}^{n_h,n_p} \otimes \mathbf{A}_{\text{act}}^{n_h,n_p}\right), \tag{5.9}$$

where

$$\left(\mathbf{A}_{\text{inact}}^{n_h,n_p}\right)_{ab\cdots,kl\cdots}^{ij\cdots,cd\cdots} = -\left(\mathbf{1}_{\text{inact}}^{n_h-1,n_p}\otimes_{\text{A}}\mathbf{f}_{\text{eff}}^{1,0}\right)_{kl\cdots ab\cdots}^{ij\cdots cd\cdots} + \left(\mathbf{1}_{\text{inact}}^{n_h,n_p-1}\otimes_{\text{A}}\mathbf{f}_{\text{eff}}^{0,1}\right)_{kl\cdots ab\cdots}^{ij\cdots cd\cdots} \tag{5.10}$$

and

$$\left(\mathbf{A}_{\text{act}}^{n_h,n_p}\right)_{v\cdots,w\cdots}^{u\cdots,x\cdots} = \sum_{\substack{y\cdots z\cdots \\ u'\cdots v'\cdots}} \left(\left(\mathbf{X}_{\text{act}}^{n_h,n_p}\right)^\dagger\right)_{v\cdots,y\cdots}^{u\cdots,z\cdots} \langle\Psi_0|\hat{a}_{z\cdots}^{y\cdots}[\hat{H}_{\text{D,act}},\hat{a}_{u'\cdots}^{v'\cdots}]|\Psi_0\rangle \left(\mathbf{X}_{\text{act}}^{n_h,n_p}\right)_{v'\cdots,w\cdots}^{u'\cdots,x\cdots}.$$
$$\tag{5.11}$$

In Eq. (5.10), $\mathbf{A}_{\text{inact}}^{n_h,n_p}$ contains the hole-hole and particle-particle blocks of the effective Fock operator and hence becomes diagonal in case of pseudocanonical CASSCF orbitals. The diagonal elements of $\mathbf{A}$ are of practical relevance, as they constitute an efficient choice for the denominators in the amplitude update step (see Ref. [P3] and app. A.1.4).[19]

As will become apparent in sec. 5.3.2, the use of $\hat{H}_{\text{D}}$ as zeroth-order Hamiltonian leads to an undesired coupling between excitations that receive a direct contribution from the Hamiltonian through the first-order residual and other excitations. In order to allow a clear-cut assignment of perturbation orders to excitation operators, it is advantageous to simplify $\hat{H}_{\text{D}}$ as follows. Application of the resolution of the identity

$$\hat{I} = \sum_m |\Psi_m\rangle\langle\Psi_m| + \sum_\rho \hat{\tau}'^\rho|\Psi_0\rangle\langle\Psi_0|(\hat{\tau}'^\rho)^\dagger \tag{5.12}$$

to both sides of $\hat{H}_{\text{D}}$ yields a representation in the basis of model space eigenfunctions and excited functions,

$$\hat{H}_{\text{D}} = \sum_m |\Psi_m\rangle E_m\langle\Psi_m| + \sum_\rho \hat{\tau}'^\rho|\Psi_0\rangle(E_0 + A_\rho^\rho)\langle\Psi_0|(\hat{\tau}'^\rho)^\dagger + \sum_{\rho\neq\sigma} \hat{\tau}'^\rho|\Psi_0\rangle A_\rho^\sigma\langle\Psi_0|(\hat{\tau}'^\sigma)^\dagger.$$
$$\tag{5.13}$$

---

[19]This observation was in fact the main motivation for choosing $\hat{H}_{\text{D}}$ as starting point for defining perturbative triples models. Attempts to neglect the effective active two-electron interaction in the amplitude update step resulted in much worse rates of convergence.

The following perturbative analysis assumes that the off-diagonal elements A_ρ^σ are much less important than the diagonal elements A_ρ^ρ and neglects these terms in the zeroth-order Hamiltonian:

$$\hat{H}^{(0)} = \sum_m |\Psi_m\rangle E_m \langle\Psi_m| + \sum_\rho \hat{\tau}'^\rho |\Psi_0\rangle (E_0 + A_\rho^\rho) \langle\Psi_0|(\hat{\tau}'^\rho)^\dagger. \tag{5.14}$$

For numerical evidence in support of this assumption, see sec. 5.6.1. While simplifying the perturbative analysis, this form of $\hat{H}^{(0)}$ is unsuited for a definition of practical ic-MRCC models with perturbative triples, because Eq. (5.13) relies on Eq. (5.7) and is therefore only valid as long as the reference coefficients c_μ are fixed at their CASSCF values. The actual methods presented in sec. 5.4.2 will thus employ the full Dyall Hamiltonian.

It should be noted that, in principle, one can introduce an additional transformation of the excitation operators $\hat{\tau}'^\rho$ that diagonalizes the active part of $\mathbf{A}$ and thereby eliminates the third term of Eq. (5.13) analytically (cf. Ref. [193]). However, in the present context this would lead to a mixing of excitations of different rank and is not pursued further.

5.3.2 Order-by-order analysis

The Lagrangian from Eq. (5.5) has four sets of stationarity conditions:

$$0 \stackrel{!}{=} \frac{\partial\mathcal{L}}{\partial\bar{c}_\mu} = \langle\Phi_\mu|(\bar{H} - E)|\Psi_0\rangle, \tag{5.15}$$

$$0 \stackrel{!}{=} \frac{\partial\mathcal{L}}{\partial c_\mu} = \langle\overline{\Psi_0}|(\bar{H} - E)|\Phi_\mu\rangle + \langle\Phi_\mu|\hat{\Lambda}\bar{H}|\Psi_0\rangle + \langle\Psi_0|\hat{\Lambda}\bar{H}|\Phi_\mu\rangle, \tag{5.16}$$

$$0 \stackrel{!}{=} \frac{\partial\mathcal{L}}{\partial\lambda'^\rho} = \Omega'_\rho = \langle\Psi_0|(\hat{\tau}'^\rho)^\dagger\bar{H}|\Psi_0\rangle, \tag{5.17}$$

$$0 \stackrel{!}{=} \frac{\partial\mathcal{L}}{\partial t'_\rho} = \langle\overline{\Psi_0}|\frac{\partial\bar{H}}{\partial t'_\rho}|\Psi_0\rangle + \langle\Psi_0|\hat{\Lambda}\frac{\partial\bar{H}}{\partial t'_\rho}|\Psi_0\rangle. \tag{5.18}$$

Condition (5.15) is the eigensystem for the reference coefficients [Eq. (3.5)], and (5.17) are the amplitude equations [Eq. (3.8)]. The conditions (5.16) and (5.18) determine the Lagrange multipliers $\bar{c}_\mu$ and λ'^ρ, respectively, and are not needed in a regular computation of the ic-MRCC energy.

Let us now partition the Hamiltonian into $\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}$ and expand all parameters in orders of perturbation, e.g., $t'_\rho = t'^{(0)}_\rho + t'^{(1)}_\rho + t'^{(2)}_\rho + \dots$. A shorthand notation will be used to indicate perturbed reference functions, e.g., $|\Psi_0^{(n)}\rangle = \sum_\mu |\Phi_\mu\rangle c_\mu^{(n)}$.

Our starting point is to recognize that $\hat{T}^{(0)} = 0$ and $\hat{\Lambda}^{(0)} = 0$ is a possible solution to Eqs. (5.15) to (5.18) in zeroth order:

$$0 = \langle\Phi_\mu|(\hat{H}^{(0)} - E^{(0)})|\Psi_0^{(0)}\rangle = \langle\Phi_\mu|(E_0 - E^{(0)})|\Psi_0^{(0)}\rangle, \tag{5.19}$$

$$0 = \langle\overline{\Psi_0}^{(0)}|(\hat{H}^{(0)} - E^{(0)})|\Phi_\mu\rangle = \langle\Psi_0^{(0)}|(E_0 - E^{(0)})|\Phi_\mu\rangle, \tag{5.20}$$

$$0 = \langle\Psi_0^{(0)}|(\hat{\tau}'^\rho)^\dagger\hat{H}^{(0)}|\Psi_0^{(0)}\rangle, \tag{5.21}$$

$$0 = \langle\overline{\Psi_0}^{(0)}|[\hat{H}^{(0)}, \hat{\tau}'^\rho]|\Psi_0^{(0)}\rangle. \tag{5.22}$$

Equations (5.19) and (5.20) are in accordance with taking the CASSCF coefficients for both $c_\mu^{(0)}$ and $\bar{c}_\mu^{(0)}$ and yield the CASSCF energy as zeroth-order energy, $E^{(0)} = E_0$. The terms in Eqs. (5.21) and (5.22) vanish, because $\hat{\tau}'^\rho$ has at least one inactive index, which cannot be compensated by $\hat{H}^{(0)}$ or the reference function.

In first order, the stationarity conditions take the following form:

$$0 = \langle \Phi_\mu | \left(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}^{(1)}] \right) | \Psi_0^{(0)} \rangle - E^{(1)} \langle \Phi_\mu | \Psi_0^{(0)} \rangle + \langle \Phi_\mu | (\hat{H}^{(0)} - E^{(0)}) | \Psi_0^{(1)} \rangle, \quad (5.23)$$

$$0 = \langle \Psi_0^{(0)} | \left(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}^{(1)}] \right) | \Phi_\mu \rangle - E^{(1)} \langle \Psi_0^{(0)} | \Phi_\mu \rangle + \langle \overline{\Psi_0}^{(1)} | (\hat{H}^{(0)} - E^{(0)}) | \Phi_\mu \rangle$$
$$+ \langle \Phi_\mu | \hat{\Lambda}^{(1)} \hat{H}^{(0)} | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | \hat{\Lambda}^{(1)} \hat{H}^{(0)} | \Phi_\mu \rangle, \quad (5.24)$$

$$0 = \langle \Psi_0^{(1)} | (\hat{\tau}'^\rho)^\dagger \hat{H}^{(0)} | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | (\hat{\tau}'^\rho)^\dagger \hat{H}^{(0)} | \Psi_0^{(1)} \rangle + \langle \Psi_0^{(0)} | (\hat{\tau}'^\rho)^\dagger \hat{H}^{(1)} | \Psi_0^{(0)} \rangle$$
$$+ \langle \Psi_0^{(0)} | (\hat{\tau}'^\rho)^\dagger [\hat{H}^{(0)}, \hat{T}^{(1)}] | \Psi_0^{(0)} \rangle, \quad (5.25)$$

$$0 = \langle \overline{\Psi_0}^{(1)} | [\hat{H}^{(0)}, \hat{\tau}'^\rho] | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | [\hat{H}^{(0)}, \hat{\tau}'^\rho] | \Psi_0^{(1)} \rangle + \langle \Psi_0^{(0)} | [\hat{H}^{(1)}, \hat{\tau}'^\rho] | \Psi_0^{(0)} \rangle$$
$$+ \langle \Psi_0^{(0)} | \left([[\hat{H}^{(0)}, \hat{T}^{(1)}], \hat{\tau}'^\rho] + [[\hat{H}^{(0)}, \hat{\tau}'^\rho], \hat{T}^{(1)}] \right) | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | \hat{\Lambda}^{(1)} [\hat{H}^{(0)}, \hat{\tau}'^\rho] | \Psi_0^{(0)} \rangle. \quad (5.26)$$

Again, a number of terms vanish (printed in blue), in which operators with inactive indices (including $\hat{H}^{(1)}$) annihilate a left-hand or right-hand reference function. The condition (5.23) is a set of equations projected onto the reference determinants $\langle \Phi_\mu |$. We can split it into two distinct conditions by separately considering projections onto $\langle \Psi_0^{(0)} |$ and its complementary space spanned by $\{ \langle \Phi_\mu | \hat{Q} \}$, where we have introduced the projector $\hat{Q} = 1 - | \Psi_0^{(0)} \rangle \langle \Psi_0^{(0)} |$. Using Eq. (5.20), the projection onto $\langle \Psi_0^{(0)} |$ directly leads to $E^{(1)} = 0$, whereas the projection onto the complementary model space yields a homogeneous system of linear equations,

$$\sum_\nu \langle \Phi_\mu | \hat{Q} (\hat{H}^{(0)} - E^{(0)}) | \Phi_\nu \rangle c_\nu^{(1)} = 0, \quad (5.27)$$

with the trivial solution $c_\mu^{(1)} = 0$. In analogy to this, Eq. (5.24) results in $\bar{c}_\mu^{(1)} = 0$. Hence, there is no first-order contribution to both the right-hand and left-hand reference functions.

The third term of Eq. (5.25) is the initial contribution to the ic-MRCC residual. Due to the generalized Brillouin theorem, Eq. (2.34), and the current choice of $\hat{H}^{(0)}$, it is zero for singly excited functions. In addition, the sequential orthogonalization procedure (sec. 3.2.1) rules out projections onto triply and higher excited functions, since $\hat{H}^{(1)}$ is a two-body operator. Consequently, this term vanishes for all but doubly excited functions,

$$\langle \Psi_0^{(0)} | (\hat{\tau}_n'^\rho)^\dagger \hat{H}^{(1)} | \Psi_0^{(0)} \rangle = 0 \qquad \text{if} \quad n = \{1, 3, 4, \ldots\}. \quad (5.28)$$

Owing to the exclusion of the off-diagonal elements A_ρ^σ from $\hat{H}^{(0)}$, the last term in Eq. (5.25) simplifies to

$$\langle \Psi_0^{(0)} | (\hat{\tau}'^\rho)^\dagger [\hat{H}^{(0)}, \hat{T}^{(1)}] | \Psi_0^{(0)} \rangle = A_\rho^\rho t_\rho'^{(1)}. \quad (5.29)$$

As a result, the first-order contribution to the doubles amplitudes takes the simple form

$$t_{2,\rho}'^{(1)} = -\frac{1}{A_\rho^\rho} \langle \Psi_0^{(0)} | (\hat{\tau}_2'^\rho)^\dagger \hat{H}^{(1)} | \Psi_0^{(0)} \rangle. \quad (5.30)$$

This is analogous to SRCC theory, where the Brillouin theorem ensures that only double excitations arise in first order of perturbation theory. Importantly, the first-order equations for the Lagrange multipliers λ'^ρ, Eq. (5.26), turn out to be the Hermitian conjugate of Eq. (5.25), such that $\boldsymbol{\lambda}_2'^{(1)} = (\mathbf{t}_2'^{(1)})^\dagger$.

In second order, Eqs. (5.15) to (5.18) become

$$0 = \langle\Phi_\mu|[\hat{H}^{(1)}, \hat{T}_2^{(1)}]|\Psi_0^{(0)}\rangle - E^{(2)}\langle\Phi_\mu|\Psi_0^{(0)}\rangle + \langle\Phi_\mu|(\hat{H}^{(0)} - E^{(0)})|\Psi_0^{(2)}\rangle, \tag{5.31}$$

$$0 = \langle\Psi_0^{(0)}|[\hat{H}^{(1)}, \hat{T}_2^{(1)}]|\Phi_\mu\rangle - E^{(2)}\langle\Psi_0^{(0)}|\Phi_\mu\rangle + \langle\overline{\Psi}_0^{(2)}|(\hat{H}^{(0)} - E^{(0)})|\Phi_\mu\rangle$$
$$+ \langle\Phi_\mu|(\hat{T}_2^{(1)})^\dagger\big(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}_2^{(1)}]\big)|\Psi_0^{(0)}\rangle + \langle\Psi_0^{(0)}|(\hat{T}_2^{(1)})^\dagger\big(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}_2^{(1)}]\big)|\Phi_\mu\rangle, \tag{5.32}$$

$$0 = \langle\Psi_0^{(0)}|(\hat{\tau}'^\rho)^\dagger[\hat{H}^{(1)}, \hat{T}_2^{(1)}]|\Psi_0^{(0)}\rangle + \tfrac{1}{2}\langle\Psi_0^{(0)}|(\hat{\tau}'^\rho)^\dagger[[\hat{H}^{(0)}, \hat{T}_2^{(1)}], \hat{T}_2^{(1)}]|\Psi_0^{(0)}\rangle + A_\rho^\rho t_\rho'^{(2)}, \tag{5.33}$$

$$0 = \langle\Psi_0^{(0)}|\big([[\hat{H}^{(1)}, \hat{T}_2^{(1)}], \hat{\tau}'^\rho] + [[\hat{H}^{(1)}, \hat{\tau}'^\rho], \hat{T}_2^{(1)}]\big)|\Psi_0^{(0)}\rangle + \lambda'^{\rho(2)}A_\rho^\rho$$
$$+ \langle\Psi_0^{(0)}|(\hat{T}_2^{(1)})^\dagger\big([\hat{H}^{(1)}, \hat{\tau}'^\rho] + [[\hat{H}^{(0)}, \hat{T}_2^{(1)}], \hat{\tau}'^\rho] + [[\hat{H}^{(0)}, \hat{\tau}'^\rho], \hat{T}_2^{(1)}]\big)|\Psi_0^{(0)}\rangle, \tag{5.34}$$

where vanishing terms have been omitted from the start. We can obtain an expression for the second-order energy by considering the projections of Eqs. (5.31) and (5.32) onto the zeroth-order reference functions. In Eq. (5.32), this is formally achieved by multiplying with $c_\mu^{(0)}$ and summing over μ. Inserting Eqs. (5.29) and (5.30) into the result, the expressions derived from the last two terms of Eq. (5.32) turn out to cancel according to

$$\langle\Psi_0^{(0)}|(\hat{T}_2^{(1)})^\dagger\big(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}_2^{(1)}]\big)|\Psi_0^{(0)}\rangle = -\sum_\rho A_\rho^\rho(t_{2,\rho}'^{(1)})^2 + \sum_\rho A_\rho^\rho(t_{2,\rho}'^{(1)})^2 = 0. \tag{5.35}$$

Consequently, both Eqs. (5.31) and (5.32) yield the same second-order energy contribution:

$$E^{(2)} = \langle\Psi_0^{(0)}|[\hat{H}^{(1)}, \hat{T}_2^{(1)}]|\Psi_0^{(0)}\rangle = -\sum_\rho \frac{1}{A_\rho^\rho}\big|\langle\Psi_0^{(0)}|(\hat{\tau}_2'^\rho)^\dagger\hat{H}^{(1)}|\Psi_0^{(0)}\rangle\big|^2. \tag{5.36}$$

In the single-reference limit ($N_{\text{act}} = 0$), $E^{(2)}$ reduces to the second-order energy correction of Møller–Plesset perturbation theory (MP2). Studying the performance of this variant of multireference perturbation theory is beyond the scope of the present work.

The projections of Eqs. (5.31) and (5.32) onto the space $\{\hat{Q}|\Phi_\mu\rangle\}$ lead to two different sets of linear equations determining $c_\mu^{(2)}$ and $\bar{c}_\mu^{(2)}$:[20]

$$\sum_\nu \langle\Phi_\mu|\hat{Q}(\hat{H}^{(0)} - E^{(0)})|\Phi_\nu\rangle c_\nu^{(2)} = -\langle\Phi_\mu|\hat{Q}[\hat{H}^{(1)}, \hat{T}_2^{(1)}]|\Psi_0^{(0)}\rangle, \tag{5.37}$$

$$\sum_\nu \bar{c}_\nu^{(2)}\langle\Phi_\nu|(\hat{H}^{(0)} - E^{(0)})\hat{Q}|\Phi_\mu\rangle = -\langle\Psi_0^{(0)}|[\hat{H}^{(1)}, \hat{T}_2^{(1)}]\hat{Q}|\Phi_\mu\rangle$$
$$- \langle\Phi_\mu|\hat{Q}(\hat{T}_2^{(1)})^\dagger\big(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}_2^{(1)}]\big)|\Psi_0^{(0)}\rangle$$
$$- \langle\Psi_0^{(0)}|(\hat{T}_2^{(1)})^\dagger\big(\hat{H}^{(1)} + [\hat{H}^{(0)}, \hat{T}_2^{(1)}]\big)\hat{Q}|\Phi_\mu\rangle. \tag{5.38}$$

---

[20]A rigorous way to achieve this projection for Eq. (5.32) is to evaluate $(\partial\mathcal{L}/\partial c_\mu)^{(2)} - c_\mu^{(0)}\sum_\nu(\partial\mathcal{L}/\partial c_\nu)^{(2)}$.

Turning to Eqs. (5.33) and (5.34), a comparison of these equations suggests that the second-order contributions to $\hat{T}$ and $\hat{\Lambda}$ are also generally different. This stands in contrast to SRCC theory, where $\boldsymbol{\lambda}_n'^{(2)} = (\mathbf{t}_n^{(2)})^\dagger$ for $n = \{1, 2, 3\}$ [177]. Additionally, in SRCC theory, second-order contributions to the cluster amplitudes only affect singles, doubles and triples. While the commutator $[\hat{H}^{(1)}, \hat{T}_2^{(1)}]$ in the first term of Eq. (5.33) has a maximum rank of three and can only yield contributions to $\hat{T}_1$, $\hat{T}_2$ and $\hat{T}_3$ due to the sequential orthogonalization approach, the second term in Eq. (5.33) also gives a nonzero residual for quadruple excitations with at least two active indices. In order to reformulate the latter term, we expand the double commutator, insert Eq. (5.14) and exploit that the two excitation operators in the expectation value $\langle\Psi_0^{(0)}|(\hat{\tau}_4'^\rho)^\dagger\hat{\tau}_2'^\sigma\hat{\tau}_2'^v|\Psi_0^{(0)}\rangle$ commute. This yields

$$
\begin{aligned}
\tfrac{1}{2}\langle\Psi_0^{(0)}|(\hat{\tau}_4'^\rho)^\dagger&[[\hat{H}^{(0)}, \hat{T}_2^{(1)}], \hat{T}_2^{(1)}]|\Psi_0^{(0)}\rangle \\
&= \tfrac{1}{2}\sum_{\sigma v} t_{2,\sigma}'^{(1)} t_{2,v}'^{(1)} \left(A_\rho^\rho - A_\sigma^\sigma - A_v^v\right) \langle\Psi_0^{(0)}|(\hat{\tau}_4'^\rho)^\dagger\hat{\tau}_2'^\sigma\hat{\tau}_2'^v|\Psi_0^{(0)}\rangle.
\end{aligned}
\tag{5.39}
$$

According to Eq. (5.9), the diagonal elements of $\mathbf{A}$ are sums of inactive effective Fock matrix elements and active parts. As the inactive indices of $(\hat{\tau}_4'^\rho)^\dagger$ must match the ones of $\hat{\tau}_2'^\sigma$ and $\hat{\tau}_2'^v$ in the expectation value, the elements from $\mathbf{f}_{\text{eff}}$ cancel each other out in $(A_\rho^\rho - A_\sigma^\sigma - A_v^v)$, leaving a purely active contribution. In light of the fact that triple excitations also receive a contribution from the first term in Eq. (5.33), it is fair to assume that the second-order contribution to $\hat{T}_4$ is smaller than the one to $\hat{T}_3$. Therefore, it is justified to restrict the perturbative improvement of the ic-MRCCSD method to the approximate inclusion of triple excitations.

5.4 Approximations to the ic-MRCC Lagrangian

5.4.1 Neglect of triples with many active indices

Figure 27 shows all excitation operators present in the ic-MRCCSDT method. In the excitation classes with $n_h = 3$ or $n_p = 3$, three is the lowest excitation rank. Those triple excitations, drawn in orange in Fig. 27, are analogs of the triple excitations in SRCC theory and all of them should be treated on the same footing.[21]

The situation is somewhat different for the other triple excitations depicted in Fig. 27, which involve active-to-active excitations. A significant portion of the corresponding triply excited functions in these excitation classes is already covered by excitations of lower rank (cf. the discussion in sec. 4.3). Even when the reference function is close to a single determinant, an excited function that would be accessed by a triple excitation in SRCC theory is often reached via a double excitation from a configuration with small weight. Likewise, single excitations partially account for doubles with active-to-active excitations. It is important to realize that this is not an artifact of ic-MRCC theory but entirely due to the multiconfigurational nature of the reference function, which renders even single excitations from minor configurations second-order quantities due to the generalized Brillouin theorem, whereas the corresponding double excitations in SRCC theory arise in first order of perturbation theory. The partial redundancy of triples with

---

[21]Of course, some types of triples with $n_h = 3$ or $n_p = 3$ may be redundant for small active spaces, e.g., $\hat{a}_{ijk}^{uvw}$ and $\hat{a}_{uvw}^{abc}$ in case of a CAS(2,2).
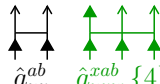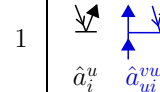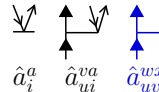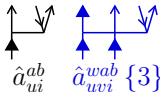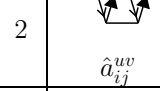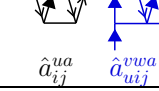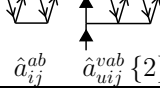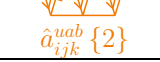
n_h	$n_p = 0$	$n_p = 1$	$n_p = 2$	$n_p = 3$
0	—	$\hat{a}_u^a$ $\hat{a}_{uv}^{wa}$ $\hat{a}_{uvw}^{xya}$ {5}	$\hat{a}_{uv}^{ab}$ $\hat{a}_{uvw}^{xab}$ {4}	$\hat{a}_{uvw}^{abc}$ {2}
1	$\hat{a}_i^u$ $\hat{a}_{ui}^{vw}$ $\hat{a}_{uvi}^{wxy}$ {5}	$\hat{a}_i^a$ $\hat{a}_{ui}^{va}$ $\hat{a}_{uvi}^{wxa}$ {4}	$\hat{a}_{ui}^{ab}$ $\hat{a}_{uvi}^{wab}$ {3}	$\hat{a}_{uvi}^{abc}$ {2}
2	$\hat{a}_{ij}^{uv}$ $\hat{a}_{uij}^{vwx}$ {4}	$\hat{a}_{ij}^{ua}$ $\hat{a}_{uij}^{vwa}$ {3}	$\hat{a}_{ij}^{ab}$ $\hat{a}_{uij}^{vab}$ {2}	$\hat{a}_{uij}^{abc}$ {2}
3	$\hat{a}_{ijk}^{uvw}$ {2}	$\hat{a}_{ijk}^{uva}$ {2}	$\hat{a}_{ijk}^{uab}$ {2}	$\hat{a}_{ijk}^{abc}$ {2}

Figure 27: Excitations in the cluster operator for ic-MRCCSDT. For each triple excitation, the lowest triples model $\{n\}$ is specified that considers this excitation.

active-to-active excitations suggests that the different types of triple excitations are not all equally important. In fact, some of these excitation types are exactly redundant for certain small active spaces (as long as η is set small enough to include all nearly linearly dependent excitations in the lower ranks). The excitations highlighted in green in Fig. 27 are redundant for a CAS(4,4), and in case of a CAS(2,2) one may additionally exclude the excitations shown in blue.

While we can expect triple excitations with many active indices to have little impact on the ic-MRCC energy, they may largely increase the computational effort. In particular, the excitations $\hat{a}_{uvi}^{wxy}$ and $\hat{a}_{uvw}^{xya}$ require the diagonalization of five-body density matrices. As we have seen in Fig. 9, this is likely the rate determining step for large active spaces.

It is therefore recommended to neglect excitations with many active lines in the design of an efficient ic-MRCC method with perturbative triples. To this end, we consider a hierarchy of triples models $\{n\}$, where n is the maximum number of active indices in a triple excitation with $n_h \leq 2$ and $n_p \leq 2$. The choice advocated in this work is model $\{3\}$. Like ic-MRCCSD, this triples model involves only diagonalizations of up to three-body metric blocks. At the same time, it recovers most of the triples contribution, as will be demonstrated in sec. 5.6.1.

5.4.2 Perturbative and commutator-based truncation

The ic-MRCC methods with perturbative treatment of triple excitations studied in this work are based on the simple recipe to assign perturbation orders to the various operators in the ic-MRCC Lagrangian, Eq. (5.5), followed by a truncation of this Lagrangian at a specific order. Similar strategies have been used with success in the development of SRCC methods (e.g., Refs. [174, 196, 197]). The perturbative analysis in sec. 5.3.2 provides some guidance in the assignment of perturbation orders, but it takes the CASSCF wave function as a starting point and is therefore only of limited value when devising perturbative improvements on top of ic-MRCCSD. This also applies to SRCC theory. A notable example for shortcomings of the perturbative analysis based on a Hartree–Fock wave function is its inability to explain the importance of the famous 'fifth-order

term' in CCSD(T), which results from the Lagrangian through the term $\langle 0|\hat{\Lambda}_1[\hat{H}, \hat{T}_3]|0\rangle$. In contrast to the Hartree–Fock based perturbation theory, an analysis focused on the CCSD wave function as zeroth-order solution can justify the inclusion of this term and suggests that $\hat{\Lambda}_1$ should be treated on the same footing as $\hat{\Lambda}_2$ [175]. In addition to this, the relaxed reference function in ic-MRCCSD generally differs from the CASSCF wave function. This implies an increased significance of $\hat{T}_1$, as the generalized Brillouin theorem does not hold for the relaxed reference. Further, the zeroth-order Hamiltonian depends on the reference function, rendering the definition from Eq. (5.14) inconvenient for practical applications.

These considerations motivate the following assignment of perturbation orders: Both the left- and right-hand reference functions and the energy E are treated as zeroth-order quantities. The Hamiltonian is split into the zeroth-order Dyall Hamiltonian $\hat{H}_{\mathrm{D}}$ and a first-order remainder. The singles and doubles amplitudes and Lagrange multipliers in $\hat{T}_1$, $\hat{\Lambda}_1$, $\hat{T}_2$ and $\hat{\Lambda}_2$ are assigned first order, while $\hat{T}_3$ and $\hat{\Lambda}_3$ are second-order operators. With these assignments, the first terms scaling like N^8 (with respect to the inactive orbital space) arise in fifth order of perturbation theory, e.g., $\langle\Psi_0|\hat{\Lambda}_3[\hat{H}, \hat{T}_3]|\Psi_0\rangle$. By truncating the ic-MRCCSDT Lagrangian at fourth order, we exclude these terms and obtain a new method called ic-MRCCSDT-1 [P3].

The method ic-MRCCSDT-1 is related to CCSDT-1a [172], but an application of the described recipe to SRCC theory does not directly lead to CCSDT-1a. The reason for this is that the three- and fourfold commutator terms in CCSD (and CCSDT-1a) have a perturbation order higher than four and would be excluded in the present approach. A modified procedure that applies the perturbative truncation only to terms not present in the singles-and-doubles model is able to cure this deficiency. However, such a modification makes no difference in ic-MRCC theory as long as we adhere to a truncation of the BCH expansion after the twofold commutator. The presented ic-MRCCSDT-1 model is thus a multireference analog of the quadratic approximation to CCSDT-1a. The ic-MRCCSDT-1 Lagrangian reads

$$
\begin{aligned}
\mathcal{L}_{\text{ic-MRCCSDT-1}} = &\langle\overline{\Psi_0}|\big(\hat{H}_{\mathrm{D}} + [\hat{H}, \hat{T}] + \tfrac{1}{2}[[\hat{H}, \hat{T}], \hat{T}_1 + \hat{T}_2] + \tfrac{1}{2}[[\hat{H}, \hat{T}_1 + \hat{T}_2], \hat{T}_3]\big)|\Psi_0\rangle \\
&+ \langle\Psi_0|\big(\hat{\Lambda}_1 + \hat{\Lambda}_2\big)\big(\hat{H} + [\hat{H}, \hat{T}] + \tfrac{1}{2}[[\hat{H}, \hat{T}_1 + \hat{T}_2], \hat{T}_1 + \hat{T}_2] \\
&\qquad\qquad + \tfrac{1}{2}[[\hat{H}_{\mathrm{D}}, \hat{T}_1 + \hat{T}_2], \hat{T}_3] + \tfrac{1}{2}[[\hat{H}_{\mathrm{D}}, \hat{T}_3], \hat{T}_1 + \hat{T}_2]\big)|\Psi_0\rangle \\
&+ \langle\Psi_0|\hat{\Lambda}_3\big(\hat{H} + [\hat{H}, \hat{T}_1 + \hat{T}_2] + [\hat{H}_{\mathrm{D}}, \hat{T}_3] + \tfrac{1}{2}[[\hat{H}_{\mathrm{D}}, \hat{T}_1 + \hat{T}_2], \hat{T}_1 + \hat{T}_2]\big)|\Psi_0\rangle \\
&- E\big(\langle\overline{\Psi_0}|\Psi_0\rangle - 1\big).
\end{aligned}
\tag{5.40}
$$

In order to allow the three- and fourfold commutator terms from SRCC theory in the ic-MRCC methods while simultaneously avoiding a large number of terms with high powers in $\hat{T}_2$ (cf. Tab. 1), one has to apply different truncation criteria for single and double excitations. The most elegant way to achieve this is to switch to a sequential ansatz for the ic-MRCC wave function [P5],

$$
|\Psi\rangle = e^{\hat{T}_1} e^{\hat{T}_2 + \hat{T}_3}|\Psi_0\rangle.
\tag{5.41}
$$

The similarity transformation of the Hamiltonian then falls into a rank-conserving trans-

formation with the singles operator,

$$\tilde{H} = e^{-\hat{T}_1}\hat{H}e^{\hat{T}_1}, \tag{5.42}$$

and a second transformation with the remaining excitation operators,[22]

$$\bar{H} = e^{-\hat{T}_2-\hat{T}_3}\tilde{H}e^{\hat{T}_2+\hat{T}_3}. \tag{5.43}$$

It is straightforward to truncate Eq. (5.42) at fourfold commutators and retain the truncation at twofold commutators for Eq. (5.43), thereby including all terms present in SRCC theory. This sequential ansatz gives rise to an ic-MRCC theory with the same formal properties as the one based on the simultaneous ansatz from Eq. (2.41), and the numerical differences between the two are often insignificant [P5] (see also Ref. [198] for a valuable discussion).

It is possible to combine the sequential ansatz with a perturbative truncation while treating $\hat{T}_1$ as first-order quantity. However, this requires the expansion of $\tilde{H}$ in the Lagrangian, in order to apply the truncation criterion consistently. A conceptually simpler solution is to assign zeroth order to $\hat{T}_1$ and introduce a $\hat{T}_1$-transformed Dyall Hamiltonian, $\tilde{H}_{\mathrm{D}} = e^{-\hat{T}_1}\hat{H}_{\mathrm{D}}\,e^{\hat{T}_1}$. The resulting ic-MRCC method then becomes analogous to the CC3 model [174]. This model emphasizes the role of singles amplitudes as orbital relaxation parameters and is especially suited for the calculation of response properties [199]. In this work, the corresponding ic-MRCC3 approach will also treat $\hat{\Lambda}_1$ as zeroth-order quantity, although this is not strictly necessary for achieving equivalence of ic-MRCC3 with CC3 in the single-reference limit. The ic-MRCC3 Lagrangian takes the following form:

$$\begin{aligned}
\mathcal{L}_{\text{ic-MRCC3}} = {}&\langle\overline{\Psi_0}|\big(\tilde{H} + [\tilde{H},\hat{T}_2+\hat{T}_3] + \tfrac{1}{2}[[\tilde{H},\hat{T}_2+\hat{T}_3],\hat{T}_2] + \tfrac{1}{2}[[\tilde{H},\hat{T}_2],\hat{T}_3]\big)|\Psi_0\rangle \\
&+ \langle\Psi_0|\hat{\Lambda}_1\big(\tilde{H} + [\tilde{H},\hat{T}_2+\hat{T}_3] + \tfrac{1}{2}[[\tilde{H},\hat{T}_2+\hat{T}_3],\hat{T}_2] + \tfrac{1}{2}[[\tilde{H},\hat{T}_2],\hat{T}_3]\big)|\Psi_0\rangle \\
&+ \langle\Psi_0|\hat{\Lambda}_2\big(\tilde{H} + [\tilde{H},\hat{T}_2+\hat{T}_3] + \tfrac{1}{2}[[\tilde{H},\hat{T}_2],\hat{T}_2] \\
&\qquad + \tfrac{1}{2}[[\tilde{H}_{\mathrm{D}},\hat{T}_2],\hat{T}_3] + \tfrac{1}{2}[[\tilde{H}_{\mathrm{D}},\hat{T}_3],\hat{T}_2]\big)|\Psi_0\rangle \\
&+ \langle\Psi_0|\hat{\Lambda}_3\big(\tilde{H} + [\tilde{H},\hat{T}_2] + [\tilde{H}_{\mathrm{D}},\hat{T}_3] + \tfrac{1}{2}[[\tilde{H}_{\mathrm{D}},\hat{T}_2],\hat{T}_2]\big)|\Psi_0\rangle \\
&- E\big(\langle\overline{\Psi_0}|\Psi_0\rangle - 1\big).
\end{aligned} \tag{5.44}$$

An important aspect of both CCSDT-1a and CC3 regarding computational efficiency is that the only term by which $\hat{T}_3$ contributes to the triples residual is $\langle 0|\hat{\Lambda}_3[\hat{F},\hat{T}_3]|0\rangle$. With a diagonal Fock operator, the triples amplitudes can be directly computed in each iteration from $\hat{T}_2$ and the ($\hat{T}_1$-transformed) Hamiltonian in a single step. By inserting the equations for the triples into the residual expressions of the singles and doubles, one can avoid the storage of triples amplitudes altogether [172, 174, 199]. This attractive feature is lost in the current formulation of the ic-MRCCSDT-1 and ic-MRCC3 models. It might be possible to simplify these methods by replacing $[\hat{H}_{\mathrm{D}},\hat{T}_3]$ in Eq. (5.40) and $[\tilde{H}_{\mathrm{D}},\hat{T}_3]$ in Eq. (5.44) by $[\hat{H}^{(0)},\hat{T}_3]$, but this is beyond the scope of the current work.

---

[22]Reference [P5] postulates a fully sequential ansatz with the wave operator $e^{\hat{T}_1}e^{\hat{T}_2}e^{\hat{T}_3}\cdots$. However, similarity transformations of the Hamiltonian with $\hat{T}_2$ or higher excitation operators are not rank-conserving and do not offer a clear computational advantage over using $e^{\hat{T}_1}e^{\hat{T}_2+\hat{T}_3+\cdots}$.

5.5 The noniterative triples correction: ic-MRCCSD(T)

The essence of CCSD(T) is to perform the first iteration of CCSDT-1a using $\hat{T}_1$ and $\hat{T}_2$ from a converged CCSD calculation and then evaluate the CCSDT-1a Lagrangian, with $\hat{\Lambda}_1$ and $\hat{\Lambda}_2$ replaced by $\hat{T}_1^\dagger$ and $\hat{T}_2^\dagger$, respectively. As discussed in sec. 5.3.2, the replacement of $\hat{\Lambda}$ by $\hat{T}^\dagger$ in ic-MRCC is correct to first order in perturbation theory. This also applies to the substitution of $\langle\Psi_0|$ for $\langle\overline{\Psi_0}|$. Starting from the ic-MRCCSDT-1 Lagrangian, Eq. (5.40), these replacements lead to a Lagrangian of the following form:

$$\mathcal{L}_{\text{ic-MRCCSD(T)}} = E_{\text{ic-MRCCSD}} + \Delta E_{(T)}^{\text{direct}} + \mathbf{t}_1^{\prime\dagger}\mathbf{\Omega}_{1,T}' + \mathbf{t}_2^{\prime\dagger}\mathbf{\Omega}_{2,T}' + \boldsymbol{\lambda}_3'\mathbf{\Omega}_3'. \tag{5.45}$$

In Eq. (5.45), no perturbative truncation is applied to the ic-MRCCSD energy $E_{\text{ic-MRCCSD}}$. This requirement plays a role when using the sequential ansatz from Eq. (5.41). $\Delta E_{(T)}^{\text{direct}}$ is a direct contribution of $\hat{T}_3$ to the energy and has no counterpart in single-reference CCSD(T). $\mathbf{t}_1'$ and $\mathbf{t}_2'$ are the converged singles and doubles amplitudes from ic-MRCCSD and Eq. (5.45) implicitly uses the relaxed reference function from ic-MRCCSD. Since the corresponding ic-MRCCSD residuals are zero, $\mathbf{\Omega}_{1,T}'$ and $\mathbf{\Omega}_{2,T}'$ are restricted to terms involving the operator $\hat{T}_3$.

The triples amplitudes follow from the condition $\mathbf{\Omega}_3' = \mathbf{0}$. With ic-MRCCSDT-1 as starting point, the triples residual contains one $\hat{T}_3$ dependent term,

$$\langle\Psi_0|(\hat{\tau}_3^{\prime\rho})^\dagger[\hat{H}_D, \hat{T}_3]|\Psi_0\rangle = \sum_\sigma A_\rho^\sigma t_{3,\sigma}', \tag{5.46}$$

where we have used Eq. (5.8). This term introduces a coupling between different triples amplitudes and requires an iterative solution of the equations for $\hat{T}_3$. The corresponding triples model will be denoted as ic-MRCCSD(T)$_{\text{it}}$. In practice, it is advantageous to neglect the off-diagonal matrix elements of $\mathbf{A}$, which is in the spirit of the zeroth-order Hamiltonian defined in Eq. (5.14). The resulting triples residual has the form

$$\Omega_{3,\rho}' = A_\rho^\rho t_{3,\rho}' + \Omega_{3,\rho}'\big|_{\hat{T}_3=0} \tag{5.47}$$

and allows to compute $\hat{T}_3$ in a single step:

$$t_{3,\rho}' = -\frac{1}{A_\rho^\rho}\Omega_{3,\rho}'\big|_{\hat{T}_3=0}. \tag{5.48}$$

With this choice, the last term in Eq. (5.45) vanishes and we obtain the following expression for the perturbative correction to the ic-MRCCSD energy:

$$\Delta E_{(T)} = \Delta E_{(T)}^{\text{direct}} + \underbrace{\mathbf{t}_1^{\prime\dagger}\mathbf{\Omega}_{1,T}' + \mathbf{t}_2^{\prime\dagger}\mathbf{\Omega}_{2,T}'}_{\mathbf{t}^{\prime\dagger}\mathbf{\Omega}_T'}. \tag{5.49}$$

Equations (5.48) and (5.49) are the working equations for ic-MRCCSD(T). In connection with a simultaneous exponential ansatz for the wave function, we have

$$\mathbf{\Omega}_3'\big|_{\hat{T}_3=0} = \langle\Psi_0|\boldsymbol{\tau}_3^{\prime\dagger}\big(\hat{H} + [\hat{H}, \hat{T}_1 + \hat{T}_2] + \tfrac{1}{2}[[\hat{H}_D, \hat{T}_1 + \hat{T}_2], \hat{T}_1 + \hat{T}_2]\big)|\Psi_0\rangle, \tag{5.50}$$

$$\Delta E_{(T)}^{\text{direct}} = \langle\Psi_0|\big([\hat{H}, \hat{T}_3] + \tfrac{1}{2}[[\hat{H}, \hat{T}_1 + \hat{T}_2], \hat{T}_3] + \tfrac{1}{2}[[\hat{H}, \hat{T}_3], \hat{T}_1 + \hat{T}_2]\big)|\Psi_0\rangle, \tag{5.51}$$

$$\mathbf{t}^{\prime\dagger}\mathbf{\Omega}_T' = \langle\Psi_0|\big(\hat{T}_1^\dagger + \hat{T}_2^\dagger\big)\big([\hat{H}, \hat{T}_3] + \tfrac{1}{2}[[\hat{H}_D, \hat{T}_1 + \hat{T}_2], \hat{T}_3] + \tfrac{1}{2}[[\hat{H}_D, \hat{T}_3], \hat{T}_1 + \hat{T}_2]\big)|\Psi_0\rangle. \tag{5.52}$$

In case of the sequential ansatz from Eq. (5.41), these expressions read

$$\mathbf{\Omega}_3'\big|_{\hat{T}_3=0} = \langle\Psi_0|\boldsymbol{\tau}_3'^\dagger\big(\hat{H} + [\hat{H},\hat{T}_1+\hat{T}_2] + [[\hat{H}_{\mathrm{D}},\hat{T}_1],\hat{T}_2]$$
$$+ \tfrac{1}{2}[[\hat{H}_{\mathrm{D}},\hat{T}_1],\hat{T}_1] + \tfrac{1}{2}[[\hat{H}_{\mathrm{D}},\hat{T}_2],\hat{T}_2]\big)|\Psi_0\rangle, \tag{5.53}$$

$$\Delta E_{(\mathrm{T})}^{\mathrm{direct}} = \langle\Psi_0|\big([\hat{H},\hat{T}_3] + [[\hat{H},\hat{T}_1],\hat{T}_3] + \tfrac{1}{2}[[\hat{H},\hat{T}_2],\hat{T}_3] + \tfrac{1}{2}[[\hat{H},\hat{T}_3],\hat{T}_2]\big)|\Psi_0\rangle, \tag{5.54}$$

$$\mathbf{t}'^\dagger\mathbf{\Omega}_{\mathrm{T}}' = \langle\Psi_0|\big(\hat{T}_1^\dagger + \hat{T}_2^\dagger\big)\big([\hat{H},\hat{T}_3] + [[\hat{H}_{\mathrm{D}},\hat{T}_1],\hat{T}_3]$$
$$+ \tfrac{1}{2}[[\hat{H}_{\mathrm{D}},\hat{T}_2],\hat{T}_3] + \tfrac{1}{2}[[\hat{H}_{\mathrm{D}},\hat{T}_3],\hat{T}_2]\big)|\Psi_0\rangle. \tag{5.55}$$

The ic-MRCCSD(T) method used in this work involves two steps scaling like N^7 with respect to the inactive orbital spaces. These steps occur in the evaluation of the terms $\langle\Psi_0|\boldsymbol{\tau}_3'^\dagger[\hat{H},\hat{T}_2]|\Psi_0\rangle$ and $\langle\Psi_0|\hat{T}_2^\dagger[\hat{H},\hat{T}_3]|\Psi_0\rangle$. In single-reference CCSD(T), one can exploit $\mathbf{\Omega}_3 = \mathbf{0}$ and insert the former term into the latter in order to avoid a second N^7 step. A similar simplification in ic-MRCCSD(T) requires additional approximations. A first step is the neglect of the terms printed in blue in Eqs. (5.50) to (5.55). Due to the sequential orthogonalization technique, this, in fact, is exact whenever η is set small enough to include all nearly linearly dependent single and double excitations. A further simplification results from treating $\hat{T}_1$ as second-order quantity. This makes no difference for terms present in the single-reference limit, but leads to an exclusion of the orange contributions from Eqs. (5.50) to (5.55). The green term is the last obstacle to a straightforward simplification of the second N^7 step and its neglect allows a reformulation according to

$$\langle\Psi_0|\hat{T}_2^\dagger[\hat{H},\hat{T}_3]|\Psi_0\rangle = -\sum_\rho A_\rho^\rho(t_{3,\rho}')^2 + \langle\Psi_0|[\hat{H},\hat{T}_2^\dagger\hat{T}_3]|\Psi_0\rangle. \tag{5.56}$$

This simplified version of ic-MRCCSD(T) will be denoted as ic-MRCCSD(T)$_{\mathrm{simp}}$.

5.6 Benchmark applications of ic-MRCC with perturbative triples

Unless otherwise noted, the results presented in this section rely on the same default approximations as the ones from chap. 4, namely a truncation of the BCH expansion after the double commutator and an elimination of redundant excitations by the sequential orthogonalization technique in conjunction with the threshold $\eta = 10^{-6}$. Whether a simultaneous exponential or the sequential ansatz from Eq. (5.41) is used will be individually specified.

5.6.1 Applications to model systems

Figures 28, 29 and 30 compare the performance of various ic-MRCC methods with perturbative triples with that of ic-MRCCSD and ic-MRCCSDT for the model systems BeH_2 (sec. 4.1.1), H_2O (sec. 4.1.4) and N_2 (sec. 4.1.5), respectively. In case of the BeH_2 model, both the simultaneous exponential $e^{\hat{T}_1+\hat{T}_2+\hat{T}_3}$ (Fig. 28a) and the sequential exponential $e^{\hat{T}_1}e^{\hat{T}_2+\hat{T}_3}$ (Fig. 28b) were used. A comparison of the energy curves for these two choices reveals that the sequential ansatz offers a slightly improved description of the energy surface in the region of maximum multireference character (see also Ref. [P5]). The reason for this is the inclusion of up to fourfold commutators in $\tilde{H}$, defined in Eq. (5.42),

(a) Simultaneous exponential ansatz.

(b) Sequential ansatz.
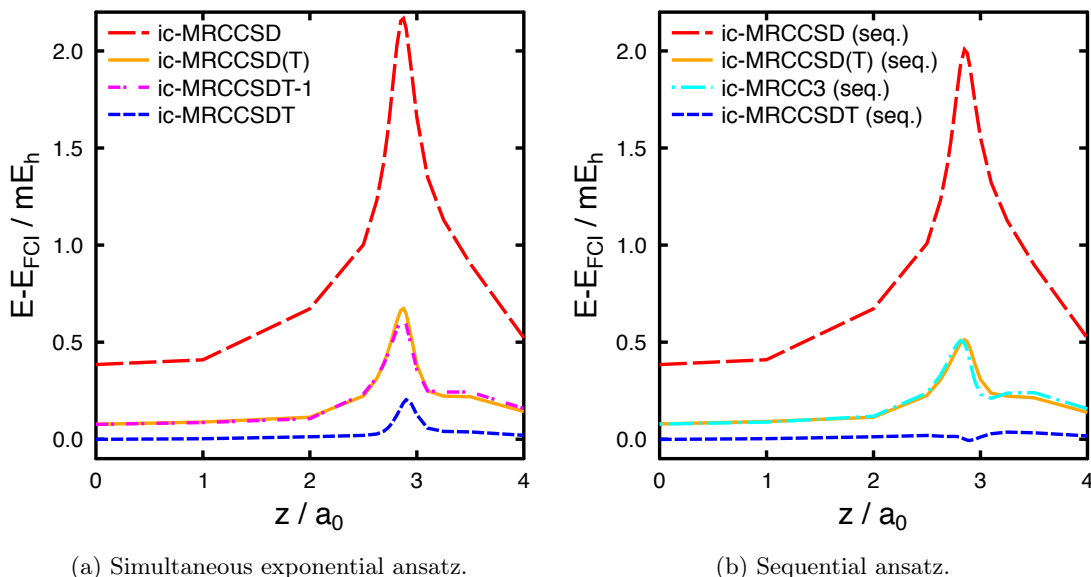
Figure 28: Performance of ic-MRCC methods with triples for the BeH_2 model system.

which takes into account higher powers of $\hat{T}_1$. A similar improvement can be obtained within the simultaneous exponential ansatz by lifting the commutator approximation in the eigenvalue equations (cf. Fig. 23b). The differences between the two ansätze are negligible for the H_2O and N_2 model systems [P5], and the present discussion arbitrarily employs the simultaneous ansatz for H_2O and the sequential one in case of N_2.

As shown in Figs. 28, 29a and 30a, the noniterative (T) correction is in excellent agreement with the iterative methods ic-MRCCSDT-1 and ic-MRCC3, with maximum deviations in the order of 0.1 mE_h for BeH_2 and H_2O and around 0.25 mE_h for N_2. The absolute errors of all triples models are much smaller than those of ic-MRCCSD. For both BeH_2 and H_2O, perturbative inclusion of triple excitations also reduces the NPE by more than a factor of three and recovers between 72% and 91% of the full triples contribution. In the N_2 dissociation, the perturbative triples models overestimate the effect of triple excitations between $R = 2.3\ a_0$ and $R = 3\ a_0$. However, the irregularities observed in this region mainly stem from the inclusion of small metric eigenvalues (cf. Fig. 22b), as becomes obvious from a comparison with results for $\eta = 10^{-4}$, shown in Fig. 30b.

While triple excitations seem unable to cure anomalies due to small metric eigenvalues near the equilibrium distance, they efficiently mitigate the drastic energy changes caused by the exclusion of small metric eigenvalues when approaching the dissociation limit. The rise in the ic-MRCCSD energy error observed when increasing the O–H bond length from $2.6R_e$ to $3.8R_e$ is reduced from 0.2 mE_h to less than 0.04 mE_h upon inclusion of triples. In case of the energy change that occurs when stretching the N–N bond beyond 5 a_0, triple excitations yield a similar reduction from 0.6 mE_h to 0.15 mE_h (for $\eta = 10^{-6}$).

Since the BeH_2 system involves a CAS(2,2), all triple excitations beyond the triples model {2} (defined in sec. 5.4.1) are exactly redundant and naturally excluded. A compilation of the nonredundant excitation types for this example is found in Fig. 24 (diagrams
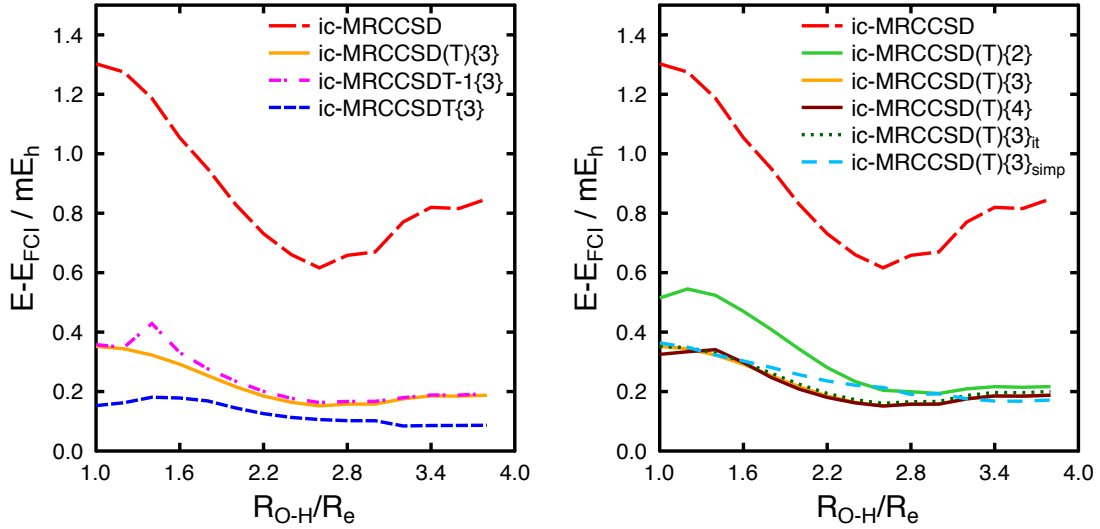
(a) Comparison of iterative and noniterative triples models.

(b) Comparison of various (T) corrections.

Figure 29: Performance of ic-MRCC methods with triples (using a simultaneous exponential ansatz) for the H_2O model system.

drawn in black). The results for the systems H_2O and N_2 shown in Figs. 29a and 30a are based on the model {3} and lead to similar conclusions as the higher model {4} [P3].

Figures 29b and 30b offer a comparison of the models {2}, {3} and {4} for ic-MRCCSD(T). The close agreement between the models {3} and {4} confirms that excitations with more than three active indices can be safely neglected. The difference between these models is somewhat larger for N_2 (0.15 mE_h) than for H_2O (0.03 mE_h), indicating that the importance of excitations with many active indices may increase with the size of the active space. On the other hand, the deviations of the model {2} values from those of {3} are significantly larger and suggest that one should not exclude excitations of the types $\hat{a}_{uij}^{vwa}$ and $\hat{a}_{uvi}^{wab}$. Table 5 collects the mean absolute errors (MAEs) and NPEs of the various ic-MRCCSD(T) methods. While both the MAEs and NPEs for H_2O and the MAEs for N_2 exhibit systematic improvements along the hierarchy ic-MRCCSD

Table 5: Mean average error (MAE) and nonparallelity error (NPE) of ic-MRCCSD and various ic-MRCCSD(T) models for the H_2O and N_2 model systems, given in mE_h.

Method	H_2O ($\eta = 10^{-6}$)		N_2 ($\eta = 10^{-4}$)	
	MAE	NPE	MAE	NPE
ic-MRCCSD	0.88	0.69	1.50	0.45
ic-MRCCSD(T){2}	0.32	0.35	0.65	0.29
ic-MRCCSD(T){3}	0.22	0.20	0.32	0.41
ic-MRCCSD(T){4}	0.22	0.19	0.27	0.44
ic-MRCCSD(T){3}$_{it}$	0.23	0.19	0.32	0.42
ic-MRCCSD(T){3}$_{simp}$	0.24	0.20	0.41	0.13
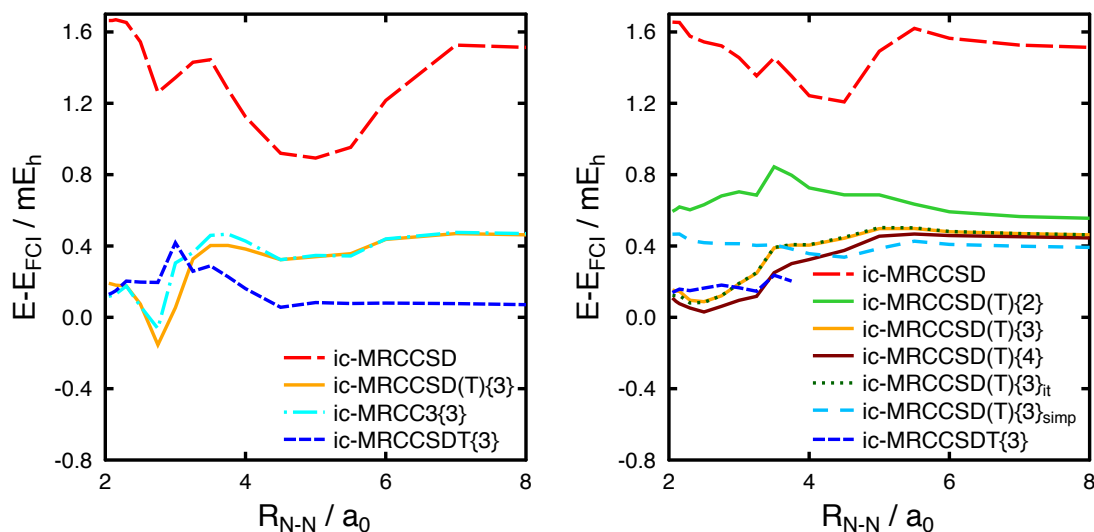
(a) Comparison of iterative and noniterative triples models, using $\eta = 10^{-6}$.

(b) Comparison of various (T) corrections, using $\eta = 10^{-4}$.

Figure 30: Performance of ic-MRCC methods with triples (using a sequential exponential ansatz) for the N_2 model system.

and ic-MRCCSD(T){n} with $n = 2, 3, 4$, the {2} model achieves the lowest NPE for N_2 within this hierarchy.

Remarkably, the noniterative (T) correction as enabled by Eq. (5.47) is in excellent agreement with the full-blown iterative variant ic-MRCCSD(T)$_{it}$ that also takes into account the off-diagonal elements of $\mathbf{A}$.[23] Further simplification of the ic-MRCCSD(T) method by omitting the blue and orange terms in Eqs. (5.53) to (5.55) was also found to yield only negligible changes (not shown), but additional neglect of the double commutator term in Eq. (5.53) (printed in green) has significant impact on the energy curves. Interestingly, this ic-MRCCSD(T){3}$_{simp}$ method yields the smallest NPE for the N_2 model system among the studied methods. To which extent this may be fortuitous is unclear at present. A possible explanation is that the green double commutator term in Eq. (5.53) facilitates an unphysical effect of doubles amplitudes associated with small metric eigenvalues on the triples amplitudes (cf. the discussion in sec. 4.3). This renders ic-MRCCSD(T)$_{simp}$ a promising candidate for future studies.

5.6.2 Structure and vibrational frequencies of ozone

Let us now turn to the ozone molecule for a first comparison of ic-MRCC results with experimental values. This example is scientifically relevant, as the detailed shape of the potential energy surface of O_3 has pronounced influence on the kinetics of the ozone formation and is still subject to active research [200–202]. The equilibrium structure and harmonic vibrational frequencies of O_3 have often served as a litmus test for multirefer-

---

[23]The number of iterations needed in the solution of the equations for $\hat{T}_3$ in ic-MRCCSD(T)$_{it}$ is relatively small (between 5 and 11 iterations in the present examples), which provides further evidence for the diagonal dominance of $\mathbf{A}$.

ence approaches [91, 189, 203–206], ever since there are accurate experimental benchmarks derived from high-resolution infrared [207] and ro-vibrational spectra [208], respectively. Accurate and reliable theoretical predictions of these properties call for a full-fledged and efficient MRCC method, as the computed values, in particular those for the frequencies of the stretching modes, are very sensitive to the treatment of both static and dynamic correlation and to the basis set used.

Adhering to the paradigm that dynamic correlation should be taken care of by the cluster amplitudes, the natural choice for the active space is the minimal active space consisting of two electrons in the orbitals $1a_2$ and $2b_1$, as suggested by an analysis of UHF natural orbital occupation numbers [63]. The following ic-MRCC results [P1, P3] are therefore based on a CAS(2,2). While an efficient calculation of equilibrium structures and vibrational frequencies generally requires an implementation of analytic energy gradients (e.g., Ref. [209]), a numerical differentiation is still feasible for the C_{2v}-symmetric ozone molecule, as there are only two total symmetric vibrational modes and one asymmetric mode. Hence, the numerical structure optimizations carried out for the ic-MRCC methods involved four calculations per iteration in a quasi-Newton–Raphson procedure (at a step size of $2 \cdot 10^{-6} a_0$). The computation of the frequencies $\omega_1(a_1)$ and $\omega_2(a_1)$ required a diagonalization of the Hessian matrix obtained from nine single-point calculations (at a step size of $10^{-3} a_0$). The evaluation of the asymmetric stretching frequency $\omega_3(b_2)$ needed only one additional calculation in C_s symmetry after having analytically eliminated the rotational degree of freedom (see, e.g., Ref. [210]).

Figure 31 shows the basis set convergence of the computed properties of O_3 for ic-MRCC methods with and without triple excitations in comparison to experimental data [207, 208] and the results from selected single-reference coupled-cluster methods [91, 211] and Mk-MRCC theory [91, 189]. All results are based on a frozen core that comprises the three O $1s$ orbitals and the cc-pVXZ basis sets, abbreviated by XZ. In case of the pentuple-zeta basis set, the ic-MRCC data as well as the vibrational frequencies from CCSD(T) rely on the simplified basis set 5Z', a variant of 5Z without h functions.[24]

Except for the bond angle θ_e, all of the studied properties exhibit significant contributions from triple excitations at the ic-MRCC level of theory. The ic-MRCCSDT results for the DZ and TZ basis sets allow a direct assessment of the perturbative triples models, revealing that the iterative models ic-MRCCSDT-1 and ic-MRCC3 slightly overestimate the effect of triples by up to 19% whereas ic-MRCCSD(T) recovers between 62% and 93% of the full triples contribution. This is in contrast to the good agreement between iterative and noniterative methods for the systems discussed in sec. 5.6.1. The likely reason is a coupling of triples amplitudes to doubles amplitudes, since the compliance of ic-MRCCSD(T) with ic-MRCCSD(T)$_{it}$ (see Tab. A15) rules out a significant coupling among triple excitations through the matrix $\mathbf{A}$. Considering the remarkable agreement between ic-MRCCSDT-1 and ic-MRCC3 (which is the only method using a sequential ansatz in the present section) on the other hand, higher powers of single excitations are negligible.

In connection with large basis sets, ic-MRCCSD(T) yields an excellent equilibrium structure, outperforming both CCSD(T) and Mk-MRCCSD(T). It also leads to more accurate vibrational frequencies than Mk-MRCCSD(T), although the absolute deviations

---

[24]In the ic-MRCC calculations, this is due to current technical limitations: While the integral code of GAMESS is restricted to g functions, the DALTON interface to GeCCo is limited to 255 orbitals.

(a) Ic-MRCC energies and legend.

(b) Equlibrium bond length.

(c) Equilibrium bond angle.

(d) Symmetric stretching mode.

(e) Symmetric bending mode.

(f) Asymmetric stretching mode.
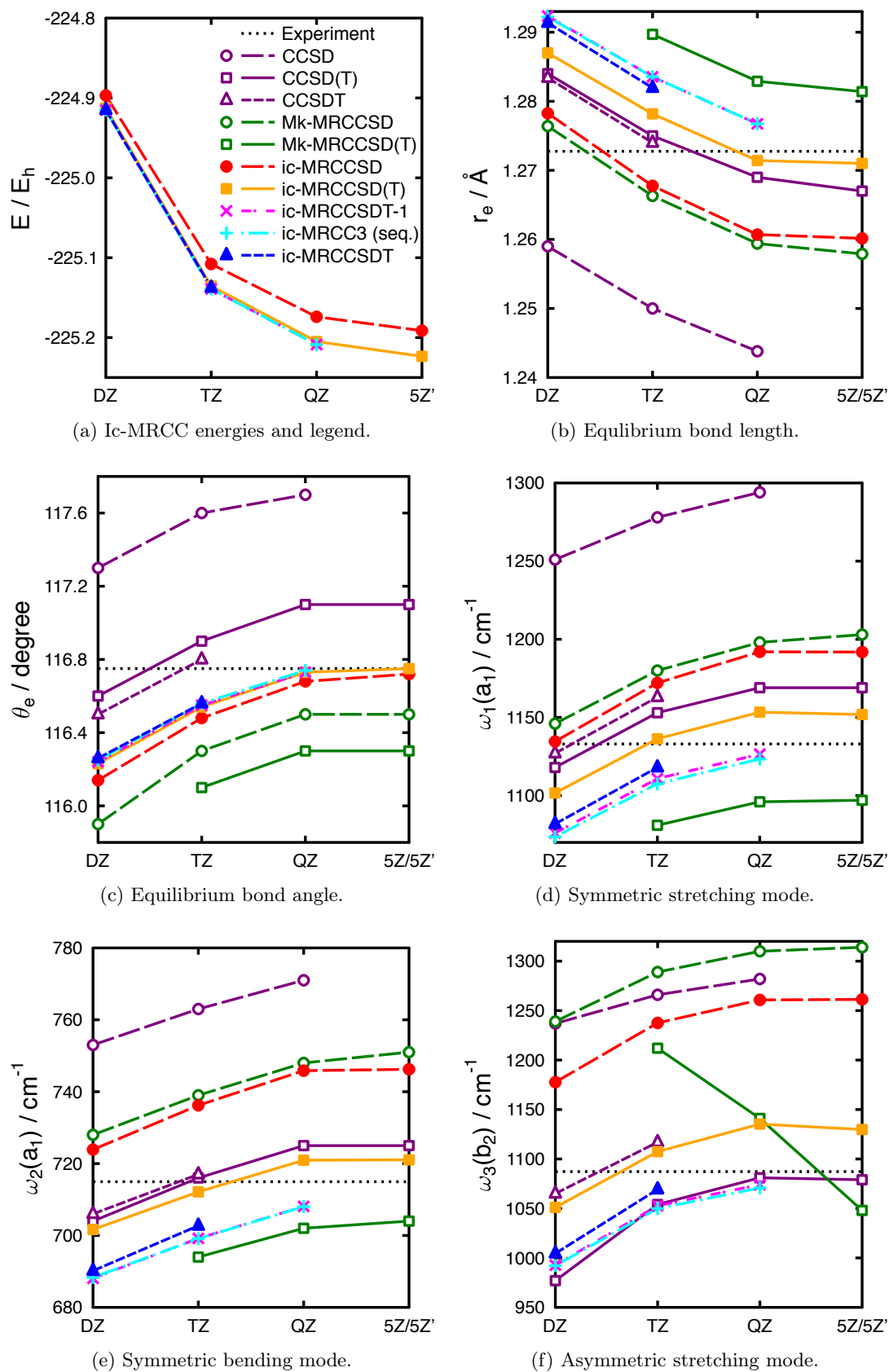
Figure 31: Energy, structure and harmonic vibrational frequencies of ozone.

Table 6: Comparison of the experimental equilibrium structure (in Å and degree) and harmonic vibrational frequencies (in cm^{-1}) of ozone with the best theoretical estimate derived from ic-MRCC theory (published in Ref. [P3]).

	r_e	θ_e	$\omega_1(a_1)$	$\omega_2(a_1)$	$\omega_3(b_2)$
ic-MRCCSDT/cc-pVTZ	1.2819	116.56	1118.0	703.0	1069.3
Basis set correction[a]	−0.0072	+0.21	+15.5	+8.9	+22.3
Core-correlation effect[b]	−0.0011	0.00	+1.2	+1.5	0.0
Best theoretical estimate	1.2737	116.77	1135	713	1092
Experiment[c]	1.2728	116.75	1133	715	1087

 [a] Difference between ic-MRCCSD(T)/cc-pVTZ and ic-MRCCSD(T)/cc-pV5Z'.
 [b] Difference between ic-MRCC3/cc-pVDZ using frozen core approximation and ic-MRCC3/cc-pCVDZ (see Ref. [212] for the basis set) with all electrons correlated.
 [c] Equilibrium structure from Ref. [207] and harmonic vibrational frequencies from Ref. [208].

from experiment for both stretching frequencies at the 5Z' level are still rather large (19 cm^{-1} for ω_1 and 43 cm^{-1} for ω_3). However, especially the asymmetric stretching frequency ω_3 is well known to be even more sensitive to the treatment of higher excitations in SRCC theory, as the results from the methods CCSDT-n (with n = 1a, 1b, 1c, 2, 3), CC3, CCSD(T) and CCSDT scatter over a range of several hundred wavenumbers [211]. It is also a showcase for the lack of orbital invariance in Mk-MRCC theory, where it varies dramatically with a rotation of the active orbitals 2a″ and 3a″ (in C_s symmetry) [91]. The anomalous basis set convergence of ω_3 at the Mk-MRCCSD(T) level is likely a consequence of this lack of orbital invariance.

The present data suggests that a reliable prediction of the vibrational frequencies of O_3 with ic-MRCC theory demands an iterative and preferably full account of triple excitations. Table 6 compiles best theoretical estimates, obtained from adding corrections for basis set and core-correlation effects to the ic-MRCCSDT/cc-pVTZ results and demonstrates that these estimates are in very good agreement with the experimental values.

5.6.3 Singlet-triplet splitting of benzynes

The diradicals *ortho-*, *meta-* and *para-*benzyne, illustrated in Fig. 32, are the parent systems of aryne chemistry [18, 213–215]. Their singlet-triplet splittings have been the subject of various experimental [216, 217] and theoretical studies [60, 91, 185, 218–222]. As the overlap between the essentially non-bonding orbitals at the radicalic centers shrinks from *o-* to *p-*benzyne, the multireference character increases, rendering particularly *p-*benzyne a valuable test system for multireference methods.

The ic-MRCC results for the singlet-triplet splittings [P3] are collected in Tab. 7 and stem from single-point calculations, using structures optimized at the CCSD(T)/cc-pVDZ and Mk-MRCCSD/cc-pVTZ levels of theory [P5, 91] (see the table for details) and keeping the six lowest CASSCF orbitals of C $1s$ type frozen. The data for the ic-MRCCSD and ic-MRCCSD(T) methods in connection with a CAS(2,2) and the basis sets cc-pVDZ and cc-pVTZ shows that the basis set error decreases from *o-* to *p-*benzyne while the effect of triple excitations increases. When approximately corrected for differences in the zero-point vibrational energies (ZPVE) between the singlet and triplet states,
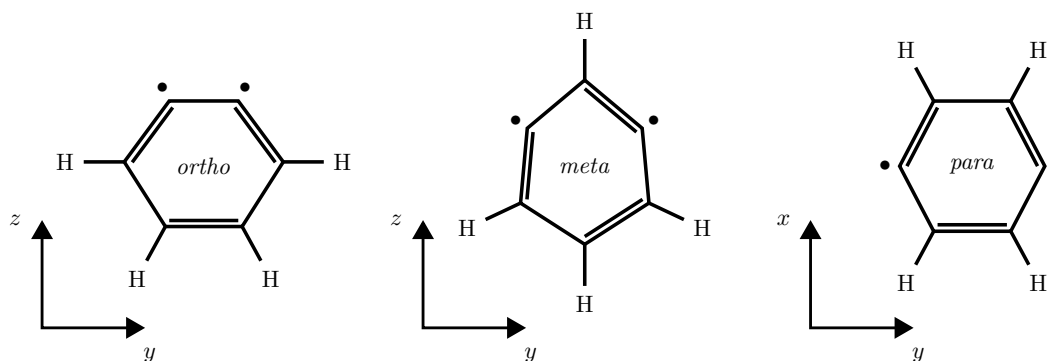
Figure 32: Structure of *o*-, *m*- and *p*-benzyne.

the ic-MRCCSD(T)/cc-pVTZ values for *o*- and *m*-benzyne already agree well with the experimental numbers within 1 kcal/mol and are likely to further improve with increased basis set size.

Whereas the perturbative triples correction is essential for an accurate singlet-triplet splitting of *m*-benzyne, in the case of *p*-benzyne it leads to an overshooting of the experimental value by about 1.4 kcal/mol. In order to judge whether an insufficient correlation treatment is the reason for this, let us analyze the (T) correction in greater detail. Table 8 lists the individual contributions of the various types of triple excitations to the (T) correction for the absolute energy of singlet *p*-benzyne and the singlet-triplet splitting.

Table 7: Singlet-triplet splittings of *o*-, *m*- and *p*-benzyne in kcal/mol (published in Ref. [P3]).

Method	Basis set	*ortho*[a] CAS(2,2) $M_s^{\text{trip}} = 0$	*meta*[a] CAS(2,2) $M_s^{\text{trip}} = 0$	*para*[b] CAS(2,2) $M_s^{\text{trip}} = 0$	CAS(2,2) $M_s^{\text{trip}} = 1$	CAS(8,8) $M_s^{\text{trip}} = 0$
ic-MRCCSD	cc-pVDZ	33.685	17.368	3.715	3.729	4.652
	cc-pVTZ	36.615	18.617	3.727		
	cc-pVQZ[c]			3.742		
ic-MRCCSD(T)	cc-pVDZ	33.832	18.274	4.763	4.776	4.772[d]
	cc-pVTZ	36.842	19.642	4.878		
ic-MRCCSDT[c]	cc-pVDZ			4.935		
ZPVE corr.[e]		−0.3	+0.7	+0.3		
Experiment[f]		37.5±0.3	21.0±0.3	3.8±0.4		

[a] Using RHF-CCSD(T)/cc-pVDZ structures for the singlet states and UHF-CCSD(T)/cc-pVDZ structures for the triplet states from Ref. [91].
[b] Using Mk-MRCCSD/cc-pVTZ structures for the singlet and triplet states from Ref. [P5].
[c] Published in Ref. [P5].
[d] Based on the triples model {3}. The values for {2} and {4} are 4.697 and 4.949, respectively.
[e] Zero-point vibrational energy (ZPVE) correction at CCSD(T)/cc-pVDZ level from Ref. [91].
[f] From Ref. [217].

Table 8: Individual (T) energy corrections from the various types of triple excitations for the singlet state of p-benzyne in kcal/mol, using the cc-pVTZ basis set. The contributions to the singlet-triplet splitting are given in parentheses.

n_h	$n_p = 1$	$n_p = 2$	$n_p = 3$
1	—	—	-0.172 $(+0.128)$
2	—	-2.294 $(+0.994)$	-4.965 (-0.055)
3	-0.041 $(+0.032)$	-2.195 $(+0.244)$	-23.168 (-0.193)

The largest contributions to the total energies come from the excitation types $\hat{a}_{uij}^{abc}$ and $\hat{a}_{ijk}^{abc}$, but they essentially cancel out when taking the difference between the energies of the singlet and triplet state. The dominant term in the (T) correction is instead due to excitations of the type $\hat{a}_{uij}^{vab}$, which we may interprete as an external double excitation coupled to an excitation within the active space. The importance of this excitation for the singlet-triplet splitting is physically reasonable, since $\hat{a}_{uij}^{vab}$ effectively provides individual double excitations for different reference functions and can therefore probe the difference between the two states. In line with the expected trend that singlet states have a larger correlation energy than triplet states [2], the resulting triples correction increases the singlet-triplet splitting. It is also in good agreement with the full triples contribution, as a comparison between ic-MRCCSD(T) and ic-MRCCSDT reveals for the cc-pVDZ basis set.

Another critical issue is the choice of the active space. Expanding the active space to a CAS(8,8) by inclusion of the π system leads to an ic-MRCCSD result close to the ic-MRCCSD(T) value for a CAS(2,2). This demonstrates that ic-MRCC theory in principle offers two strategies for achieving higher accuracy: An improvement of the coupled-cluster model or an increase of the active space's size. Remarkably, the effect of triple excitations is much smaller for the CAS(8,8) than for the CAS(2,2). In fact, both active spaces give about the same ic-MRCCSD(T) results, indicating that the ic-MRCCSD(T) method is accurate enough to work well in connection with minimal active spaces.

All of these findings support the high quality of the ic-MRCCSD(T) results and suggest that the correlation treatment is not the source for their deviation from the experimental singlet-triplet splitting of p-benzyne. The insensitivity with respect to basis set sizes between cc-pVDZ and cc-pVQZ also rules out a large basis set error. Core-correlation effects are also sufficiently small ($+0.037$ kcal/mol at the ic-MRCCSD/cc-pCVDZ level, see Ref. [212] for the basis set). If we assume that the use of the Mk-MRCCSD/cc-pVTZ equilibrium structures is unproblematic, the most likely source of error in the theoretical value is the correction from the difference of zero-point vibrational energies (cf. Refs. [218, 223, 224]). A computation of the structure and ZPVE correction using ic-MRCC theory could resolve this issue, but this requires the implementation of analytic energy gradients.

However, the problem might also lie in the experimentally derived value [217]. The singlet-triplet splitting appears in the UV photoelectron spectrum of the p-benzyne anion as a difference between the 0-0 transitions to the singlet and triplet states of the neutral molecule. Unfortunately, the origin of the triplet band cannot be clearly distinguished from the vibrational peaks of the singlet band. Therefore one cannot rule out a mis-

assignment of the triplet origin by one quantum of vibrational energy. This gives rise to the alternative values 2.1 kcal/mol and 5.5 kcal/mol for the singlet-triplet splitting of p-benzyne, the latter being in closer agreement with the current best estimate from ic-MRCC theory (5.2 kcal/mol).

Regarding the present application of ic-MRCC theory to triplet states, a few remarks on the spin purity of the ic-MRCC wave function are in place. The present spin orbital formalism is not spin-adapted and is found to give spin contamination in the wave function of nonsinglets. It also yields different energies for the various components of a spin multiplet. An example for this is the slight deviation between the energies of the triplet components with $M_s = 0$ and $M_s = 1$ for p-benzyne (denoted by M_s^{trip} in Tab. 7). In order to evaluate the spin contamination in the wave function, one could compute the spin expectation value $\langle\Psi_0|(1 + \hat{\Lambda})e^{-\hat{T}}\hat{S}^2 e^{\hat{T}}|\Psi_0\rangle$ [225]. Considering that it is straightforward to enforce the correct spin symmetry for the reference function (see sec. A.1.3), we can omit the trivial term $\langle\Psi_0|e^{-\hat{T}}\hat{S}^2 e^{\hat{T}}|\Psi_0\rangle = \langle\Psi_0|\hat{S}^2 e^{\hat{T}}|\Psi_0\rangle = S(S+1)$. Replacement of Λ by $\hat{T}^\dagger$ for convenience then leads to the following measure for the spin contamination:

$$\Delta\langle\hat{S}^2\rangle_{\text{a}} = \langle\Psi_0|\hat{T}^\dagger e^{-\hat{T}}\hat{S}^2 e^{\hat{T}}|\Psi_0\rangle. \tag{5.57}$$

Alternatively, we could simply look at the deviation of the spin expectation value of the excited function $\hat{T}|\Psi_0\rangle$ from the expected result:

$$\Delta\langle\hat{S}^2\rangle_{\text{b}} = \frac{\langle\Psi_0|\hat{T}^\dagger\hat{S}^2\hat{T}|\Psi_0\rangle}{\langle\Psi_0|\hat{T}^\dagger\hat{T}|\Psi_0\rangle} - \langle\Psi_0|\hat{S}^2|\Psi_0\rangle = \frac{\langle\Psi_0|\hat{T}^\dagger[\hat{S}^2,\hat{T}]|\Psi_0\rangle}{\langle\Psi_0|\hat{T}^\dagger\hat{T}|\Psi_0\rangle}. \tag{5.58}$$

In a rigorously spin-adapted theory based on a cluster operator expressed in terms of the generators of the unitary group [34,94,106,226–229], $E_p^q = \sum_{pq}(m_s(p) + m_s(q))^2\hat{a}_p^q$ (where $m_s(p)$ is the m_s quantum number of spin orbital φ_p), the cluster operator commutes with $\hat{S}^2$ such that both $\Delta\langle\hat{S}^2\rangle_{\text{a}}$ and $\Delta\langle\hat{S}^2\rangle_{\text{b}}$ vanish. While the degree of spin contamination in the ic-MRCCSD/cc-pVDZ wave function for the triplet state of p-benzyne with $M_s = 0$ is negligible ($\Delta\langle\hat{S}^2\rangle_{\text{a}} = 1.5 \cdot 10^{-5}$, $\Delta\langle\hat{S}^2\rangle_{\text{b}} = 8.1 \cdot 10^{-6}$), the $M_s = 1$ component exhibits somewhat larger spin contamination ($\Delta\langle\hat{S}^2\rangle_{\text{a}} = 9.2 \cdot 10^{-4}$, $\Delta\langle\hat{S}^2\rangle_{\text{b}} = 3.2 \cdot 10^{-3}$).[25]

It is worth noting that ic-MRCC calculations on the discussed benzyne systems are computationally feasible, despite a large number of terms in the ic-MRCC equations. For instance, on a single 3.0 GHz Intel® Xeon processor (and using 3.7 GB of virtual memory), the ic-MRCCSD/cc-pVTZ calculation on the singlet state of p-benzyne took about 8 h, which is only a factor of 3.3 more than the duration of a SRCCSD calculation performed with GeCCo under the same conditions. The (T) step required about 4 h (using 12.7 GB of virtual memory).[26] There is still room for improving the program's efficiency by factoring out suitable intermediates from the ic-MRCC equations, as the present work only employs intermediates of the types discussed in sec. 3.1.3.

---

[25]Generally, I have not observed any spin contamination in ic-MRCC wave functions for singlets (independent of the active space size).

[26]These timings should be taken with some caution, as they are sensitive to the amount of allocated virtual memory and to other processes running simultaneously on the same node. A more careful investigation of computational timings is certainly desirable, once more effort has been spent on making the implementation more efficient.

5.6.4   Structure and vibrational frequencies of Ni_2O_2

First-row transition metal compounds are challenging systems for multireference methods, as they often show strong static correlation due to partially filled d shells [19–21,230]. The binuclear transition metal oxide Ni_2O_2 has a rhombic structure of D_{2h} symmetry and an open-shell ^1A_g ground state [231]. Hübner and Himmel [232] have studied it with MRCI+Q in connection with a basis set of quadruple zeta quality and a large active space of 28 electrons in 16 orbitals (O $2p$, Ni $3d$), providing valuable benchmarks for several properties of Ni_2O_2. The ANO-RCC-VQZ basis set used in Ref. [232] is a subset of the relativistic atomic natural orbital basis set by Roos *et al.* [233, 234] and consists of a [$7s$,$6p$,$4d$,$3f$,$2g$] contraction on Ni and a [$5s$,$4p$,$3d$,$2f$] contraction on O. This basis set is also feasible for ic-MRCCSD calculations, but the presented applications of ic-MRCCSD(T) employ the smaller ANO-RCC-VTZ basis set (with a [$6s$,$5p$,$3d$,$2f$] contraction on Ni and a [$4s$,$3p$,$2d$] contraction on O) in order to keep the virtual memory requirements manageable. Like in Ref. [232], scalar relativistic effects are included by a second-order Douglas–Kroll–Hess transformation of the Hamiltonian [235, 236]. The frozen core in the correlation treatment contains the Ni $1s$, $2s$, $2p$, $3s$ and O $1s$ orbitals. This choice differs from the one in Ref. [232], in which the Ni $3p$ orbitals were also frozen.

The active space in the ic-MRCC calculations is a CAS(4,4) comprising the orbitals $1a_u$, $4b_{2g}$, $2b_{1g}$ and $6b_{3u}$ (depicted in Fig. 33). For these orbitals, UHF/ANO-RCC-VTZ yields occupation numbers of 1.304, 0.869, 1.131 and 0.696, respectively, whereas the remaining natural orbitals have occupancies below 0.002 or above 1.998, indicating that a CAS(4,4) is the minimal active space. The CASSCF calculations with 28 electrons in 18 orbitals (O $2p$, Ni $3d$, $4s$) from Ref. [232] support this, showing occupation numbers beyond 0.05 and 0.95 for all natural orbitals except those four.

Table 9 compares the equilibrium structure and the symmetric harmonic vibrational frequencies of Ni_2O_2 obtained from MRCI+Q/ANO-RCC-VQZ with the current results from ic-MRCCSD/ANO-RCC-VQZ and those from both ic-MRCCSD and ic-MRCCSD(T){2} in connection with the ANO-RCC-VTZ basis set and the sequential ansatz from Eq. (5.41). Energy gradients in the structure optimizations were evaluated by numerical differentiation, and the harmonic frequencies were determined by diagonalizing the mass-weighed Hessian matrix, following a computation of the second derivatives in a basis of symmetry-adapted coordinates. The ic-MRCCSD values for the minimal active space differ from the MRCI+Q/CAS(28,16) results by up to 0.03 Å in the structural parameters and 22 cm^{-1} in the vibrational frequencies. The basis set error for ANO-
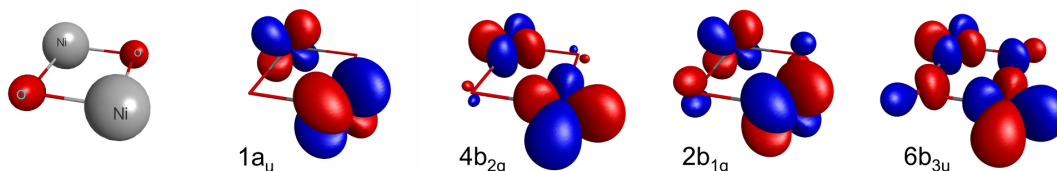
Figure 33: Structure of Ni_2O_2 (from ic-MRCCSD(T){2}/ANO-RCC-VTZ) and contour plots of the active orbitals (contour value=0.06).

Table 9: Equilibrium structure (in Å and degrees) and harmonic vibrational frequencies for the symmetric modes (in cm^{-1}) of the ^1A_g ground state of $^{58}Ni_2^{16}O_2$ (published in Ref. [P3]). The basis sets ANO-RCC-VXZ are abbreviated by XZ.

Method	Basis	CAS	R_{Ni-Ni}	R_{O-O}	$\angle_{Ni-O-Ni}$	$\omega_1(a_g)$	$\omega_2(a_g)$
MRCI+Q[a]	QZ	(28,16)	2.289	2.750	79.6	706	342
ic-MRCCSD	QZ	(4,4)	2.26	2.73	79	728	358
ic-MRCCSD (seq.)	TZ	(4,4)	2.269	2.733	79.4	725	350
ic-MRCCSD(T){2} (seq.)	TZ	(4,4)	2.278	2.751	79.3	703	337
Experiment					79.3 ± 2[b]	—[c]	—[c]

[a] From Ref. [232], based on a larger frozen core including also the Ni $3p$ orbitals.
[b] From Ref. [231], estimated from the isotopic effects on the vibrational frequencies of the b_{3g} and b_{1u} modes for $^{58}Ni_2^{16}O^{18}O$ and $^{60}Ni_2^{16}O^{18}O$ (measured in a Ne matrix at 3 K).
[c] Not experimentally observed (IR-inactive vibrational modes).

Table 10: Vibrational frequency ω_1 (in cm^{-1}) for selected isotopologues of Ni_2O_2 (published in Ref. [P3]). The harmonic frequencies obtained at the ic-MRCC/ANO-RCC-VTZ level are compared to fundamental frequencies from experiment.

Isotopologue	ic-MRCCSD (seq.)	ic-MRCCSD(T){2} (seq.)	Experiment[a]
$^{58}Ni_2^{16}O_2$	725.3	703.1	—
$^{58}Ni_2^{16}O^{18}O$	706.6	684.9	685.16
$^{58}Ni^{60}Ni^{16}O^{18}O$	706.1	684.4	684.50
$^{60}Ni_2^{16}O^{18}O$	705.6	683.9	683.88
$^{58}Ni_2^{18}O_2$	687.5	666.2	—

[a] Fundamental frequencies from Ref. [231] (measured in a Ne matrix at 3 K).

RCC-VTZ is rather small compared to these differences, as switching to this basis set changes the Ni–Ni and O–O distances by less than 0.01 Å and the vibrational frequencies by at most 8 cm^{-1}. The perturbative triples correction brings the ic-MRCC results very close to the MRCI+Q benchmark, with interatomic distances agreeing within 0.01 Å and vibrational frequencies within 5 cm^{-1}. This very good agreement emphasizes the power of ic-MRCCSD(T) to provide high accuracy in combination with minimal active spaces.

While a rough estimate for the Ni–O–Ni angle is available from an experimental study [231] and is in accordance with the theoretical results, the symmetric vibrational modes of the most common isotopologue $^{58}Ni_2^{16}O_2$ induce no change in the dipole moment and are not accessible by IR spectroscopy. However, they are observable for those isotopologues of Ni_2O_2 that do not belong to the point group D_{2h}. For the frequency ω_1 of various isotopologues, Table 10 shows a comparison of the computed harmonic frequency to the fundamental frequency measured in a neon matrix at 3 K [231]. For ω_2, no experimental data is available. There is a stunning concordance of the ic-MRCCSD(T){2} results with the experimental values, showing deviations below 0.3 cm^{-1}. However, we may expect this to be caused by fortuitous error compensation of various neglected contributions, e.g., those from anharmonicity, matrix effects, larger basis sets, core-correlation or spin-orbit coupling.

5.6.5   Ring-opening of the azirine compound C_6H_7NO

In order to demonstrate the applicability of the ic-MRCCSD(T) method to larger systems than the ones in the previous sections, let us consider the ring-opening reaction of the bicyclic azirine compound C_6H_7NO depicted in Fig. 34. Banert *et al.* [237] attempted to synthesize this azirine compound by photolysis of 3-azido-2-methyl-cyclopent-2-en-1-one, but only found evidence for the formation of the vinyl nitrene shown on the right of Fig. 34. Theoretical calculations using CASSCF and CCSD(T) provided a possible explanation for this, as the triplet state of the vinyl nitrene was found to be energetically favored over the singlet azirine. The calculations by Banert *et al.* also indicated small activation energies for the corresponding ring-opening process.

The current ic-MRCC results [P3], collected in Tab. 11, are based on single-point calculations at the structures provided by Ref. [237]. Those are optimized with CASSCF, using a CAS(6,6) and the split-valence double-zeta basis set 6-31G*, also known as 6-31G(d), which has additional *d*-type correlation functions on second-row atoms (C,N,O) [238, 239]. All of the six active orbitals are mainly composed of basis functions that are centered at the atoms of the three-membered azirine ring. The same basis set and active space was employed in the present ic-MRCC calculations, with the only difference that five spherical *d* functions were used per atom instead of six Cartesian components.

The ic-MRCCSD(T){3} results confirm the observations of Banert *et al.* and the relative energies agree well with the CCSD(T) values from Ref. [237] (within 0.9 kcal/mol), except for the transition state of the singlet. However, a meaningful comparison is hindered by the fact that CCSD(T) was applied to structures obtained with B3LYP [240] and additional diffuse functions for non-hydrogen atoms in the basis set (6-31+G*). We should also keep in mind that one would have to optimize the structures at the ic-MRCC level of theory in order to tap the method's full potential. Although the perturbative triples correction has a large effect on the absolute energies (about -20 kcal/mol), the relative energies from ic-MRCCSD and ic-MRCCSD(T){3} differ by less than 1 kcal/mol. In particular the results for the simplified variant ic-MRCCSD(T){3}$_{\text{simp}}$ are promising as they agree with the regular ic-MRCCSD(T){3} values within 0.15 kcal/mol for both absolute and relative energies.

The ic-MRCC calculations on C_6H_7NO were carried out on single Intel® Xeon and i7 processors with CPU frequencies ranging from 2.2 GHz to 3.3 GHz. Each ic-MRCCSD calculation employed 7 GB of virtual memory and took between 2.3 d and 10 d, mainly
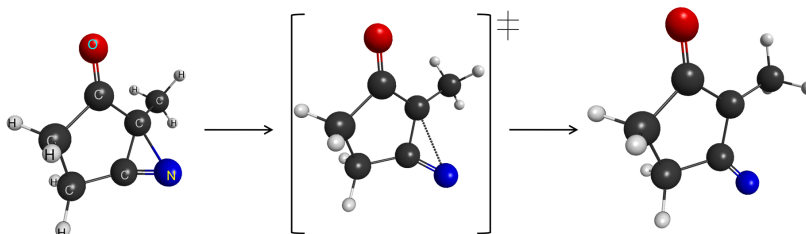
Figure 34: Ring-opening reaction of C_6H_7NO. Reproduced with permission from Ref. [P3]. Copyright 2012, American Institute of Physics.

Table 11: Reaction energies for the ring-opening of the bicyclic azirine compound C_6H_7NO through a transition state (TS) to a vinyl nitrene, for both the singlet and the triplet state. Absolute energies for the singlet azirine are given in E_h, all other energies are given relative to these values in kcal/mol.

Method	CASSCF[a]	CCSD(T)[b]	ic-MRCCSD[a]	ic-MRCCSD(T){3}	
Basis set	6-31G*	6-31+G*	6-31G*	6-31G*	
Structure[c]	CASSCF	B3LYP	CASSCF	CASSCF	
				regular[a]	simp
singlet azirine	-360.59242		-361.64114	-361.67342	-361.67319
singlet TS	6.62	12.04	7.58	7.65	7.57
singlet nitrene	3.70	5.06	5.24	5.26	5.23
triplet azirine	59.49	44.78	46.28	45.68	45.73
triplet TS	69.52	56.39	56.92	56.61	56.50
triplet nitrene	-3.63	-6.02	-6.98	-6.03	-6.17

[a] Published in Ref. [P3].

[b] From Ref. [237].

[c] Structures optimized with CASSCF(6,6)/6-31G* are from Ref. [237] and are based on Cartesian Gaussians. In case of the singlet transition state, the active space used in Ref. [237] seems to have accidentally contained one C–O π orbital. The structure optimization was therefore revised with an active space located at the C–C–N heterocycle, which also gave a lower CASSCF energy.

depending on the number of iterations needed in the solution of the ic-MRCC equations. The number of iterations required for converging the residual norm below $\epsilon = 10^{-8}$ varied from 32 (for the triplet nitrene) to 88 (for the singlet nitrene), stressing the need for further improvements in the implemented solving algorithm (see app. A.1.4 for details). A single iteration typically took 2 h, which is about a factor of 40 longer than an iteration in a single-reference CCSD calculation under comparable conditions. Although this seems large, let us contrast it with the fact that MRCC methods based on the Jeziorski–Monkhorst ansatz multiply the timings of SRCC by the number of reference determinants, i.e., by a much larger factor of 400 in the current case. The rate determining step in each iteration was the evaluation of the ic-MRCC residual, Eq. (3.15). The solution of the eigenvalue equation, Eq. (3.5), which is performed in each iteration, took only about 4 min. The (T) step used 24 GB of virtual memory and required about 21 h on average. The neglect of the colored terms in Eqs. (5.50) to (5.52) reduced this duration to 9 h in an ic-MRCCSD(T){3}$_{simp}$ run.[27] These computational timings emphasize that especially the (T) correction is an efficient way to obtain highly accurate ic-MRCC results. Future efforts to speed up the ic-MRCC calculations have to focus on improvements in the solution of the ic-MRCCSD equations.

---

[27]Note that the reformulation proposed in Eq. (5.56) is not implemented yet.

6 Cumulant approximations in ic-MRCC theory

6.1 Motivation

According to the analysis from sec. 3.1.3, the computational scaling of the ic-MRCC method is polynomial with respect to the size of the active space, except for the evaluation of reduced density matrices and the relaxation of the reference function. As ic-MRCCSD truncated at the double commutator level involves reduced density matrices of the reference function up to rank five [see Eq. (3.14)], the evaluation of $\boldsymbol{\gamma}_5$ becomes the rate determining step for very large active spaces (cf. Fig. 9).

Because the $\boldsymbol{\gamma}_n$ are not extensive quantities, a simple neglect of high-rank density matrices ruins the size-extensivity of ic-MRCC theory and is therefore not an option. A neglect of high-rank cumulant tensors, on the other hand, offers a hierarchy of rigorously size-extensive approximations with a reduced scaling with respect to the number of active orbitals.

Approximations based on neglecting cumulants beyond $\boldsymbol{\lambda}_2$ or $\boldsymbol{\lambda}_3$ are quite common in quantum chemistry. They form the basis of methods using the contracted Schrödinger equation [241–244] or the related irreducible Brillouin conditions and irreducible contracted Schrödinger equations [245–248] and belong to the defining approximations of canonical transformation theory [41]. The extensive nature of cumulants has also motivated their use as central quantities of interest in density cumulant functional theory [249, 250].

6.2 Meaning of density cumulants

Cumulants are measures for statistical correlation. Given a set of random variables and a corresponding probability distribution, they can be derived through a cumulant generating function, the logarithm of the moment generating function [251–253]. The elements of the tensors $\boldsymbol{\lambda}_n$ are n-body cumulants in this statistical sense [30, 111] and they may be defined through a moment generating function in which the elements of $\boldsymbol{\gamma}_n$ are the moments [53, 54, 254–256].

A concrete interpretation of the cumulants is possible if we identify the density operator $\hat{\rho} = |\Psi\rangle\langle\Psi|$ with the underlying multivariate probability distribution [257, 258] and the occupation numbers of the spin orbitals with the random variables [P2]. The diagonal elements of $\boldsymbol{\gamma}_n$ are then readily interpreted as joint probabilities for finding n spin orbitals occupied, and the corresponding diagonal elements from $\boldsymbol{\lambda}_n$ are the covariances of the orbital occupation numbers [57]. For a generalization of this interpretation to the off-diagonal elements of $\boldsymbol{\gamma}_n$ and $\boldsymbol{\lambda}_n$, the reader is referred to Ref. [P2].

6.3 Formulation of cumulant approximations

The most straightforward way to realize cumulant approximations in the current formulation of ic-MRCC theory is to expand the reduced density matrices that appear in the residual expression in terms of cumulants and then neglect cumulants beyond a maximum rank $n_{\max}$. In a general form, the relations between density matrices and cumulants

read [P2]

$$\gamma_n = \sum_{\sum_{j=1}^n k_j \cdot j = n} \frac{1}{\prod_{j=1}^n k_j!} \bigotimes_{j=1}^n {}_A \boldsymbol{\lambda}_j^{\otimes_A k_j}, \tag{6.1}$$

where the superscripted $\otimes_A$ denotes powers of the modified wedge product from Eq. (3.55), e.g., $\boldsymbol{\lambda}_j^{\otimes_A 3} = \boldsymbol{\lambda}_j \otimes_A \boldsymbol{\lambda}_j \otimes_A \boldsymbol{\lambda}_j$. Up to rank five, these expressions are

$$\boldsymbol{\gamma}_1 = \boldsymbol{\lambda}_1, \tag{6.2}$$

$$\boldsymbol{\gamma}_2 = \boldsymbol{\lambda}_2 + \tfrac{1}{2}\boldsymbol{\lambda}_1^{\otimes_A 2}, \tag{6.3}$$

$$\boldsymbol{\gamma}_3 = \boldsymbol{\lambda}_3 + \boldsymbol{\lambda}_2 \otimes_A \boldsymbol{\lambda}_1 + \tfrac{1}{6}\boldsymbol{\lambda}_1^{\otimes_A 3}, \tag{6.4}$$

$$\boldsymbol{\gamma}_4 = \boldsymbol{\lambda}_4 + \boldsymbol{\lambda}_3 \otimes_A \boldsymbol{\lambda}_1 + \tfrac{1}{2}\boldsymbol{\lambda}_2^{\otimes_A 2} + \tfrac{1}{2}\boldsymbol{\lambda}_2 \otimes_A \boldsymbol{\lambda}_1^{\otimes_A 2} + \tfrac{1}{24}\boldsymbol{\lambda}_1^{\otimes_A 4}, \tag{6.5}$$

$$\boldsymbol{\gamma}_5 = \boldsymbol{\lambda}_5 + \boldsymbol{\lambda}_4 \otimes_A \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_3 \otimes_A \boldsymbol{\lambda}_2 + \tfrac{1}{2}\boldsymbol{\lambda}_3 \otimes_A \boldsymbol{\lambda}_1^{\otimes_A 2}$$
$$+ \tfrac{1}{2}\boldsymbol{\lambda}_2^{\otimes_A 2} \otimes_A \boldsymbol{\lambda}_1 + \tfrac{1}{6}\boldsymbol{\lambda}_2 \otimes_A \boldsymbol{\lambda}_1^{\otimes_A 3} + \tfrac{1}{120}\boldsymbol{\lambda}_1^{\otimes_A 5}. \tag{6.6}$$

A reduction of the computational scaling requires the insertion of the approximated expressions for $\boldsymbol{\gamma}_k$ with $k > n_{\max}$ into the ic-MRCC residual equation. However, this strategy leads to a dramatic increase in the number of terms, since for a given diagram, one has to consider all topologically unique possibilities for assigning indices from different cumulant tensors to the distinct indices of the original density matrix.

The number of additional terms can be significantly reduced by a compact factorization of the approximated expressions for the $\boldsymbol{\gamma}_k$. In a first step, we expand the remaining cumulant tensors again in terms of reduced density matrices, which is equivalent to taking the following relations as a starting point:

$$\boldsymbol{\gamma}_2 = \boldsymbol{\lambda}_2 + \tfrac{1}{2}\boldsymbol{\gamma}_1^{\otimes_A 2}, \tag{6.7}$$

$$\boldsymbol{\gamma}_3 = \boldsymbol{\lambda}_3 + \boldsymbol{\gamma}_2 \otimes_A \boldsymbol{\gamma}_1 - \tfrac{1}{3}\boldsymbol{\gamma}_1^{\otimes_A 3}, \tag{6.8}$$

$$\boldsymbol{\gamma}_4 = \boldsymbol{\lambda}_4 + \boldsymbol{\gamma}_3 \otimes_A \boldsymbol{\gamma}_1 + \tfrac{1}{2}\boldsymbol{\gamma}_2^{\otimes_A 2} - \boldsymbol{\gamma}_2 \otimes_A \boldsymbol{\gamma}_1^{\otimes_A 2} + \tfrac{1}{4}\boldsymbol{\gamma}_1^{\otimes_A 4}, \tag{6.9}$$

$$\boldsymbol{\gamma}_5 = \boldsymbol{\lambda}_5 + \boldsymbol{\gamma}_4 \otimes_A \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_3 \otimes_A \boldsymbol{\gamma}_2 - \boldsymbol{\gamma}_3 \otimes_A \boldsymbol{\gamma}_1^{\otimes_A 2}$$
$$- \boldsymbol{\gamma}_2^{\otimes_A 2} \otimes_A \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 \otimes_A \boldsymbol{\gamma}_1^{\otimes_A 3} - \tfrac{1}{5}\boldsymbol{\gamma}_1^{\otimes_A 5}. \tag{6.10}$$

After neglecting the $\boldsymbol{\lambda}_k$ with $k > n_{\max}$, the expressions for $\boldsymbol{\gamma}_k$ can be conveniently inserted into the relations for the higher ranks. This already results in simplifications for $\boldsymbol{\gamma}_k$ with $k > n_{\max} + 1$ due to the cancellation of terms. Taking $n_{\max} = 3$ as an example, insertion of the expression for $\boldsymbol{\gamma}_4$ into $\boldsymbol{\gamma}_5$ leads to zero prefactors for both $\boldsymbol{\gamma}_3 \otimes_A \boldsymbol{\gamma}_1^{\otimes_A 2}$ and $\boldsymbol{\gamma}_2 \otimes_A \boldsymbol{\gamma}_1^{\otimes_A 3}$, leaving only

$$\boldsymbol{\gamma}_5 \overset{n_{\max}=3}{\approx} \boldsymbol{\gamma}_3 \otimes_A \boldsymbol{\gamma}_2 - \tfrac{1}{2}\boldsymbol{\gamma}_2^{\otimes_A 2} \otimes_A \boldsymbol{\gamma}_1 + \tfrac{1}{20}\boldsymbol{\gamma}_1^{\otimes_A 5}. \tag{6.11}$$

In a second step, we factorize the resulting expressions by introducing appropriate

intermediates. The final expressions used in the present work are

$$\gamma_3 \approx \gamma_2 \otimes_A \gamma_1 - \tfrac{1}{3}\gamma_1^{\otimes_A 3} = \left(\gamma_2 - \tfrac{1}{3}\gamma_1^{\otimes_A 2}\right)\otimes_A \gamma_1, \tag{6.12}$$

$$\gamma_4 \approx \tfrac{1}{2}\gamma_2^{\otimes_A 2} - \tfrac{1}{12}\gamma_1^{\otimes_A 4} = \tfrac{1}{2}\left(\gamma_2 + \tfrac{1}{\sqrt{6}}\gamma_1^{\otimes_A 2}\right)\otimes_A \left(\gamma_2 - \tfrac{1}{\sqrt{6}}\gamma_1^{\otimes_A 2}\right), \tag{6.13}$$

$$\gamma_5 \approx \tfrac{1}{2}\gamma_2^{\otimes_A 2}\otimes_A \gamma_1 - \tfrac{1}{3}\gamma_2\otimes_A \gamma_1^{\otimes_A 3} + \tfrac{1}{20}\gamma_1^{\otimes_A 5}$$
$$= \tfrac{1}{2}\left(\gamma_2 - \tfrac{\sqrt{10}+1}{3\sqrt{10}}\gamma_1^{\otimes_A 2}\right)\otimes_A \left(\gamma_2 - \tfrac{\sqrt{10}-1}{3\sqrt{10}}\gamma_1^{\otimes_A 2}\right)\otimes_A \gamma_1 \tag{6.14}$$

for $n_{\max} = 2$,

$$\gamma_4 \approx \gamma_3\otimes_A \gamma_1 + \tfrac{1}{2}\gamma_2^{\otimes_A 2} - \gamma_2\otimes_A \gamma_1^{\otimes_A 2} + \tfrac{1}{4}\gamma_1^{\otimes_A 4}$$
$$= \gamma_3\otimes_A \gamma_1 + \tfrac{1}{2}\left(\gamma_2 - \tfrac{\sqrt{2}+1}{\sqrt{2}}\gamma_1^{\otimes_A 2}\right)\otimes_A \left(\gamma_2 - \tfrac{\sqrt{2}-1}{\sqrt{2}}\gamma_1^{\otimes_A 2}\right), \tag{6.15}$$

$$\gamma_5 \approx \gamma_3\otimes_A \gamma_2 - \tfrac{1}{2}\gamma_2^{\otimes_A 2}\otimes_A \gamma_1 + \tfrac{1}{20}\gamma_1^{\otimes_A 5}$$
$$= \gamma_3\otimes_A \gamma_2 - \tfrac{1}{2}\left(\gamma_2 + \tfrac{1}{\sqrt{10}}\gamma_1^{\otimes_A 2}\right)\otimes_A \left(\gamma_2 - \tfrac{1}{\sqrt{10}}\gamma_1^{\otimes_A 2}\right)\otimes_A \gamma_1 \tag{6.16}$$

for $n_{\max} = 3$, and

$$\gamma_5 \approx \gamma_4\otimes_A \gamma_1 + \gamma_3\otimes_A \gamma_2 - \gamma_3\otimes_A \gamma_1^{\otimes_A 2} - \gamma_2^{\otimes_A 2}\otimes_A \gamma_1 + \gamma_2\otimes_A \gamma_1^{\otimes_A 3} - \tfrac{1}{5}\gamma_1^{\otimes_A 5}$$
$$= \left(\gamma_4 - \left(\gamma_2 - \tfrac{\sqrt{5}+1}{2\sqrt{5}}\gamma_1^{\otimes_A 2}\right)\otimes_A \left(\gamma_2 - \tfrac{\sqrt{5}-1}{2\sqrt{5}}\gamma_1^{\otimes_A 2}\right)\right)\otimes_A \gamma_1 + \gamma_3\otimes_A \left(\gamma_2 - \gamma_1^{\otimes_A 2}\right) \tag{6.17}$$

for $n_{\max} = 4$.

The described approximations provide a systematic reduction of the computational scaling with respect to the active space at a moderate increase of the number of terms in the residual expression. While the current formalism allows the numerical assessment of cumulant approximations in ic-MRCC theory and pilot applications to systems with large active spaces, a truly efficient use of cumulant approximations might require a fully GNO-based approach. This is because the evaluation of the GNO-based residual with the extended Wick theorem yields expressions in terms of cumulants from the start, such that a neglect of cumulant tensors goes along with a simplification of the equations.

6.4 Impact of cumulant approximations in ic-MRCC theory

Before studying the impact of cumulant approximations in ic-MRCC theory, it is worth investigating the magnitude of the cumulant tensors for CASSCF wave functions. A detailed analysis for a variety of systems was published in Ref. [P2]. The current discussion will be limited to the H_2O and N_2 model systems described in secs. 4.1.4 and 4.1.5. The cumulant norm

$$\|\boldsymbol{\lambda}_n\| = \sum_{\substack{u_1<\cdots<u_n \\ v_1<\cdots<v_n}} \left(\lambda_{u_1\cdots u_n}^{v_1\cdots v_n}\right)^2 \tag{6.18}$$

is an orbital invariant quantity [P2] and serves as a convenient measure for a cumulant tensor's magnitude (cf. Ref. [259]).

Figure 35a shows how the cumulant norms vary along the symmetric dissociation of the H_2O molecule. As the underlying CASSCF wave function is based on a CAS(4,4), the highest nonzero reduced density matrix is γ_4. However, the rank of the cumulant

(a) $\|\boldsymbol{\lambda}_n\|$ along H_2O dissociation.   (b) $\|\boldsymbol{\lambda}_n\|$ for H_2O (log. scale).   (c) $\|\boldsymbol{\lambda}_n\|$ for N_2 (log. scale).
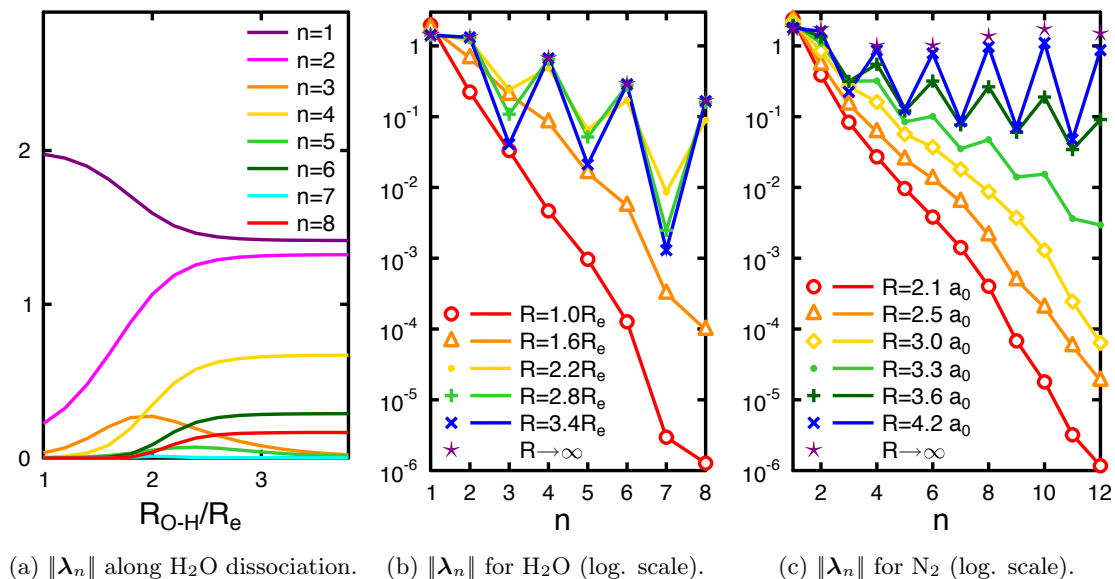
Figure 35: Cumulant norms for CASSCF wave functions of model systems.

tensors is only restricted by the number of active spin orbitals, hence up to eight-body tensors appear. While the high-rank cumulants are close to zero near the equilibrium structure, those with even rank converge towards significant values in the dissociation limit. Even the eight-body cumulant, which consists of a single element only, assumes a value of $\boldsymbol{\lambda}_8 \approx -0.166667$. Although CASSCF gives a size-consistent description for the isolated atoms, cumulant tensors beyond $\boldsymbol{\lambda}_4$ (the highest possible cumulant tensor on the O fragment) do not vanish despite their extensivity. This is because the atoms' spins are coupled to a singlet, rendering the CASSCF wave function not multiplicatively separable and causing nonzero values for both $\|\boldsymbol{\lambda}_6\|$ and $\|\boldsymbol{\lambda}_8\|$.

Figure 35b reveals how the cumulant norms depend on the rank for selected points on the potential energy surface. Figure 35c plots similar data for the N_2 model, where the CAS(6,6) allows cumulant tensors up to $\boldsymbol{\lambda}_{12}$. Close to the equilibrium structures, the cumulant norms decay approximately exponentially. This decreasing importance of cumulants with increasing rank is expected for regions dominated by dynamic electron correlation [54] and provides a justification for truncating the cumulant series. However, with rising multireference character along the dissociation path, the decay of the cumulant series becomes less pronounced. In the dissociation limit, cumulants with high ranks are found to take exceedingly large values, such as $\boldsymbol{\lambda}_{12} \approx -1.492119$ in the case of N_2.

An interesting aspect visible in Figs. 35b and 35c is that the cumulants with odd ranks (excluding $\boldsymbol{\lambda}_1$) vanish in the dissociation limit. While this is not a general feature of the cumulant series, it is characteristic for MCSCF wave functions with an electron-hole symmetry in the active space. In the dissociation limits of both H_2O and N_2, each active spin orbital assumes an occupation number of $1/2$, and for each determinant in the wave function there is another determinant with the same weight but with opposite occupancies of the active orbitals (i.e., with electrons and holes exchanged). We may then think of $\hat{\rho}$ as a symmetric distribution function centered around its first moment

(γ_1, i.e., half occupancy for each spin orbital), which implies that all odd cumulants vanish exactly [P2].

Let us now turn to a numerical assessment of cumulant approximations in ic-MRCC theory. Figure 36 shows the errors of the ic-MRCCSD energies with respect to FCI for the H_2O model system, using cumulant approximations with $n_{max} = \{2, 3, 4\}$ in connection with three different thresholds for discarding small metric eigenvalues, $\eta = \{10^{-2}, 10^{-3}, 10^{-4}\}$. Since the highest reduced density matrix in this case is γ_4, the results for $n_{max} = 4$ use no cumulant approximation and are identical to the data presented in Fig. 22a. Results for a similar study on the N_2 molecule, using the sequential ansatz from Eq. (5.41), are plotted in Fig. 37. Here, up to γ_5 arises for residual equations truncated at the double commutator level. The choice $n_{max} = 4$ constitutes a first approximation in this case, namely the neglect of λ_5.

Unfortunately, the current optimization algorithm (see app. A.1.4) failed to converge to the ic-MRCC solution in many cases. This problem affected especially calculations with a low n_{max} and became more severe when decreasing the threshold η. For this reason, many of the curves shown in Figs. 36 and 37 are incomplete. The available data indicates that the neglect of λ_5 is an excellent approximation, inducing energy errors less than 0.1 mE_h for the N_2 model. Discarding λ_4 in addition yields energy errors up to 0.3 mE_h, rendering $n_{max} = 3$ also a viable approximation. Although the tensors λ_4 exhibit large norms in the dissociation limits (0.7 for H_2O and 1.0 for N_2, respectively), their neglect only leads to small energy changes (about 0.01 mE_h for H_2O and 0.06 mE_h for N_2). However, the more drastic approximation $n_{max} = 2$ deteriorates the calculated energies for the N_2 model by up to 3.4 mE_h (Fig. 37a) and therefore cannot be recommended for calculations aiming at high accuracy.

It is striking that even the neglect of λ_5 triggers convergence problems for $R_{N-N} =$

(a) $\eta = 10^{-2}$.　　　　(b) $\eta = 10^{-3}$.　　　　(c) $\eta = 10^{-4}$.
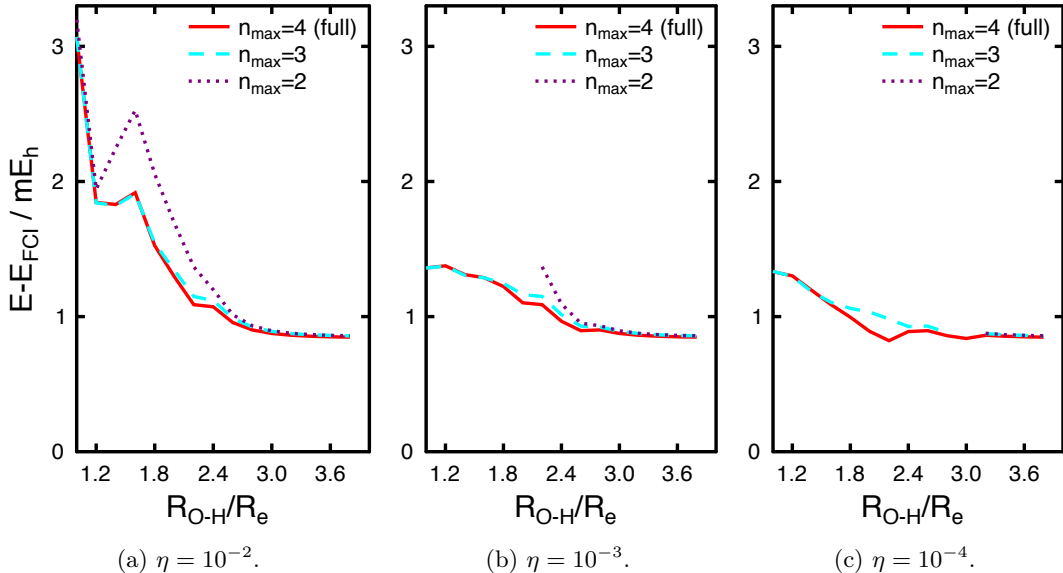
Figure 36: Impact of cumulant approximations in ic-MRCCSD calculations on the H_2O model system.

(a) $\eta = 10^{-2}$.   (b) $\eta = 10^{-3}$.   (c) $\eta = 10^{-4}$.
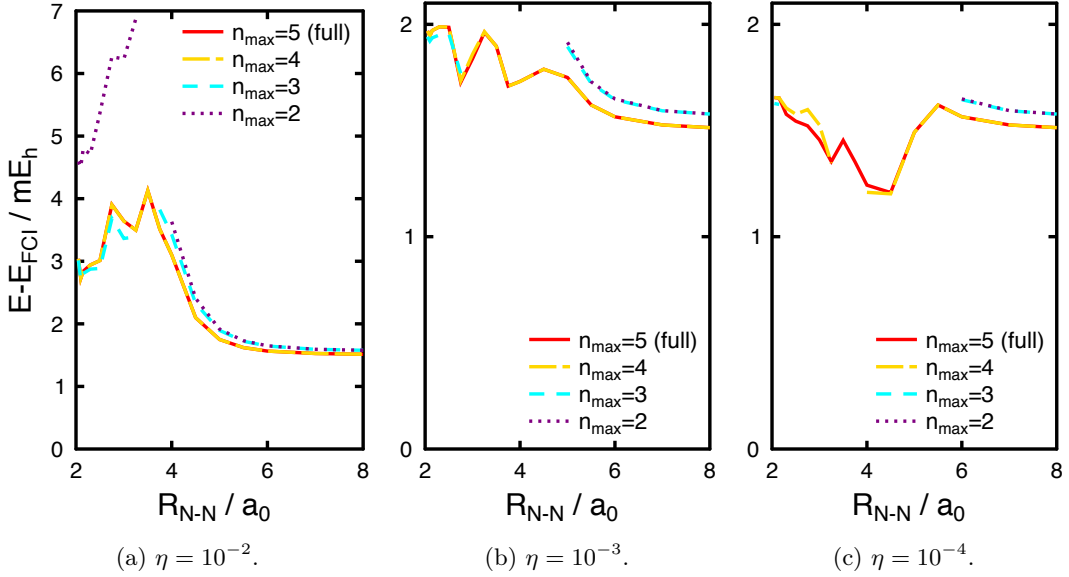
Figure 37: Impact of cumulant approximations in ic-MRCCSD calculations on the N_2 model system. Note the different scale for $\eta = 10^{-2}$.

$3.5\ a_0$ and $R_{\text{N--N}} = 3.75\ a_0$ when using $\eta = 10^{-4}$. The only neglected terms in this case are quadratic in $\hat{T}$. The fact that approximations in quadratic terms may spoil the convergence is remarkable, considering that the amplitude update step described in app. A.1.4 employs a linearized Jacobian and thereby only takes into account the impact of cluster amplitudes on the residual at linear order. A possible explanation for the deteriorating effect of cumulant approximations on the convergence is an overestimation of residual elements corresponding to small metric eigenvalues, which would be in line with the observed sensitivity of this problem with respect to the threshold η. We may view such a residual element as a projection of $\bar{H}|\Psi_0\rangle$ onto

$$\hat{\tau}'^\rho|\Psi_0\rangle = s_\rho^{-\frac{1}{2}} \cdot \left(\sum_\sigma \hat{\tau}^\sigma U_\sigma^\rho\right)|\Psi_0\rangle, \qquad (6.19)$$

i.e., the inverse of the square root of a metric eigenvalue multiplied with a nearly linearly dependent function that has a small norm. As $s_\rho^{-1/2}$ is large for a near-linear dependency, the smallness of the resulting residual element relies on the fact that the linear combination $\sum_\sigma \hat{\tau}^\sigma U_\sigma^\rho$ almost annihilates the reference function. However, the neglect of cumulants approximates the reduced density matrix, i.e., the contracted product of $\hat{C}_0^\dagger$ and $\hat{C}_0$. The operator $\hat{\tau}'^\rho$ then does not act on the original reference function any more and might loose its ability to compensate for the large value of $s_\rho^{-1/2}$. Attempts to restore the convergence of the ic-MRCC equations by damping the effect of nearly linearly dependent residual elements on the amplitude update have not been successful yet. Further effort is needed to resolve this problem.

6.5 Application to the chromium dimer

The chromium dimer is generally recognized as one of the most challenging systems for multireference approaches [19,260,261]. As in the dissociation limit each chromium atom has a ^{7}S high-spin ground state with six unpaired electrons in the $4s$ and $3d$ orbitals, the description of the $^1\Sigma_g^+$ ground state energy surface of Cr_2 requires at least an active space with twelve electrons in the orbitals $4s\sigma_g$, $4s\sigma_u$, $3d\sigma_g$, $3d\sigma_u$ and the orbital pairs $3d\pi_u$, $3d\pi_g$, $3d\delta_g$ and $3d\delta_u$. The resulting CAS(12,12) consists of as many as $107\,216$ determinants or $28\,784$ configuration state functions [262]. The bonding in Cr_2 is complex and has been the subject of controversial discussions: On the one hand, the six electrons from each Cr atom formally offer the formation of a hextuple bond, consisting of a $4s$–$4s$ σ bond and $3d$–$3d$ bonds of the types σ, π and δ. The δ bonds are somewhat weak, such that the effective bond order estimated from the orbitals' occupation numbers assumes a value of around 4.5 [19,261]. Nevertheless, this is still a high bond order and in line with the exceptionally short bond length of $r_e = 1.6788$ Å [263]. On the other hand, there are several factors that weaken the bond [260,264]: The $4s$ orbitals hardly contribute to the bonding at the equilibrium distance as they are much larger than the $3d$ orbitals and have an optimum distance for bonding around $R_{\text{Cr–Cr}} = 3$ Å. The filled $3p$ orbitals, on the contrary, have about the same size as the $3d$ orbitals and cause additional electron repulsion in the region of the bonding electrons. These effects may explain why the Cr_2 molecule has a relatively small harmonic force constant, $k_e = 353$ N/m corresponding to the vibrational frequency $\omega_e = 480.6 \pm 0.5$ cm^{-1} [264], and a dissociation energy of only $D_0 = 35.3$ kcal/mol [265], both of which are even lower than the corresponding values for the weakly bonded F_2 molecule ($k_e = 470$ N/m, $D_0 = 37$ kcal/mol) [260].

The chromium dimer has been extensively studied experimentally [263–269] and theoretically [260,262,270–281]. Of particular importance is the photoelectron spectroscopic study of Cr_2^- by Casey and Leopold [264], who have reconstructed the potential energy curve of Cr_2 from the transitions to 30 vibrational states. This potential is plotted in Fig. 38, using the value for r_e from Ref. [263] and $D_e = 1.50 \pm 0.04$ eV, which is an error-weighted average from the dissociation energies reported in the literature (1.47 ± 0.06 eV [268], 1.45 ± 0.10 eV [269] and 1.56 ± 0.06 eV [265]). The potential is highly anharmonic and exhibits a characteristic shelf below $R_{\text{Cr–Cr}} = 3$ Å, indicating regions with different types of bonding. While the $3d$–$3d$ bonds become weaker when separating the atoms, the interaction between the $4s$ orbitals first becomes more favorable. The resulting $4s$–$4s$ bond finally breaks when stretching the molecule beyond $R_{\text{Cr–Cr}} = 3$ Å. Between $R_{\text{Cr–Cr}} = 2.1$ Å and $R_{\text{Cr–Cr}} = 3$ Å, the shape of the potential is not uniquely determined by the experiment as the available spectral data in the corresponding energy range is sparse. Even the existence of a local minimum in the $4s$–$4s$ bonding region is in accordance with experimental data [264,266,267] and has remained an unresolved issue for decades.

The one common conclusion from the long history of theoretical studies on Cr_2 is that its description requires an accurate treatment of both static and dynamic correlation (including correlation of the $3d$ core orbitals) and the use of large basis sets. While CASSCF does not yield a bound potential at the experimental equilibrium bond length (see also Fig. 38), single-reference CCSD(T) calculations either fatally underestimate the bond energy ($D_e = 0.38$ eV) or dramatically overestimate the bond length ($r_e = 2.54$ Å), depending on whether a restricted or unrestricted Hartree–Fock reference is used [272]. The

failure of single-reference methods is understandable, considering that even at the experimental equilibrium distance, the leading configuration in the CASSCF wave function has a weight of only 47% [278].[28] Recent calculations using very large basis sets and advanced multireference approaches such as MR-AQCC based on a restricted active space of 1 516 configuration state functions [277] or CASPT2 in connection with density matrix renormalization group (DMRG, Refs. [282–284]) reference functions for a CAS(12,28) [279] agree well with the experimental potential energy curve (within 0.15 eV and 0.3 eV, respectively). Even though these studies present a convincing argument against a second minimum in the $4s$–$4s$ bonding region, it should be noted that some other multireference methods do predict a second minimum. Those include CIPT2 [275], a hybrid method with a CI-like treatment of excitations from active orbitals and a perturbative treatment of other excitations, and the third-order n-electron valence state perturbation theory (NEVPT3, Ref. [276]).

Figure 38 presents ic-MRCCSD and ic-MRCCSD(T){3} potential energy curves computed with the cc-pwCVTZ-DK basis set [285], keeping only $1s$, $2s$ and $2p$ orbitals uncorrelated and including scalar relativistic effects by a second-order Douglas–Kroll–Hess transformation [236]. The results are based on the sequential ic-MRCC ansatz, Eq. (5.41). The five-body cumulant tensor $\boldsymbol{\lambda}_5$ was neglected ($n_{\max} = 4$), as a full evaluation of $\boldsymbol{\gamma}_5$ would have been prohibitively expensive with the available resources. Figure 38 shows results for $\eta = 10^{-2}$ and $\eta = 10^{-3}$. Unfortunately, the latter choice did not yield
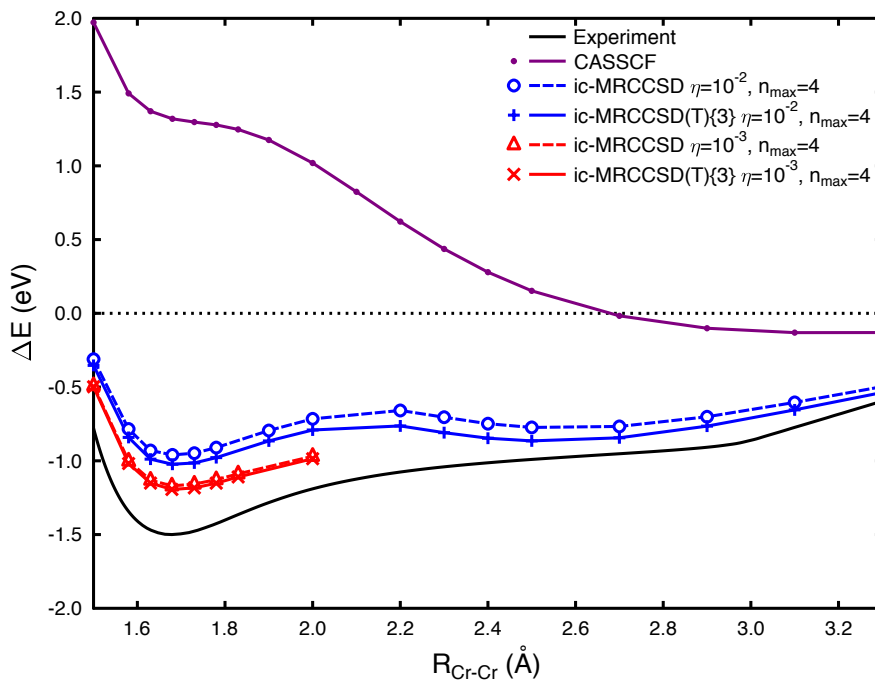
Figure 38: Potential energy surface of the $^1\Sigma_g^+$ ground state of Cr_2. The plotted energies $\Delta E = E - E(R_{Cr-Cr} = \infty)$ are relative to the dissociation limit.

---

[28]The same value is found in the present work. Interestingly, it increases to 53% during an ic-MRCCSD calculation.

convergence beyond $R_{\text{Cr–Cr}} = 2.0$ Å, whereas with $\eta = 10^{-2}$ most of the considered single-point calculations converged (except for $R_{\text{Cr–Cr}} = 1.83$ Å and $R_{\text{Cr–Cr}} = 2.1$ Å). The ic-MRCCSD(T){3} curve for $\eta = 10^{-2}$ differs from the experimental curve by up to 0.5 eV and has a second minimum at around $R_{\text{Cr–Cr}} = 2.5$ Å. While a significant portion of these deviations may be attributed to a too high threshold η (the results for $\eta = 10^{-2}$ and $\eta = 10^{-3}$ differ by up to 0.2 eV), the basis set error for cc-pwCVTZ-DK is still expected to be large. In particular, the accurate account of dynamic correlation in the $3d$–$3d$ bonding region requires basis functions with high angular momentum. At the DMRG-CASPT2 level of theory, the basis set error around r_e is as large as 0.6 eV [279], letting us expect an important basis set effect also for ic-MRCC theory. At present we cannot conclude whether the second minimum in the ic-MRCC curve is an artifact due to either the large threshold η or the insufficient basis set.[29]

Although the perturbative triples correction to ic-MRCCSD lowers the absolute energies by as much as 0.7 eV (see Tab. A20), its impact on the relative energies is minute (< 0.03 eV for $\eta = 10^{-3}$). This suggests that higher-order correlation effects beyond those recovered by ic-MRCCSD are negligible in the present example. Nevertheless, the difference between the curves obtained with $\eta = 10^{-2}$ and $\eta = 10^{-3}$ is disturbingly large. The increasing impact of the (T) correction when going from $\eta = 10^{-3}$ to $\eta = 10^{-2}$ indicates that the exclusion of metric eigenvalues beyond 10^{-2} is partially compensated by the triple excitations. However, within the triples model {3}, triple excitations with more than three active indices are neglected, allowing no compensation for excluded double excitations in the corresponding excitation classes.

We can derive spectroscopic constants from each ic-MRCCSD(T){3} curve by constructing a third-order polynomial in accordance with the four data points between $R_{\text{Cr–Cr}} = 1.63$ Å and $R_{\text{Cr–Cr}} = 1.78$ Å.[30] For $\eta = 10^{-2}$, this procedure gives $r_e = 1.692$ Å, $D_e = -1.02$ eV and $\omega_e = 407$ cm^{-1}. A decrease of the threshold η to 10^{-3} brings all of these three properties closer to the experimental values, yielding $r_e = 1.691$ Å, $D_e = -1.20$ eV and $\omega_e = 443$ cm^{-1}, respectively. However, the remaining deviations from experiment are still large, with an equilibrium distance that is too large by more than 0.01 Å, a dissociation energy too small in absolute terms by 0.3 eV and a harmonic vibrational frequency that is too low by about 40 cm^{-1}. An improvement of the ic-MRCC results requires either a full account of $\boldsymbol{\gamma}_5$ or an implementation of cumulant approximations that works in connection with low thresholds η. In addition, reliable ic-MRCC calculations call for a reduction of the basis set error, which can be achieved either by using larger basis sets or by combining the ic-MRCC method with explicitly correlated techniques, as described in Ref. [P7].

---

[29]The use of larger basis sets with ic-MRCC is currently hampered by the limitations of the GeCCo interfaces to Dalton and Gamess to either 255 orbitals or up to g functions, respectively.

[30]Note, however, that the present analysis can only provide rough estimates for the derivatives of the potential energy curve, since the spacing between the data points is rather large (≈ 0.05 Å) and a varying number of included metric eigenvalues causes discontinuities in the energy curve.

7 Summary and outlook

7.1 Summary and conclusions

The main subject of this thesis is the development and benchmarking of internally contracted multireference coupled-cluster (ic-MRCC) methods for the accurate theoretical description of systems with complicated electronic structure. The work accomplished in this effort embodies the following original contributions:

- The first implementation of ic-MRCC methods with the correct polynomial scaling of computational cost with respect to system size. This has been achieved through the automated expansion and termwise factorization of the working equations using the GeCCo program. The implementation is in principle general, allowing for arbitrary excitation levels, and is currently used for ic-MRCC with singles and doubles (ic-MRCCSD) and ic-MRCC with singles, doubles and triples (ic-MRCCSDT).

- A detailed analysis of ic-MRCC theory that provides a deep understanding of the theory's formal structure and complexity, its computational scaling and its most important formal properties, with an emphasis on size-extensivity.

- The development of two techniques for removing redundancies from the parameterization of the wave function. While the first technique is core-extensive (i.e., extensive with respect to a growing number of inactive electrons and orbitals), the second is fully size-extensive (also with respect to a growing active space) and constitutes the first realization of an ic-MRCC theory that takes the multiconfigurational reference function as Fermi vacuum in the second-quantized formalism.

- The development of both iterative and noniterative ic-MRCC methods that approximately account for triple excitations, in particular the ic-MRCCSD(T) method.

- A reformulation of the theory of density matrix cumulants that simplifies their use (e.g., for formulating cumulant approximations within the current ic-MRCCSD(T) method) and offers a clear interpretation of their meaning in quantum mechanics, and a numerical assessment of the cumulants' magnitudes for multiconfigurational wave functions.

- The investigation of different variants of ic-MRCC theory through benchmark applications to model systems. This includes the study of different schemes for removing redundancies from the cluster operator and different treatments of excitations within the active space, as well as an assessment of hierarchies of approximations, such as the truncation of the cluster operator after double or triple excitations, the neglect of nearly linearly dependent excitations with corresponding metric eigenvalues below the threshold η, truncations of the commutator series of the similarity transformed Hamiltonian, and cumulant approximations.

- Applications of the ic-MRCCSD(T) method to the equilibrium structure and vibrational frequencies of ozone and Ni_2O_2, the singlet-triplet splittings of benzynes, the ring-opening of the azirine compound C_6H_7NO, and the ground state energy surface of the chromium dimer.

The following summarizes the main findings of the present work and the conclusions drawn from them.

Due to the introduction of active orbitals that are only partially filled in the reference function, the cluster operator in ic-MRCC theory consists of a non-commuting set of excitation operators. This leads to a dramatic surge in the length and complexity of the amplitude equations as compared to single-reference coupled-cluster (SRCC) theory. The number of individual terms (i.e., unique products of tensor contractions) rises sharply with the level of nested commutators considered in the expansion of the similarity transformed Hamiltonian. A full implementation of ic-MRCCSD would involve up to eightfold commutators and hundreds of thousands of individual terms. However, truncation after the twofold commutator is found to preserve the inherent accuracy of the method (corroborating results by Evangelista and Gauss [38]) while reducing the number of terms to less than 10 000. Though this is still a large number of terms, it is manageable with the help of automated techniques like the ones provided by the GECCo program.

Key to an efficient implementation of this ic-MRCCSD method is a reduction of the number of those terms that dominate the computational expense. Through a suitable choice of intermediate expressions, the terms scaling most steeply with respect to the inactive orbital spaces can be combined to yield a scaling comparable to single-reference CCSD (N^6) in the limit of small active spaces. When increasing the size of the active space, the method's cost will eventually scale exponentially with the number of active electrons due to the necessity to compute the reduced density matrices of the reference function (up to rank five, $\boldsymbol{\gamma}_5$). The optional relaxation of the reference function also contributes to an exponential scaling with respect to the active space's size. In most realistic applications, however, there will be far more inactive than active orbitals. Although terms scaling with respect to both the inactive and active orbital spaces may become rate determining, their scaling is always polynomial and does not exceed $N \cdot M_{\text{act}}^{10}$. The ic-MRCC method therefore has a fundamental advantage over Jeziorski–Monkhorst-based MRCC theories in which each term from SRCC inherits an additional factorial scaling with respect to the active space. This computational advantage is already realized within the GECCo program, as ic-MRCCSD calculations on the azirine compound C_6H_7NO took by about a factor of 40 longer than a corresponding CCSD calculation, whereas the number of reference determinants (400) in this case lets one expect a Jeziorski–Monkhorst-based calculation (without internal contraction of inactive excitations) to take an additional factor of ten times longer.

Ic-MRCC theory provides a hierarchy of size-extensive methods (ic-MRCCSD, ic-MRCCSDT, etc.) that systematically converge towards the exact (FCI) solution of the Schrödinger equation within a given basis set. Further, it is invariant with respect to a rotation of the active orbitals, a feature not shared by most other MRCC approaches.

However, the internally contracted ansatz may be considered as overparameterized, as generally linear dependencies between the ic-MRCC equations prevent a unique determination of all cluster amplitudes. A solution to this problem is a well-defined procedure for eliminating redundant parameters from the cluster operator. Although there is some flexibility in setting up such a procedure, the requirements of orbital invariance and size-extensivity pose stringent constraints to its definition, rendering it not completely arbitrary. The methods studied in this work rely on a canonical orthogonalization of the set of functions that are created from the reference function through single, double,

and—if needed—higher excitation operators, thereby fulfilling orbital invariance (proven in Ref. [38]). In the course of this orthogonalization, excitations corresponding to eigenvalues of the excited functions' overlap matrix below a threshold η are discarded in order to remove exact and near-linear dependencies. In the current formalism, the excitation operators act on spin orbitals (not spatial orbitals) and are normal ordered with respect to a determinant composed of all filled inactive orbitals. While a full orthogonalization covering all excitation levels at once turns out to spoil size-extensivity, a sequential procedure that starts with single excitations and successively orthogonalizes the functions of each excitation level to the ones from lower levels regains core-extensivity. Both schemes yield similar results for ic-MRCCSD, but upon inclusion of triple excitations, the core-extensive treatment of double excitations is crucial for a systematically enhanced accuracy. The newly developed perturbative triples models therefore employ the sequential orthogonalization.

A fully size-extensive ic-MRCC method is obtained when switching to a generalized normal ordering (GNO) with respect to the multiconfigurational reference function. This is because in the evaluation of the overlap between excited functions, the extended Wick theorem involves as contraction tensors the reduced density matrix cumulants of the reference function, which are extensive quantities and thereby prevent inextensive contributions to the cluster amplitudes during the orthogonalization step. In the present implementation of the GNO-based ic-MRCCSD method, the coupled-cluster residual is still evaluated in a particle-hole normal ordering and then transformed to its representation in the GNO.

A critical issue in ic-MRCC theory is the choice of the aforementioned threshold η. Whereas the internally contracted variant of multireference configuration interaction (MRCI) is rather robust with respect to η and allows one to set η to very low values (limited only by numerical precision), the same is not true for ic-MRCC theory. When studying the N_2 molecule near its equilibrium bond length, too small thresholds are found to cause large artificial energy contributions, indicating that the elimination of linear dependencies within the linear excitation space is not perfectly compatible with the exponential cluster ansatz. Based on this study, the recommended choice for η is the range $10^{-6} \leq \eta \leq 10^{-3}$.

The benchmark applications of ic-MRCCSD to the potential energy surfaces of six selected model systems (BeH_2, HF, LiF, H_2O, N_2 and Be_3) certify that the method already yields chemical accuracy in many cases, showing absolute errors with respect to FCI of up to 10 kJ/mol and relative errors (nonparallelity errors) below 5 kJ/mol. Inclusion of triple excitations significantly reduces the remaining errors. Both absolute and relative errors for the model systems BeH_2, H_2O and N_2 are as small as 1 kJ/mol for the ic-MRCCSDT method. The noniterative incorporation of triples through ic-MRCCSD(T) recovers most of this improvement over ic-MRCCSD, yielding errors below 2 kJ/mol.[31]

The enhanced accuracy of ic-MRCCSD(T) comes at a moderate extra effort, which for small active spaces can be characterized by two steps scaling like N^7 and may be further reduced to a single N^7 step in the future. Applications to H_2O and N_2 indicate that one may safely neglect triple excitations with more than three active indices without

---

[31]As these error estimates rely on small basis sets of double-zeta quality (with polarization functions), they might increase somewhat for larger basis sets.

jeopardizing the inherent accuracy of the ic-MRCCSD(T) method. The resulting model, denoted as ic-MRCCSD(T){3}, has the computational advantage that its orthogonalization step requires only the diagonalization of reduced density matrices of the reference function up to rank three instead of five.

In applications of ic-MRCC theory to the equilibrium structure and harmonic vibrational frequencies of the ozone molecule and the transition metal oxide Ni_2O_2 as well as to the singlet-triplet splittings of benzynes and the ground state energy surface of Cr_2, the (T) correction on top of ic-MRCCSD consistently improves the agreement between computed and experimental results. The sole exception to this is the singlet-triplet splitting of p-benzyne, where the experimentally derived value is disputable. The ozone molecule turns out to be a challenging case for ic-MRCC theory. Especially the computed stretching frequencies are very sensitive to the treatment of triple excitations. The iterative and noniterative approximate triples models yield results differing by tens of cm^{-1}, while the full ic-MRCCSDT method provides the most accurate estimate, agreeing with experiment within 5 cm^{-1} for all vibrational modes.

The good agreement of the ic-MRCCSD(T) results for Ni_2O_2 based on a minimal active space (four electrons in four orbitals) with values obtained from size-extensivity corrected MRCI using a much larger active space (28 electrons in 16 orbitals) indicates that the ic-MRCCSD(T) method is able to achieve high accuracy in connection with minimal active spaces. The accordance of singlet-triplet splittings of p-benzyne, calculated with ic-MRCCSD(T) using either the minimal active space (two electrons in two orbitals) or an enlarged active space (eight electrons in eight orbitals), supports this conclusion and demonstrates that ic-MRCCSD(T) is much less sensitive to the size of the active space than ic-MRCCSD. Further recent applications of ic-MRCC methods using different active space sizes provide additional evidence for these conclusions [P7].

The n-body cumulants ($\boldsymbol{\lambda}_n$) of the reduced density matrices encode the statistical information that the reference function has about the correlation between n orbitals' occupancies. A neglect of high-rank cumulants offers a reduced computational scaling with respect to the number of active orbitals without compromising the method's size-extensivity. The development of such cumulant approximations in ic-MRCC theory is still at an experimental stage. Numerical tests indicate that the neglect of both $\boldsymbol{\lambda}_5$ and $\boldsymbol{\lambda}_4$ preserves the accuracy of ic-MRCCSD, whereas neglecting $\boldsymbol{\lambda}_3$ may significantly influence the computed energies. Unfortunately, in the current implementation, these approximations lead to convergence problems that need to be addressed in the future. Nevertheless, the neglect of $\boldsymbol{\lambda}_5$ enabled first applications of ic-MRCC theory to the ground state energy curve of the chromium dimer, using twelve electrons in twelve active orbitals.

Overall, the implemented ic-MRCC methods are already applicable in a routine manner to systems with few atoms (containing up to eight second-row elements or two first-row transition metal atoms) and moderate active spaces (up to eight electrons in eight orbitals). Especially the ic-MRCCSD(T) method reliably yields chemical accuracy in connection with minimal active spaces, making it one of the most accurate multireference methods to date [29]. Considering that ic-MRCC theory is also one of the few MRCC approaches fulfilling the most important formal requirements (size-extensivity, moderate computational scaling, orbital invariance), it is fair to regard it as the most promising MRCC method of the contemporary literature. By providing both a deep theoretical understanding and a useful implementation of ic-MRCC theory, this work

has laid the cornerstone for future development of ic-MRCC methods. Numerous applications of ic-MRCC theory, whether to accurately describe chemically relevant systems with complicated electronic structure or to create valuable benchmarks for testing less reliable but more affordable alternative methods, will undoubtedly follow.

7.2 Future perspectives

Despite the progress achieved in this work, the development of ic-MRCC methods into standard tools of quantum chemistry requires additional efforts.

On the one hand, shortcomings of the present formulation and implementation of ic-MRCC therory should be addressed. The high formal complexity of the theory demands an efficient factorization of the working equations, possibly accompanied by additional approximations. A reduction of the memory requirements, especially in the (T) step, could be achieved by a batchwise evaluation of contributions from subsets of excitations. Further, an improvement of the solving algorithm is desirable, in order to alleviate convergence problems and reduce the number of iterations taken. The implementation of GNO-based ic-MRCC theory should be extended to allow the use of the sequential orthogonalization technique and the perturbative inclusion of triple excitations. The evaluation of the GNO-based residual through the extended Wick theorem might also be needed for exploiting the full potential of cumulant approximations in ic-MRCC theory. Another important step to improve the present ic-MRCC approach would be its spin-adaptation, based on using the generators of the unitary group as excitation operators [34, 94, 106, 226–229]. Finally, the disturbing sensitivity of ic-MRCC methods to the threshold η justifies the search for alternative ways to eliminate linear dependencies (see, e.g., Ref. [47]) or the introduction of suitable regularization or level-shift techniques (e.g., Ref. [52]).

On the other hand, further theoretical advances are needed to enhance the versatility of ic-MRCC theory. Present applications of ic-MRCC theory are limited to single-point energy calculations for small systems. An extension of the applicability to structure optimizations and the routine calculation of vibrational frequencies calls for an implementation of analytic energy gradients (e.g., Ref. [209]). The development of a linear response theory for ic-MRCC methods to access a wider range of molecular properties is already in progress, showing first promising results in the computation of excitation energies [286]. It is also worthwhile to explore alternative ways to target multiple states within ic-MRCC theory, for instance through the use of many-body equations [115]. There are several possibilities to enable ic-MRCC applications for larger systems. Explicitly correlated techniques accomplish this by reducing the basis set requirements, and are already being employed in ic-MRCC theory [P7]. A strategy for allowing larger active spaces is the theory's extension to the use of incomplete (e.g., Ref. [86]) or restricted [298] active spaces. Other routes towards the description of larger systems that could be combined with ic-MRCC theory are local correlation approaches [287–291] and the treatment of different parts of a system by different electronic structure methods through embedding schemes [292].

In view of these future prospects of ic-MRCC theory, the present work is the foundation for a vivid development of powerful tools for the description of chemically interesting and theoretically demanding systems.

A Appendix

A.1 Implementation details

A.1.1 Technical setup

Most of the programming effort for this thesis was done in the General Contraction Code (GeCCo, Ref. [44]), using the Fortran 90 programming language. GeCCo is interfaced to a modified version of the DALTON program package [293], which carries out the HF or MCSCF calculations and provides the matrix elements of the Hamiltonian in the basis of molecular orbitals. A new interface to the MCSCF code of the GAMESS program [139] was created in order to allow ic-MRCC calculations with state-averaged CASSCF orbitals. The GAMESS interface was also used for some of the larger applications in this work.

A.1.2 GeCCo in a nutshell

The most essential features of GeCCo are a symbolic algebra that allows the automated derivation of second-quantized expressions [226] and an algorithm for the evaluation of these expressions, using string-based techniques [294–296]. Second-quantized operators and expressions (*formulas*) are treated in the code as objects, and GeCCo provides a multitude of operations (*rules*) for their definition, manipulation and evaluation. However, a GeCCo routine implementing a particular electronic structure theory is typically not a mere sequence of such operations. Instead, rules performing certain tasks are grouped into *targets*, lending the routine a modular structure. The individual targets are linked by a network of *dependencies*. Once the execution of a primary target is requested due to the user's input, the program evaluates the dependency tree for this target to control the order in which to carry out the involved targets. Figure A1 illustrates the dependency tree for the evaluation of a single expression, with each target depicted as a blue box.

We may divide the rules into three classes, namely (i) those operating on the algebraic level, (ii) those creating an interface between the algebraic and numeric levels of the program, and (iii) rules performing numeric operations.
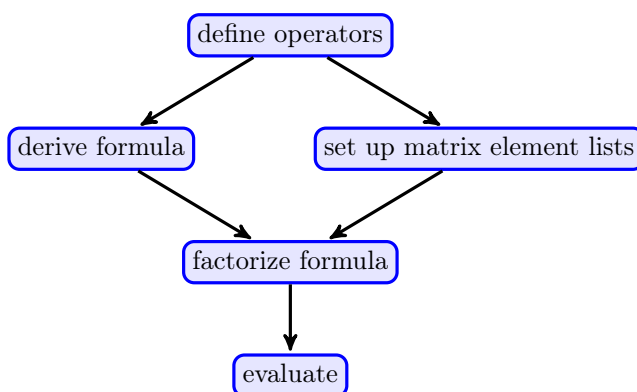
Figure A1: Dependency tree for the evaluation of a formula in GeCCo.

The first class subsumes the definition of operators, the derivation of quasi-vacuum expectation values of operator products (e.g., the method-specific Lagrangian), the differentiation of expressions with respect to Lagrange multipliers or other parameters, rules for factoring out or expanding subexpressions, and more specific operations such as the elimination of terms based on certain (e.g., perturbative) criteria. GeCCo's operator algebra is equivalent to a diagrammatic algebra that is based on antisymmetrized Brandow-type diagrams [119]. Each excitation type (*block*) of an operator is characterized by the number of creation (C) and annihilation (A) operators in the hole, particle, and valence orbital spaces, as exemplified in Fig. A2. In operator blocks consisting of an antisymmetric tensor and an excitation operator, e.g., $t_{ab}^{ij}\hat{a}_{ij}^{ab}$, all annihilators or creators belonging to the same orbital space are equivalent and need not be distinguished. In special cases, however, a tensor may not be antisymmetric with respect to the exchange of some indices of the same type. For instance, this happens for elements of the metric matrix in ic-MRCC theory such as $S_{va,wj}^{ui,xb}$ (see Tab. 3). In order to keep track of inequivalent indices of the same type, GeCCo allows a partitioning of the indices into different *vertices*. In a diagrammatic representation of an operator block, vertices appear as horizontal lines. The introduction of multiple vertices is also advantageous for preserving the topological structure of diagrams when defining expressions for operators through a differentiation. An example of an operator with two vertices is the residual in ic-MRCC theory, see Fig. A2b. It is obtained according to $\Omega_\rho = \partial\mathcal{L}/\partial\lambda^\rho$ and may inherit active annihilation lines from $\hat{C}_0^\dagger$ at the top of a diagram in addition to the lines originating from connections between Λ and the operators on its right. In the derivation of quasi-vacuum expectation values, GeCCo employs Wick's theorem and finds all possible ways to distribute binary contractions in the considered operator product. The resulting formulas are lists of individual terms (or diagrams).

In order to allow an efficient evaluation of expressions, a rule of type (ii) converts them into lists of binary tensor contractions and—if needed—index reordering steps. This factorization is often preceded by a recognition of predefined intermediates. For each term in the final expression, GeCCo systematically runs through the various possible contraction sequences and determines the one with the least computational cost. The estimation of the computational cost for a contraction explicitly takes into account the dimension of the contracted tensors and thus depends on the sizes of the orbital spaces. To this end, objects called *matrix element lists* are defined, which contain the information necessary for creating an interface between an operator and a corresponding tensor stored on hard disk. GeCCo uses graphical techniques [297, 298] for addressing and storing the

(a)

	n_h	n_p	n_v
C	0	1	1
A	0	0	2

(b)

vertex		n_h	n_p	n_v
1	C	0	0	0
	A	0	0	2
2	C	0	1	1
	A	0	0	0
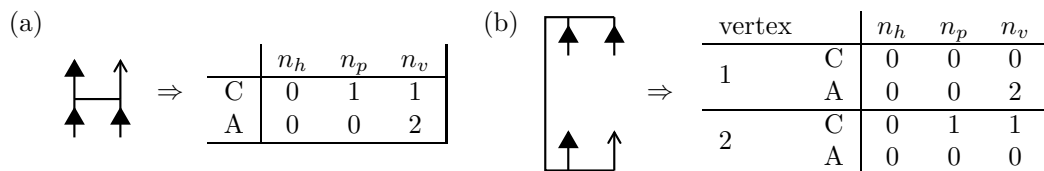
Figure A2: Representation of (a) the cluster excitation $t_{wa}^{uv}\hat{a}_{uv}^{wa}$ and (b) the corresponding residual in GeCCo.

nonredundant elements of antisymmetrized tensors with arbitrary rank. The irreducible representations of the orbitals are exploited along with the total symmetry of a tensor and its M_s quantum number, significantly reducing the number of tensor elements for Abelian point groups.

Finally, operations of type (iii) change the content of the matrix element lists. While some rules perform individual tasks like the evaluation of a formula or, for instance, the singular value decomposition of the metric matrix needed in ic-MRCC theory, others trigger the iterative solution of systems of equations. GeCCo employs subspace-based Davidson-type algorithms for eigenvalue problems and linear equation systems [299]. In the solution of nonlinear equations, the convergence is enhanced by Pulay's direct inversion in the iterative subspace (DIIS) method [300, 301].

A.1.3 Implementation of uncontracted MRCI

The implementation of (uncontracted) MRCI methods in GeCCo is based on the energy expectation value

$$E = \langle 0|\hat{C}_0^\dagger \hat{H}\hat{C}_0|0\rangle, \tag{A.1}$$

where

$$\hat{C}_0 = \hat{\boldsymbol{v}}\mathbf{c} \tag{A.2}$$

is a generalization of Eq. (3.11). The operators in the row vector $\hat{\boldsymbol{v}}$ combine the creation of N_{act} electrons with excitations of electrons out of the hole space and into the particle space, see Fig. A3 for an example. When $\hat{C}_0$ is restricted to the excitation class with $n_h = 0$, $n_p = 0$, the MRCI method reduces to CAS-CI and is used in GeCCo to compute the reference coefficients c_μ before an ic-MRCC calculation.

The coefficients in $\mathbf{c}$ are determined by the eigenvalue equation

$$\mathbf{Hc} = E\mathbf{c}, \tag{A.3}$$

with

$$\mathbf{H} = \langle 0|\hat{\boldsymbol{v}}^\dagger \hat{H}\hat{\boldsymbol{v}}|0\rangle. \tag{A.4}$$
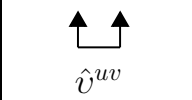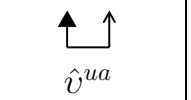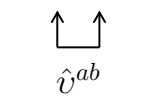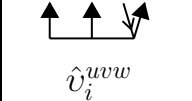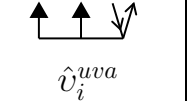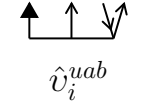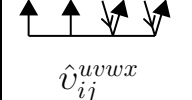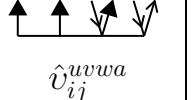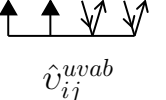
n_h	$n_p = 0$	$n_p = 1$	$n_p = 2$
0	$\hat{v}^{uv}$	$\hat{v}^{ua}$	$\hat{v}^{ab}$
1	$\hat{v}_i^{uvw}$	$\hat{v}_i^{uva}$	$\hat{v}_i^{uab}$
2	$\hat{v}_{ij}^{uvwx}$	$\hat{v}_{ij}^{uvwa}$	$\hat{v}_{ij}^{uvab}$

Figure A3: Blocks of $\hat{C}_0$ for (uncontracted) MRCISD with a CAS(2,2).

Davidson's algorithm [299] requires the evaluation of the matrix vector product $\mathbf{Hc}$ in each iteration. The formula for this step is conveniently derived in GeCCo by differentiating Eq. (A.1) with respect to the parameters in $\hat{C}_0^\dagger$.[32] Further, the diagonal elements of $\mathbf{H}$ are needed for approximating $(\mathbf{H} - E \cdot \mathbf{1})^{-1}$ in the generation of new subspace vectors. These diagonal elements are obtained by a summation of diagonal elements of the Hamiltonian.

In the current spin orbital formalism, control over the spin quantum number S of the state $\hat{C}_0|0\rangle$ requires additional effort. For states with $M_s = 0$, GeCCo can clearly distinguish between states with even and odd S by enforcing the correct spin-flip symmetry in $\hat{C}_0$. Whenever spin-flip symmetry is not sufficient to ensure the desired spin multiplicity, unwanted spin multiplicities are projected out according to

$$^{2S+1}\hat{C}_0|0\rangle = \ ^{2S+1}\hat{O}\hat{C}_0|0\rangle, \tag{A.5}$$

using Löwdin's projector [302]

$$^{2S+1}\hat{O} = \prod_{K \neq S} \frac{\hat{S}^2 - K(K+1)}{S(S+1) - K(K+1)}. \tag{A.6}$$

In second quantization, the square of the spin operator reads [3]

$$\hat{S}^2 = \tfrac{1}{2}\big(\hat{S}_+\hat{S}_- + \hat{S}_-\hat{S}_+\big) + \hat{S}_z^2, \tag{A.7}$$

where

$$\hat{S}_z = \sum_p m_s(p)\hat{a}_p^p \tag{A.8}$$

probes the M_s quantum number and the ladder operators

$$\hat{S}_\pm = \sum_{\phi_p = \phi_q} \big(m_s(q) \pm \tfrac{1}{2}\big)\big(m_s(q) - m_s(p)\big)\hat{a}_p^q \tag{A.9}$$

excite from a spin orbital with $m_s(p) = -\tfrac{1}{2}$ (spin function β) to the corresponding orbital with $m_s(q) = \tfrac{1}{2}$ (spin function α) and vice versa.

A.1.4   Implementation of ic-MRCC theory

The derivation of the ic-MRCC equations in GeCCo starts from the Lagrange function

$$\mathcal{L} = \langle 0|\hat{C}_0^\dagger(1 + \hat{\Lambda})e^{-\hat{T}}\hat{H}e^{\hat{T}}\hat{C}_0|0\rangle. \tag{A.10}$$

In contrast to Eq. (5.5), Eq. (A.10) does not explicitly contain the Lagrange multipliers $\bar{c}_\mu$. Nevertheless, the derivation of the left-hand side of the eigenvalue equations (3.5) is straightforward. It proceeds by deleting the terms that contain $\hat{\Lambda}$ and differentiating with respect to the parameters in $\hat{C}_0^\dagger$. The last term in Eq. (5.5) is also not needed as

---

[32]Note that GeCCo's operator algebra treats $\hat{C}_0^\dagger$ and $\hat{C}_0$ independently. Differentiations with respect to the parameters in either $\hat{C}_0^\dagger$ or $\hat{C}_0$ are therefore different from a mathematically rigorous differentiation with respect to c_μ.

it only contributes to the trivial right-hand side of the eigenvalue equation. The BCH expansion is carried out up to a given commutator level N_{com} in form of the summation

$$e^{-\hat{T}}\hat{H}e^{\hat{T}} = \sum_{n=0}^{N_{\text{com}}} \sum_{k=0}^{n} \frac{(-1)^k}{k!(n-k)!}\hat{T}^k\hat{H}\hat{T}^{n-k}. \tag{A.11}$$

Initially this leads to the multiple occurrence of topologically identical terms. A subsequent summation of identical terms ensures that each individual term appears only once and with its correct prefactor.[33] Since disconnected terms aquire a prefactor of zero, they naturally cancel out in this step.

When using the sequential ansatz from Eq. (5.41), the formula for the $\hat{T}_1$-transformed Hamiltonian $\tilde{H}$ is defined separately, also with help of Eq. (A.11). In the Lagrangian from Eq. (A.10), $\hat{H}$ is then replaced by $\tilde{H}$, and $\hat{T}$ by a cluster operator without singles. In order to circumvent a costly evaluation of the blocks of $\tilde{H}$ with three or four particle indices, the explicit expressions for these blocks are inserted, leading to a few additional terms containing $\hat{H}$ and $\hat{T}_1$.

The perturbative truncation of the ic-MRCCSDT Lagrangian closely follows the recipe detailed in sec. 5.4.2. Instead of explicitly defining the Dyall Hamiltonian, Eq. (5.6), the automated assignment of perturbation orders to the operators in a term takes into account whether the Hamiltonian block corresponds to one of the blocks in the Dyall Hamiltonian. In the event that a fourth-order term contains either $f_j^i\hat{a}_i^j$ or $f_b^a\hat{a}_a^b$, the Fock operator is replaced by the effective Fock operator defined in Eq. (4.1). Since $\hat{F}_{\text{eff}}$ depends on $\hat{C}_0$, it has to be evaluated after each solution of the ic-MRCC eigenvalue equations. The current procedure ensures this by inserting its explicit expression into the Lagrangian. In methods based on the sequential ansatz, one cannot directly classify the blocks of $\tilde{H}$ according to perturbation orders. It is therefore expanded in terms of $\hat{H}$ and $\hat{T}_1$ prior to the perturbative truncation and factored out again afterwards. The derivation of the ic-MRCCSD(T) Lagrangian additionally involves the removal of the terms constituting the ic-MRCCSD residuals as well as a substitution of $\hat{T}_1^\dagger$ and $\hat{T}_2^\dagger$ for $\hat{\Lambda}_1$ and $\hat{\Lambda}_2$, respectively. Further, the terms contributing to the ic-MRCCSD energy are exempt from the perturbative truncation, as they should be retained independent of the perturbation order assigned to $\hat{T}_1$. The (T) step currently requires a separate calculation and reuses matrix element lists for $\hat{C}_0$ and $\hat{T}$ from a previous ic-MRCCSD run.

The formula for the ic-MRCC residual follows from a differentiation of the Lagrangian with respect to $\boldsymbol{\lambda}$. In order to reduce the number of terms scaling highly with respect to the number of inactive orbitals, the resulting formula is differentiated separately with respect to the Hamiltonian vertices $g_{ab}^{AB}\hat{a}_{AB}^{ab}$ and $g_{IJ}^{ij}\hat{a}_{ij}^{IJ}$. This yields expressions for intermediates of the type $\mathcal{I}_{AB}^{IJ}$, which are generalizations of the intermediates discussed in sec. 3.1.3. These intermediates are then factored out prior to the factorization of the formula for the residual.

---

[33]The identification of topologically identical terms relies on a comparison of diagrams after bringing their operators into a canonical order. However, I was not able to find a canonicalization recipe that provides a unique order for all diagrams with four or more $\hat{T}$ operators. In order to obtain the correct number of terms even for high commutator levels, the termwise comparison was extended to consider all possible permutations of operators.

The metric matrix can be derived as a second derivative of the expression

$$\langle 0|\hat{C}_0^\dagger \hat{T}^\dagger \hat{T} \hat{C}_0|0\rangle = \langle 0|\hat{T}^\dagger \hat{T}|0\rangle + \sum_n \gamma_{u_1\cdots u_n}^{v_1\cdots v_n} \langle 0|\left\{\hat{a}_{u_n}\cdots\hat{a}_{u_1}\right\}\hat{T}^\dagger\hat{T}\{\hat{a}_{v_1}^\dagger\cdots\hat{a}_{v_n}^\dagger\}|0\rangle \quad \text{(A.12)}$$

with respect to both $\mathbf{t}^\dagger$ and $\mathbf{t}$. In Eq. (A.12), the expression $\left\{\hat{A}\right\}\hat{B}\{\hat{C}\}$ shall indicate that $\hat{A}\hat{C}$ is a normal ordered product that can be contracted to the operator $\hat{B}$ in between. The braces surrounding $\hat{B}$, printed in a smaller font than the ones embracing $\hat{A}$ and $\hat{C}$, locally lift the normal ordering. In the current example, $\hat{A}$ and $\hat{C}$ are the two vertices of a reduced density matrix (cf. Fig. 4a). Hence, the implementation of the metric matrix through the right-hand side of Eq. (A.12) directly leads to expressions in terms of reduced density matrices. As the operators $\hat{T}^\dagger$ and $\hat{T}$ are contracted by at least one inactive line, a mathematically correct second derivative of Eq. (A.12) leads to the appearance of unit tensors (cf. Tab. 3). Within GeCCo's algebra these unit tensors have to be explicitly inserted before the differentiation, as the lines connecting $\hat{T}^\dagger$ and $\hat{T}$ will otherwise simply disappear and leave the resulting operator with a reduced rank. This feature can be deliberately exploited for a removal of all inactive lines in the metric matrix. In this way, insertion of only the active unit tensors into contractions between $\hat{T}^\dagger$ and $\hat{T}$ followed by double differentiation straightforwardly yields the formula for the active parts of the metric matrix. Figure A4a shows a diagrammatic representation of the active metric block that corresponds to the excitation $\hat{a}_{uv}^{wa}$ and how the indices coming from the excitation and deexcitation operator can be merged to the superindices σ and ρ, in order to bring the metric tensor into matrix form.

The singular value decomposition is carried out individually for all possible subblocks of $\mathbf{S}_{\text{act}}^{n_h,n_p}$ with a given spatial symmetry and M_s quantum number for both indices σ and ρ. A subblock may have contributions from all excitation ranks within one excitation class. It is either diagonalized fully or sequentially according to the procedure detailed in sec. 3.2.1. In case of subblocks with $M_s = 0$, we can achieve a block diagonal structure by transforming into a basis of positive and negative linear combinations of excitations related by spin-flip. The singular value decomposition is then performed on the two resulting blocks. This prevents a mixing of singlet- and triplet-adapted excitations with the same singular value, which would lead to spin contamination in the cluster operator.

Initially, the transformation matrix $\mathbf{X}$, defined in Eq. (3.31), is stored on a matrix element list of an operator with the same topological structure as $\mathbf{S}$, while the matrix
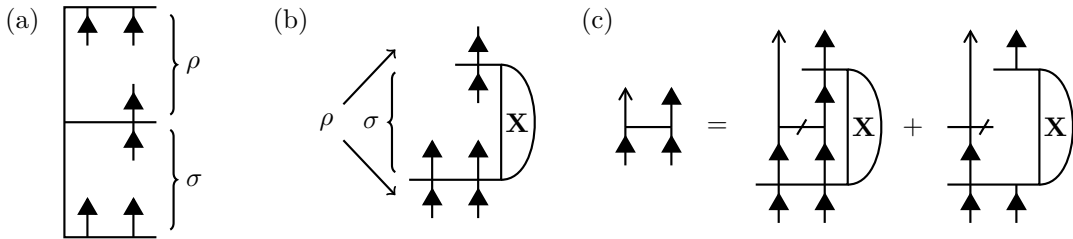
Figure A4: Sample block of (a) the active metric matrix and (b) the reordered transformation matrix, and (c) the corresponding formula for the transformation of the cluster amplitudes.

element list of $\mathbf{S}$ is overwritten with the projector $\mathbf{P}$ from Eq. (3.34). Both tensors are subsequently reordered by moving the active annihilators from the first vertex to the last vertex, as exemplified in Fig. A4b. This allows the definition of a simple formula for the transformation

$$\mathbf{t} = \mathbf{X}\mathbf{t}' \tag{A.13}$$

that preserves the topological structure of the cluster operator, see Fig. A4c for an example. Assigning the transformation operator to the matrix element list of $\mathbf{P}$, the same formula is used for carrying out the projection from Eq. (3.33). The transformation of the residual into the orthogonal excitation basis according to

$$\mathbf{\Omega}' = \mathbf{X}^\dagger \mathbf{\Omega} \tag{A.14}$$

also employs the same formula. To this end, a separate matrix element list for $\mathbf{X}^\dagger$ is created by reordering the adjoint of the initial transformation matrix. Owing to GeCCo's way to order index strings in the addressing of tensor elements, matrix element lists belonging to cluster amplitudes and residuals are isomorphic. This allows the direct application of the transformation formula to the residual.

The GNO-based removal of redundant excitation operators described in sec. 3.2.2 requires the following additional steps: The derivation of the metric matrix $\tilde{\mathbf{S}}$ from Eq. (3.66) also starts from the right-hand side of Eq. (A.12). The reduced density matrices are expanded in terms of cumulants according to Eq. (6.1). The removal of terms involving self-contractions (i.e., terms in which a cumulant tensor is only contracted to either $\hat{T}^\dagger$ or $\hat{T}$) yields the expression for $\langle\Psi_0|\hat{T}^\dagger\hat{T}|\Psi_0\rangle$ as evaluated with the extended Wick theorem. The current implementation of the GNO-based scheme is limited to ic-MRCCSD and a full diagonalization of the metric matrix,

$$\tilde{\mathbf{s}} = \tilde{\mathbf{U}}^\dagger \tilde{\mathbf{S}}\tilde{\mathbf{U}}. \tag{A.15}$$

As the cluster operator and the residual are still expressed with respect to $|0\rangle$ as Fermi vacuum, the additional transformation between the two normal orderings according to Eqs. (3.60) and (3.61) is taken into account by modifying the transformation matrix $\tilde{\mathbf{X}}$ and the projector $\tilde{\mathbf{P}}$ as follows:

$$\mathbf{X} = \mathbf{Y}\tilde{\mathbf{X}} = \begin{pmatrix} \tilde{\mathbf{X}}_1 + \mathbf{Y}_{12}\tilde{\mathbf{X}}_{21} & \tilde{\mathbf{X}}_{12} + \mathbf{Y}_{12}\tilde{\mathbf{X}}_2 \\ \tilde{\mathbf{X}}_{21} & \tilde{\mathbf{X}}_2 \end{pmatrix}, \tag{A.16}$$

$$\mathbf{P} = \mathbf{Y}\tilde{\mathbf{P}}\mathbf{Y}^{-1} = \begin{pmatrix} \tilde{\mathbf{P}}_1 + \mathbf{Y}_{12}\tilde{\mathbf{P}}_{21} & \tilde{\mathbf{P}}_{12} + \mathbf{Y}_{12}\tilde{\mathbf{P}}_2 - \tilde{\mathbf{P}}_1\mathbf{Y}_{12} - \mathbf{Y}_{12}\tilde{\mathbf{P}}_{21}\mathbf{Y}_{12} \\ \tilde{\mathbf{P}}_{21} & \tilde{\mathbf{P}}_2 - \tilde{\mathbf{P}}_{21}\mathbf{Y}_{12} \end{pmatrix}. \tag{A.17}$$

Equations (3.57) to (3.59) imply that $\mathbf{Y}_{12}$ is essentially given by the negative one-particle density matrix,

$$\mathbf{Y}_{12}^{1,1} = -\boldsymbol{\gamma}_1, \qquad \mathbf{Y}_{12}^{1,0} = -\boldsymbol{\delta}\otimes_{\mathrm{A}}\boldsymbol{\gamma}_1, \qquad \mathbf{Y}_{12}^{0,1} = -\boldsymbol{\delta}\otimes_{\mathrm{A}}\boldsymbol{\gamma}_1. \tag{A.18}$$

Figure A5a illustrates how the transformation with $\mathbf{Y}$ from the left is applied to both the initial tensors $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{P}}$. The additional transformation of $\tilde{\mathbf{P}}$ with $\mathbf{Y}^{-1}$ from the right is exemplified in Fig. A5b.

The evaluation of the *preconditioner*, a vector containing the diagonal elements of the approximate Jacobian $\mathbf{A}$ defined in Eq. (5.8), is performed in two steps: First, the
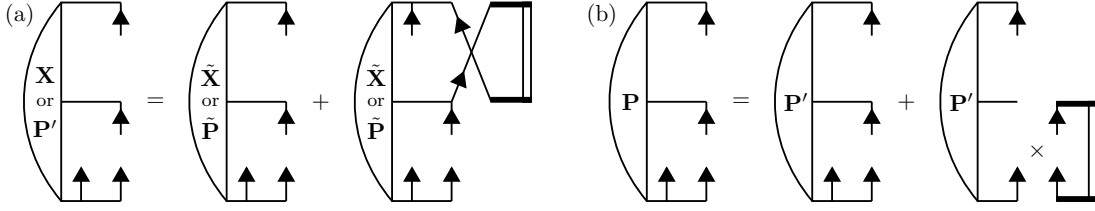
Figure A5: Sample blocks of the modified (a) transformation matrix and (b) projector when using the generalized normal ordering.

inactive part of the effective Fock operator from Eq. (4.1) is computed and its diagonal elements are added to the preconditioner according to Eq. (5.10) and the first summand in Eq. (5.9). However, the restriction of the unit tensors $\mathbf{1}_{\mathrm{act},\eta}^{n_h,n_p}$ to nonredundant excitations is dropped. This allows us to compute the preconditioner only once, as it already contains nonzero elements for nearly linearly dependent excitations that might enter the calculation after a couple of iterations. Second, the diagonal blocks of the matrices $\mathbf{A}_{\mathrm{act}}^{n_h,n_p}$ defined in Eq. (5.11) are explicitly evaluated before adding their diagonal elements as well. In GeCCo, the operator for $\mathbf{A}_{\mathrm{act}}$ has the same topological structure as the one for the active part of the metric. The formula for $\mathbf{A}_{\mathrm{act}}$ is derived starting from

$$\langle 0|\hat{C}_0^\dagger\hat{\Lambda}[\hat{H},\hat{T}]\hat{C}_0|0\rangle. \tag{A.19}$$

The insertion of the transformation formulas for $\hat{\Lambda}$ and $\hat{T}$ (see Fig. A4c for an example of the latter) and double differentiation with respect to both $\boldsymbol{\lambda}'$ and $\mathbf{t}'$ directly yields the desired expression.

The algorithm for the iterative solution of the ic-MRCC equations works as illustrated in Fig. A6. Starting from a cluster operator set to zero and the CASSCF wave function as a reference, the residual $\boldsymbol{\Omega}$ is computed. Although not strictly necessary, the current ic-MRCC energy is also evaluated at this point, see Eq. (3.3). After evaluating and diagonalizing the metric matrix, the residual is transformed to the orthogonal excitation basis with help of the matrix $\mathbf{X}^\dagger$. In this step, the residual's matrix element list is assigned to an operator with the same structure as $\hat{T}$. This implicit reordering of the residual operator goes along with an additional sign change for some blocks, which has to be explicitly taken care of. According to my experience, it is beneficial for the convergence rate to perform the amplitude update in the orthogonalized basis, using the diagonal elements of $\mathbf{A}$ as a denominator. If we estimate the impact of a change in the cluster amplitudes on the residual by considering only terms linear in $\hat{T}$, this choice corresponds to approximating the Hamiltonian by the zeroth-order Hamiltonian $\hat{H}^{(0)}$ defined in Eq. (5.14). The condition

$$0 \overset{!}{=} \Omega_\rho' + \Delta\Omega_\rho' \approx \Omega_\rho' + A_\rho^\rho\Delta t_\rho' \tag{A.20}$$

then immediately leads to the equation for the amplitude update shown in Fig. A6. Applying the DIIS method [300] for an enhanced convergence, both the amplitude guess $\mathbf{t}^{[n]} = \mathbf{t} + \Delta\mathbf{t}^{[n]}$ in the n-th iteration and the corresponding increment $\Delta\mathbf{t}^{[n]}$ are stored as subspace vectors. The amplitude vectors are then extrapolated by the linear combination $\mathbf{t}_{\mathrm{extr}} = \sum_n w_n\mathbf{t}^{[n]}$, with coefficients w_n chosen such that they minimize the norm
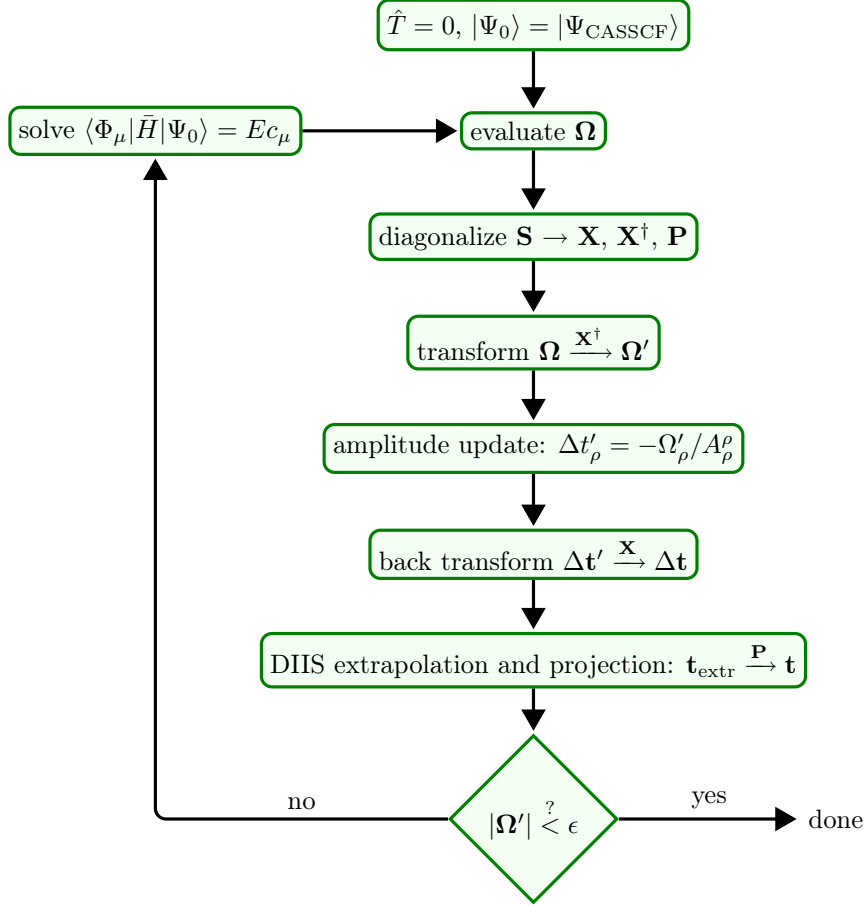
Figure A6: Iterative solution of the ic-MRCC equations.

of the extrapolated error vector $\Delta \mathbf{t}_{\text{extr}} = \sum_n w_n \Delta \mathbf{t}^{[n]}$. In the calculations performed in this work, the subspace contained vectors from up to the last eight iterations. As the amplitude vectors from different iterations belong to different metric matrices, the redundant components of $\mathbf{t}_{\text{extr}}$ with respect to the current metric are removed by applying the projector $\mathbf{P}$. Note that this step is not needed when using a fixed reference function, because the transformation matrices satisfy $\mathbf{X}^\dagger = \mathbf{X}^\dagger \mathbf{P}$ and $\mathbf{X} = \mathbf{P}\mathbf{X}$ and thus already account for the projection. The iterative cycle is continued as long as the norm of the residual in the orthogonal excitation basis still exceeds a convergence threshold ϵ. Before evaluating the next residual, the reference function is updated by solving the eigenvalue equation (3.5). This is done as described in app. A.1.3, with the difference that the matrix vector product $\mathbf{Hc}$ contains the similarity transformed Hamiltonian instead of $\hat{H}$. Upon convergence, the remaining error in the ic-MRCC energy in E_h is usually less than the threshold ϵ. I have also implemented an alternative algorithm in which each iteration involves only a single update step for the reference function, carried out simultaneously with the amplitude update according to

$$\Delta c_\mu = -\frac{\langle \Phi_\mu | (\bar{H} - E) | \Psi_0 \rangle}{\langle \Phi_\mu | \hat{H} | \Phi_\mu \rangle - E},$$
(A.21)

followed by a normalization of the reference function. However, this scheme was found to provide convergence only for CAS(2,2) cases.

The correctness of the implementation of ic-MRCC theory relies to a great extent on the reliability of the automated tools of the GeCCo program, which is supported by the successful implementation of various advanced electronic structure methods such as SRCC methods with explicitly correlated techniques (e.g., Ref. [44]). The implementation of multireference techniques started with the implementation of uncontracted MRCI and an internally contracted variant thereof, allowing direct comparisons with literature values at each stage. In order to verify the correctness of the implementation of ic-MRCCSD, it was compared with Evangelista's FCI-based pilot implementation [38] for a small example (BeH$_2$ model with the double-zeta basis set from Ref. [38] and using the full orthogonalization scheme), revealing an agreement of the total energies within 10^{-10} E_h [P5]. A number of internal checks used to verify individual parts of the implementation lend further credence to its correctness beyond this single example. These tests included random samples of terms in the ic-MRCC equations, checks of the idempotency of projectors and the fulfillment of $\mathbf{X}^{\dagger}\mathbf{S}\mathbf{X} = \mathbf{1}_\eta$ as well as a numerical verification of formal properties such as the orbital invariance of the theory and its core-extensivity (for the sequential orthogonalization) or full size-extensivity (for the GNO-based scheme) [P4].

A.2   Data corresponding to the figures

Table A1: Primary data for Fig. 12 (BeH$_2$ model system). FCI energies are reported in E_h. The results for all other methods are given as energy differences to FCI in mE_h.

z / a_0	FCI[a]	CASSCF	MRCISD[a]	CASCCSD[b]		ic-MRCCSD[c]	ic-MRCCSDT[c]
				(Φ_1)	(Φ_2)		
0	-15.8354756	62.2034	1.0819	0.4823	0.2970	0.3848	-0.0005
1	-15.8022622	64.4918	1.1675	0.4926	0.3163	0.4091	0.0025
2	-15.7366210	71.5254	1.4778	0.5561	0.5421	0.6719	0.0127
2.5	-15.6836446	76.8706	2.0498	1.2164	0.8837	1.0018	0.0197
2.625	-15.6696765	77.8854	2.2859	1.5002	1.0879	1.2288	0.0270
2.6875	-15.6633811	78.2010	2.4405	1.6404	1.2583	1.4163	0.0401
2.725	-15.6601303	78.3207	2.5635	1.7083	1.4045	1.5690	0.0557
2.75	-15.6583118	78.3736	2.6673	1.7388	1.5272	1.6905	0.0705
2.775	-15.6568705	78.4035	2.7944	1.7513	1.6714	1.8247	0.0893
2.8125	-15.6556164	78.3776	3.0329	1.7261	1.9146	2.0243	0.1234
2.85	-15.6556594	78.1739	3.3104	1.6484	2.1377	2.1546	0.1627
2.875	-15.6564315	77.8881	3.4916	1.5793	2.2386	2.1672	0.1881
2.9	-15.6577388	77.4932	3.6482	1.5081	2.2876	2.1162	0.2034
2.95	-15.6616025	76.5634	3.8517	1.3825	2.2502	1.8953	0.1817
3	-15.6665607	75.6615	3.9301	1.2866	2.1077	1.6579	0.1260
3.1	-15.6780097	74.0973	3.8985	1.1550	1.7575	1.3496	0.0570
3.25	-15.6962459	72.0535	3.6951	1.0186	1.3355	1.1301	0.0404
3.5	-15.7252893	68.9328	3.2990	0.8291	0.9112	0.9036	0.0377
4	-15.7605670	63.3380	2.6397	0.4909	0.4716	0.5228	0.0203

[a] Published in Ref. [P1].
[b] Computed with Lucia [138].
[c] Published in Ref. [P3].

Table A2: Primary data for Figs. 19a, 20a, 23a and 26a (BeH$_2$ model system, ic-MRCCSD), given as energy differences to FCI in mE$_h$.

$N_{com}(E) =$	1	4	4[a]	3	4	4	4	2	2
$N_{com}(\Omega) =$	1	1	2	3	4	8	2	2	2
$z\,/\,a_0$							GNO-based	Fixed reference	Internal excitations
0	−0.7975	−0.9628	0.3843	0.3873	0.3873	0.3873	0.3845	0.3846	0.3846
1	−0.8093	−0.9637	0.4073	0.4117	0.4118	0.4118	0.4095	0.4084	0.4092
2	−1.3493	−1.3907	0.6628	0.6809	0.6811	0.6811	0.6759	0.7068	0.6552
2.5	−1.6888	−3.0222	1.0007	1.0333	1.0329	1.0329	1.0089	1.0031	1.0022
2.625	−1.5526	−3.4790	1.2270	1.2657	1.2650	1.2650	1.2357	1.2306	1.2317
2.6875	−1.3299	−3.7727	1.4106	1.4502	1.4495	1.4495	1.4170	1.4161	1.4183
2.725	−1.0952	−4.0150	1.5572	1.5949	1.5943	1.5943	1.5601	1.5701	1.5660
2.75	−0.8756	−4.2275	1.6714	1.7060	1.7057	1.7057	1.6706	1.6962	1.6801
2.775	−0.5895	−4.4991	1.7941	1.8240	1.8241	1.8241	1.7887	1.8401	1.8025
2.8125	−0.0010	−5.0495	1.9649	1.9856	1.9868	1.9868	1.9521	2.0660	1.9767
2.85	0.8152	−5.7436	2.0505	2.0631	2.0662	2.0662	2.0343	2.2336	2.0808
2.875	1.4460	−6.1385	2.0312	2.0403	2.0448	2.0447	2.0163	2.2769	2.0818
2.9	1.9803	−6.2626	1.9596	1.9664	1.9717	1.9717	1.9470	2.2669	2.0233
2.95	1.8008	−5.4976	1.7539	1.7590	1.7639	1.7639	1.7446	2.1744	1.7946
3	0.3968	−4.4512	1.5671	1.5728	1.5758	1.5758	1.5599	2.0961	1.5602
3.1	−1.7466	−3.5714	1.3259	1.3316	1.3322	1.3322	1.3226	2.0560	1.2716
3.25	−2.8659	−3.2443	1.1288	1.1303	1.1303	1.1303	1.1271	2.0536	1.0679
3.5	−2.9825	−2.9048	0.9034	0.9007	0.9007	0.9007	0.9009	1.9160	0.8427
4	−2.5081	−2.4305	0.5227	0.5201	0.5201	0.5201	0.5213	1.4725	0.4608

[a] Published in Ref. [P1]. See this Ref. also for data on the 'highest rank only' and 'full orthogonalization' schemes.

Table A3: Primary data for Fig. 13 (dissociation of HF). FCI energies are reported in E_h. The results for all other methods are given as energy differences to FCI in mE_h.

$R_{\text{H–F}}$ / a_0	FCI[a]	CASSCF	MRCI		Mk-MRCC[b]		ic-MRCC[d]	
			SD[b]	SDTQ[c]	SD	SDT	SD	SDT
1	−99.555490	88.647	2.776	0.062	1.254	−0.004	1.169	−0.083
1.25	−99.972487	93.637	2.958	0.058	1.279	0.092	1.146	−0.065
1.5	−100.113242	97.390	3.118	0.055	1.372	0.194	1.147	−0.057
1.733	−100.146457	99.730	3.241	0.053	1.505	0.288	1.153	−0.056
1.75	−100.146980	99.855	3.249	0.053	1.516	0.295	1.154	−0.056
2	−100.138922	100.966	3.347	0.054	1.682	0.391	1.155	−0.057
2.25	−100.115746	100.746	3.412	0.057	1.836	0.484	1.151	−0.058
2.5	−100.088759	99.327	3.433	0.061	1.955	0.567	1.146	−0.061
2.75	−100.062848	96.943	3.405	0.065	2.028	0.630	1.147	−0.064
3	−100.040049	93.912	3.327	0.070	2.053	0.653	1.158	−0.069
3.466	−100.007886	87.825	3.085	0.076	1.954	0.463	1.196	−0.083
3.5	−100.006069	87.410	3.066	0.077	1.937	0.433	1.200	−0.083
4	−99.986564	82.448	2.796	0.081	1.404	−0.234	1.229	−0.094
5	−99.973063	78.570	2.561	0.085	0.146	−0.473	1.226	−0.096
5.199	−99.972218	78.325	2.546	0.085	0.251	−0.364	1.223	−0.097
6	−99.970733	77.905	2.524	0.086	0.889	−0.152	1.216	−0.095
8.665	−99.970266	77.792	2.519	0.086	1.211	−0.105	1.213	−0.094
10	−99.970264	77.792	2.519	0.086	1.211	−0.105	1.213	−0.094

[a] From Ref. [140].
[b] From Ref. [141].
[c] Published in Ref. [P1].
[d] Close to results published in Ref. [P1], which use up to fourfold commutator terms in the eigensystem.

Table A4: Primary data for Fig. 14 (dissociation of LiF). FCI energies are reported in E_h. The results for all other methods are given as energy differences to FCI in mE_h. Values in parentheses indicate that small metric eigenvalues (below $\eta = 10^{-4}$) were excluded.

$R_{\mathrm{Li-F}}$ / a_0	FCI[a]	SA-CASSCF	MRCISD[a]	ic-MRCCSD[b]	
				Sequential orth.	GNO-based
State X $^1\Sigma^+$					
2.5	−107.0960120	200.3848	10.2841	2.4595	2.7971
3	−107.1305303	202.7691	11.0058	2.3541	3.2900
3.5	−107.1190210	204.5679	11.6251	2.4196	3.5533
4	−107.0967924	205.8001	12.1408	2.5925	3.7522
6	−107.0175557	203.5872	12.9438	2.8080	4.1418
8	−106.9718998	193.5350	12.8488	2.6331	4.1714
10	−106.9451020	172.7575	12.7636	2.4743	4.1304
10.4	−106.9410635	168.8929	12.7444	2.4529	4.1266
10.8	−106.9373443	165.2732	12.7209	2.4365	4.1269
11.2	−106.9339082	161.8932	12.6891	2.4265	4.1342
11.6	−106.9307246	158.7404	12.6394	2.4257	4.1542
12	−106.9277679	155.7998	12.5421	2.4405	4.2014
12.5	−106.9243614	152.4019	12.1054	2.5104	4.3745
13	−106.9212629	149.3052	10.2471	2.7992	4.7625
13.5	−106.9186499	146.6909	7.8445	3.7045	3.9103
14	−106.9178186	145.8568	7.0674	3.2299	2.8987
14.5	−106.9177341	145.7690	7.0025	2.5307	2.5004
15	−106.9177131	145.7446	6.9896	2.4550	2.4521
15.2	−106.9177091	145.7393	6.9875	2.4484	2.4473
15.6	−106.9177043	145.7320	6.9852	2.4429	2.4432
20	−106.9176971	145.7076	6.9812	2.4395	2.4404
30	−106.9176964	145.6971	6.9801	2.4394	2.4403
State B $^1\Sigma^+$					
2.5	−106.8152906	144.4456	6.6396	2.4778	2.4954
3	−106.8713073	145.1999	6.8221	2.5476	2.5608
3.5	−106.8888907	146.9605	7.1397	2.6826	2.6915
4	−106.8957828	148.8794	7.4138	2.7930	2.7970
6	−106.9082536	154.4629	7.5995	2.8059	2.7977
8	−106.9147983	163.2704	7.3625	2.6399	2.6345
10	−106.9169531	183.5275	7.2014	2.5339	2.5313
10.4	−106.9171270	187.3352	7.1883	2.5218	2.5199
10.8	−106.9172558	190.9052	7.1833	2.5123	2.5111
11.2	−106.9173501	194.2411	7.1898	2.5052	2.5048
11.6	−106.9174175	197.3549	7.2170	2.5005	2.5010
12	−106.9174630	200.2610	7.2944	2.4985	2.5003
12.5	−106.9174916	203.6217	7.7093	2.5009	2.5062
13	−106.9174727	206.6870	9.5490	2.5124	2.5309
13.5	−106.9172010	209.2752	11.9359	2.5182 (2.3228)	(2.3173)
14	−106.9153559	210.0876	12.6999	2.4618 (2.9628)	(2.9625)
14.5	−106.9129511	210.1576	12.7541	2.2589 (2.9908)	(2.9908)
15	−106.9106515	210.1679	12.7584	(2.9905)	(2.9905)
15.2	−106.9097706	210.1683	12.7576	(2.9899)	(2.9899)
15.6	−106.9080748	210.1673	12.7550	(2.9889)	(2.9889)
20	−106.8939171	210.1631	12.7432	(2.9873)	(2.9873)
30	−106.8772326	210.1740	12.7448	(2.9867)	(2.9867)

[a] Published in Ref. [P1].

[b] Note that the ic-MRCCSD data published in Ref. [P1] is based on a different scheme to remove redundancies between singles and doubles (the 'highest rank only' scheme).

Table A5: Primary data for Fig. 16 (symmetric dissociation of H_2O). FCI energies are reported in E_h. The results for all other methods are given as energy differences to FCI in mE_h. For additional results from the literature, see Ref. [41].

R_{O-H}/R_e	FCI[a]	CASSCF	ic-MRCCSD	
		CAS(4,4)	CAS(6,5)	CAS(4,4)[a]
1.0	−76.2388502	163.5374	1.3144	1.3026
1.2	−76.1802630	160.0714	1.2280	1.2749
1.4	−76.0990342	154.1005	1.0865	1.1867
1.6	−76.0290190	146.2722	0.9339	1.0539
1.8	−75.9781502	138.4840	0.8353	0.9496
2.0	−75.9455881	132.1731	0.7173	0.8295
2.2	−75.9272311	127.8834	0.6312	0.7310
2.4	−75.9179257	125.3881	0.5720	0.6604
2.6	−75.9134107	124.0672	0.5364	0.6162
2.8	−75.9111780	123.3829	0.5415	0.6583
3.0	−75.9100279	123.0206	0.5281	0.6696
3.2	−75.9094154	122.8218	0.5973	0.7703
3.4	−75.9090795	122.7088	0.6202	0.8197
3.6	−75.9088890	122.6433	0.6175	0.8156
3.8	−75.9087777	122.6051	0.7580	0.8481

[a] Published in Refs. [P3, P5].

Table A6: Primary data for Figs. 22a and 25a (symmetric dissociation of H_2O, ic-MRCCSD), given as energy differences to FCI in mE_h.

R_{O-H}/R_e	Sequential orthogonalization			Full orth.	GNO-based
	$\eta = 10^{-2}$	$\eta = 10^{-3}$	$\eta = 10^{-4}$	$\eta = 10^{-6}$	$\eta = 10^{-6}$
1.0	3.0704	1.3601	1.3323	1.3111	1.3158
1.2	1.8450	1.3758	1.3015	1.2843	1.2900
1.4	1.8296	1.3107	1.1944	1.1935	1.1974
1.6	1.9172	1.2881	1.0888	1.0577	1.0599
1.8	1.5249	1.2226	0.9953	0.9502	0.9514
2.0	1.2961	1.1021	0.8908	0.8291	0.8308
2.2	1.0884	1.0884	0.8219	0.7304	0.7324
2.4	1.0725	0.9654	0.8891	0.6597	0.6617
2.6	0.9561	0.8960	0.8960	0.6154	0.6173
2.8	0.9013	0.9013	0.8586	0.6574	0.6591
3.0	0.8752	0.8752	0.8382	0.6686	0.6703
3.2	0.8619	0.8619	0.8619	0.7690	0.7710
3.4	0.8546	0.8546	0.8546	0.8181	0.8202
3.6	0.8504	0.8504	0.8504	0.8140	0.8161
3.8	0.8480	0.8480	0.8480	0.8464	0.8485

Table A7: Primary data for Fig. 17 (N_2 triple bond breaking). FCI energies are reported in E_h. The results for all other methods are given as energy differences to FCI in mE_h. For additional results from the literature, see Refs. [42, 93, 141, 158, 160].

R_{N-N} / a_0	FCI[a]	CASSCF	MRCI[b]		ic-MRCCSD	
			SD	SDTQ	($\eta = 10^{-4}$)	($\eta = 10^{-6}$)[c]
2.05	-109.2748399	186.5437	6.5876	0.0844	1.6303	1.6460
2.1	-109.2781206	187.3475	6.6270	0.0838	1.6296	1.6479
2.15	-109.2775711	188.1260	6.6636	0.0831	1.6295	1.6537
2.3	-109.2593821	190.3372	6.7601	0.0811	1.5544	1.6406
2.5	-109.2136504	193.0263	6.8682	0.0787	1.5233	1.5329
2.75	-109.1470625	195.8566	6.9884	0.0770	1.5155	1.2097
3	-109.0861916	197.6788	7.1006	0.0772	1.4404	1.3112
3.25	-109.0381329	197.8324	7.1820	0.0792	1.3266	1.4068
3.5	-109.0045545	195.9710	7.1870	0.0817	1.4386	1.4464
3.75	-108.9837811	192.7618	7.1073	0.0835	1.3464	1.2829
4	-108.9721625	189.4140	6.9909	0.0845	1.2407	1.1284
4.5	-108.9627305	184.7144	6.8108	0.0856	1.2071	0.9200
5	-108.9597981	182.4456	6.7293	0.0867	1.4915	0.8925
5.5	-108.9586119	181.3766	6.6915	0.0874	1.6199	0.9530
6	-108.9580173	180.8423	6.6683	0.0876	1.5645	1.2157
7	-108.9574155	180.4223	6.6417	0.0874	1.5263	1.5263
8	-108.9571750	180.3118	6.6280	0.0824	1.5135	1.5135
9	-108.9571224	180.2905	6.6306	0.0870	1.5166	1.5166
10	-108.9571109	180.2823	6.6305	0.0874	1.5166	1.5166

[a] Computed with LUCIA [138].
[b] Published in Ref. [P1].
[c] Curve published in Ref. [P5].

Table A8: Primary data for Figs. 19b, 20b, 25b and 26b (N_2 triple bond breaking, ic-MRCCSD), given as energy differences to FCI in mE_h.

$N_{com}(E)$ =	1	4	3	2	2	2	2
$N_{com}(\Omega)$ =	1	2	3	2	2	2	2
$R_{\text{N-N}}$ / a_0	$\eta = 0.994 \cdot 10^{-6}$ ($\eta = 1.02 \cdot 10^{-6}$)			Full orth.	GNO-based	Fixed reference	Internal excitations
2.05	-0.4747 (-0.2989)	1.6425	1.6277	1.6914	1.7402	2.0695	1.5011
2.1	-0.8085	1.6464	1.6283	1.6905	1.7373	2.1159	1.4695
2.15	-0.7620	1.6543	1.6330	1.6903	1.7377	2.1414	1.4165
2.3	-1.3467	1.6455	1.6238	1.6766	1.7223	1.9520	1.2487
2.5	-1.6632	1.5125	1.5223	1.6318	1.6907	1.8603	0.9485
2.75	0.4983	1.0481	1.2253	1.5340	1.6298	1.8444	0.8937
3	-1.0201	1.1927	1.1861	1.3842	1.5016	1.9127	0.8431
3.25	0.0506	1.3469	1.4365	1.4449	1.5181	2.2704	1.0579
3.5	-3.8514	1.3693	1.3823	1.4318	1.4488	2.2492	1.0575
3.75	-4.2659	1.2219	1.2311	1.2684	1.2765	2.2173	0.9311
4	-3.8518	1.0894	1.1112	1.1219	1.1309	2.0117	0.8279
4.5	-1.2744	0.9187	0.9811	0.9258	0.9387	1.5098	0.8297
5	-0.9142	0.8930	0.9222	0.8953	0.8962	1.2159	0.8701
5.5	-0.5122	0.9534	0.9940	0.9490	0.9445	1.1037	0.9154
6	-0.9656	1.2158	1.2415	1.2116	1.2083	1.3745	1.3207
7	-1.2922	1.5263	1.5432	1.5260	1.5263	1.5293	1.5252
8	-1.3023	1.5135	1.5304	1.5132	1.5135	1.5138	1.5134
9	-1.2980	1.5166	1.5334	1.5163	1.5166	1.5166	1.5166
10	-1.2975	1.5166	1.5335	1.5163	1.5166	1.5166	1.5166

Table A9: Primary data for Fig. 21 (N_2 triple bond breaking, ic-MRCCSD): Number of eigenvalues between 10^{-x} and 10^{1-x}.

$R_{N-N}/a_0 \setminus x =$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	>15
2.05	6	266	248	230	403	179	272	112	32	20	8	0	0	0	0	0	133
2.1	6	266	255	251	377	185	264	120	44	8	0	0	0	0	0	13	120
2.15	6	266	255	280	348	221	244	96	52	8	0	0	0	0	0	7	126
2.3	9	263	286	331	308	299	172	88	12	0	0	8	0	0	0	2	131
2.5	9	278	313	419	273	300	138	38	8	0	0	0	0	0	0	9	124
2.75	17	298	350	436	419	180	52	24	0	0	0	0	0	0	0	4	129
3	17	430	329	441	387	128	44	0	0	0	0	0	0	0	0	0	133
3.25	17	491	330	570	244	72	20	8	0	16	8	0	0	0	0	6	127
3.5	13	543	406	482	240	72	20	0	0	0	0	0	0	0	0	0	133
3.75	13	570	411	478	208	80	16	0	0	0	0	0	0	0	0	0	133
4	12	577	439	446	170	98	34	0	0	0	0	0	0	0	0	0	133
4.5	4	618	578	115	191	154	84	32	0	0	0	0	0	0	0	6	127
5	4	660	552	0	178	136	150	72	24	0	0	0	0	0	0	0	133
5.5	4	681	531	0	12	192	142	118	72	24	0	0	0	0	0	1	132
6	4	684	528	0	0	48	156	140	136	64	16	0	0	0	0	0	133
7	4	684	528	0	0	0	0	136	86	122	120	72	24	0	0	2	131
8	4	684	528	0	0	0	0	0	0	160	43	141	88	92	36	0	133
9	4	684	528	0	0	0	0	0	0	0	0	136	27	97	132	112	189
10	4	684	528	0	0	0	0	0	0	0	0	0	0	2	158	99	434

Table A10: Primary data for Fig. 22b (N_2 triple bond breaking, ic-MRCCSD), given as energy differences to FCI in mE_h.

$\eta \setminus R_{N-N} =$	2.1 a_0	2.3 a_0	2.5 a_0	2.75 a_0	3 a_0
$1 \cdot 10^{-2}$	2.7269	2.9362	3.0107	3.9063	3.6452
$3 \cdot 10^{-3}$	2.2053	2.4503	2.3759	2.6185	2.3597
$1 \cdot 10^{-3}$	1.9321	1.9778	1.9795	1.7177	1.8356
$3 \cdot 10^{-4}$	1.8112	1.7019	1.6521	1.5392	1.5403
$1 \cdot 10^{-4}$	1.6296	1.5544	1.5233	1.5155	1.4404
$3 \cdot 10^{-5}$	1.5632	1.5501	1.5633	1.5473	1.4382
$1 \cdot 10^{-5}$	1.5669	1.6312	1.6115	1.5559	1.3632
$3 \cdot 10^{-6}$	1.6473	1.6448	1.6224	1.3942	1.3602
$1 \cdot 10^{-6}$	1.6479	1.6406	1.5329	1.2097	1.3112
$3 \cdot 10^{-7}$	1.4937	1.5346	0.5742	1.1447	1.3112
$1 \cdot 10^{-7}$	1.4611	1.4246	0.5742	1.1447	1.3112
$3 \cdot 10^{-8}$	1.3767	1.2124[a]	0.0059	1.1447	1.3112
$1 \cdot 10^{-8}$	1.3767	-1.6505[a]	0.0059	1.1447	1.3112
$3 \cdot 10^{-9}$	1.4296	-1.6505[a]	0.0059	1.1447	1.3112
$1 \cdot 10^{-9}$	-8.4653[a]	-1.6505[a]	0.0059	1.1447	1.3112

[a] Reference coefficients c_μ fixed at converged values from previous calculation (with larger η).

Table A11: Primary data for Fig. 18 (symmetric dissociation of linear Be$_3$). FCI energies are reported in E_h. The results for all other methods are given as energy differences to FCI in mE_h.

R_{Be-Be} / a_0	FCI[a]	SRCC			CASSCF		MRCISD		D-USS(2)-BW[b]	ic-MRCCSD[c]		
		SD	SD(T)	SDT	CAS(6,12)	CAS(6,6)$_a$	CAS(6,6)$_b$	CAS(6,12)	CAS(6,6)$_a$	CAS(6,6)$_a$	CAS(6,6)$_b$	CAS(6,12)
3.80	−43.8575911	15.7056	0.4295	1.3192	31.5882	5.7524	7.5158	0.4619	0.1291	2.6435	2.2481	0.0651
3.85	−43.8602722	15.5870	0.5587	1.2846	31.8736	5.8832	7.4905	0.4581	0.0982	2.6815	2.1906	0.0628
3.90	−43.8625125	15.4605	0.6936	1.2511	32.1526	6.0254	7.4749	0.4531	0.0685	2.7227	2.1399	0.0648
3.95	−43.8643522	15.3242	0.8330	1.2184	32.4231	6.1789	7.4691	0.4469	0.0422	2.7660	2.1012	0.0668
4.00	−43.8658292	15.1763	0.9756	1.1859	32.6824	6.3440	7.4728	0.4394	0.0192	2.8117	2.0664	0.0684
4.05	−43.8669790	15.0151	1.1199	1.1534	32.9262	6.5205	7.4857	0.4305	−0.0010	2.8585	2.0347	0.0710
4.10	−43.8678353	14.8393	1.2643	1.1207	33.1482	6.7084	7.5078	0.4198	−0.0167	2.9057	2.0094	0.0744
4.15	−43.8684299	14.6476	1.4072	1.0876	33.3385	6.9077	7.5388	0.4073	−0.0271	2.9530	1.9849	0.0758
4.20	−43.8687926	14.4390	1.5468	1.0542	33.4820	7.1179	7.5786	0.3927	−0.0334	2.9980	1.9621	0.0761
4.25	−43.8689516	14.2127	1.6812	1.0202	33.5565	7.3383	7.6267	0.3766	−0.0344	3.0410	1.9372	0.0760
4.30	−43.8689335	13.9683	1.8088	0.9860	33.5343	7.5679	7.6829	0.3613	−0.0305	3.0917	1.9158	0.0755
4.35	−43.8687632	13.7056	1.9277	0.9516	33.3944	7.8054	7.7468	0.3494	−0.0198	3.1268	1.8877	0.0748
4.40	−43.8684643	13.4251	2.0364	0.9172	33.1361	8.0490	7.8181	0.3414	−0.0047	3.1570	1.8674	0.0743
4.45	−43.8680584	13.1271	2.1333	0.8830	32.7730	8.2958	7.8960	0.3353	0.0164	3.1810	1.8468	0.0736
4.50	−43.8675661	12.8129	2.2174	0.8494	32.3226	8.5428	7.9800	0.3299	0.0421	3.1979	1.8261	0.0729
4.55	−43.8670059	12.4837	2.2876	0.8165	31.8000	8.7853	8.0692	0.3244	0.0719	3.2065	1.8044	0.0721
4.60	−43.8663951	12.1413	2.3435	0.7846	31.2181	9.0181	8.1627	0.3184	0.1051	3.1894	1.7818	0.0718
4.65	−43.8657492	11.7875	2.3846	0.7540	30.5882	9.2342	8.2596	0.3119	0.1412	3.1768	1.7586	0.0708
4.70	−43.8650821	11.4246	2.4112	0.7247	29.9199	9.4258	8.3588	0.3047	0.1781	3.1517	1.7321	0.0694
4.75	−43.8644061	11.0549	2.4236	0.6968	29.2219	9.5836	8.4590	0.2970	0.2161	3.1140	1.7054	0.0682
4.80	−43.8637319	10.6808	2.4226	0.6706	28.5020	9.6983	8.5593	0.2888	0.2539	3.0627	1.6773	0.0665
4.85	−43.8630685	10.3049	2.4092	0.6459	27.7672	9.7629	8.6583	0.2803	0.2895	2.9971	1.6454	0.0659
4.90	−43.8624234	9.9294	2.3845	0.6226	27.0237	9.7778	8.7549	0.2715	0.3224	2.9285	1.6123	0.0645
4.95	−43.8618026	9.5564	2.3498	0.6007	26.2771	9.7557	8.8480	0.2625	0.3496	2.8356	1.5768	0.0635
5.0	−43.8612110	9.1883	2.3068	0.5801	25.5327	9.7203	8.9369	0.2535	0.3740	2.7348	1.5405	0.0619
5.2	−43.8591874	7.7945	2.0772	0.5063	22.6510	9.6789	9.2339	0.2182	0.4154	2.3119	1.3877	0.0546
5.4	−43.8577355	6.5730	1.8107	0.4409	20.0395	9.6431	9.4228	0.1869	0.3575	1.9002	1.2086	0.0464
5.6	−43.8567804	5.5402	1.5542	0.3799	17.7633	9.6619	9.5101	0.1606	0.2144	1.7406	1.0351	0.0400
5.8	−43.8562003	4.6800	1.3266	0.3238	15.8211	9.4964	9.5205	0.1392	0.0223	1.4608	0.8789	0.0334
6.0	−43.8558796	3.9654	1.1314	0.2738	14.1825	9.2716	9.4811	0.1216	−0.1974	1.4036	0.7423	0.0266
6.2	−43.8557268	3.3698	0.9657	0.2308	12.8086	9.0474	9.4132	0.1073	−0.4212	1.2159	0.6255	0.0216
6.4	−43.8556754	2.8706	0.8253	0.1947	11.6605	8.8439	9.3316	0.0954	−0.6356	1.0353	0.5318	0.0168
6.6	−43.8556796	2.4502	0.7061	0.1647	10.7030	8.6667	9.2456	0.0855	−0.8314	0.8875	0.4648	0.0131
6.8	−43.8557090	2.0947	0.6046	0.1399	9.9049	8.5151	9.1610	0.0772	−1.0070	0.7500	0.4045	0.0091
7.0	−43.8557447	1.7931	0.5182	0.1195	9.2399	8.3870	9.0813	0.0702	−1.1603	0.6380	0.3526	0.0062

Table A11 — continued from previous page

| $R_{\text{Be-Be}}$ | FCI[a] | SRCC | | | CASSCF | MRCISD | | | D-USS(2)-BW[b] | ic-MRCCSD[c] | | |
| | | SD | SD(T) | SDT | CAS(6,12) | CAS(6,6)$_a$ | CAS(6,6)$_b$ | CAS(6,12) | CAS(6,6)$_a$ | CAS(6,6)$_a$ | CAS(6,6)$_b$ | CAS(6,12) |
/ a_0												
7.2	−43.8557753	1.5368	0.4443	0.1026	8.6850	8.2788	9.0083	0.0643	−1.2917	0.5445	0.3068	0.0032
7.4	−43.8557950	1.3186	0.3814	0.0885	8.2216	8.1876	8.9428	0.0592	−1.4030	0.4644	0.2611	0.0010
7.6	−43.8558014	1.1325	0.3275	0.0766	7.8337	8.1105	8.8848	0.0548	−1.4966	0.4131	0.2275	−0.0002
7.8	−43.8557946	0.9739	0.2816	0.0666	7.5084	8.0452	8.8343	0.0511	−1.5754	0.4219	0.2107	0.0007
8.0	−43.8557757	0.8385	0.2424	0.0581	7.2348	7.9897	8.7906	0.0478	−1.6403	0.3713	0.1872	0.0006
8.2	−43.8557467	0.7230	0.2090	0.0508	7.0041	7.9424	8.7531	0.0450	−1.6933	0.3281	0.1718	0.0006
8.4	−43.8557098	0.6243	0.1805	0.0446	6.8091	7.9021	8.7212	0.0426	−1.7372	0.2854	0.1539	0.0012
8.6	−43.8556670	0.5399	0.1561	0.0392	6.6437	7.8674	8.6940	0.0405	−1.7800	0.2487	0.1379	0.0024
8.8	−43.8556203	0.4676	0.1352	0.0345	6.5029	7.8375	8.6710	0.0386	−1.8027	0.2171	0.1218	0.0023
9.0	−43.8555715	0.4058	0.1173	0.0304	6.3828	7.8118	8.6517	0.0369	−1.8265	0.1900	0.1110	0.0032
9.2	−43.8555220	0.3528	0.1020	0.0269	6.2801	7.7895	8.6354	0.0355	−1.8460	0.1666	0.1026	0.0031
9.4	−43.8554730	0.3074	0.0889	0.0239	6.1920	7.7704	8.6218	0.0343	−1.8630	0.1465	0.0923	0.0034
9.6	−43.8554252	0.2683	0.0776	0.0211	6.1162	7.7536	8.6103	0.0331	−1.8768	0.1290	0.0870	0.0032
9.8	−43.8553794	0.2347	0.0679	0.0188	6.0509	7.7392	8.6007	0.0322	−1.8886	0.1139	0.0870	0.0029
10.0	−43.8553360	0.2058	0.0596	0.0167	5.9945	7.7266	8.5927	0.0314	−1.8990	0.1009	0.0828	0.0026
10.2	−43.8552952	0.1808	0.0524	0.0149	5.9456		8.5858	0.0306			0.0758	0.0024
10.4	−43.8552572	0.1592	0.0462	0.0132	5.9031		8.5801	0.0299			0.0713	0.0021
10.6	−43.8552221	0.1406	0.0408	0.0119	5.8663		8.5753	0.0293			0.0674	0.0019
10.8	−43.8551898	0.1244	0.0362	0.0106	5.8342		8.5713	0.0288			0.0639	0.0017
11.0	−43.8551602	0.1103	0.0321	0.0095	5.8061		8.5678	0.0283			0.0607	0.0015
11.2	−43.8551332	0.0980	0.0285	0.0085	5.7815		8.5649	0.0279			0.0579	0.0013
11.4	−43.8551087	0.0873	0.0254	0.0077	5.7601		8.5624	0.0276			0.0555	0.0012
11.6	−43.8550864	0.0779	0.0227	0.0068	5.7411		8.5602	0.0272			0.0289	0.0011
11.8	−43.8550663	0.0697	0.0203	0.0062	5.7246		8.5583	0.0269			0.0270	0.0010
12.0	−43.8550482	0.0625	0.0183	0.0056	5.7100		8.5568	0.0267			0.0253	0.0009
12.2	−43.8550318	0.0561	0.0164	0.0050	5.6971		8.5554	0.0264			0.0238	0.0008
12.4	−43.8550171	0.0506	0.0148	0.0046	5.6858		8.5543	0.0263			0.0225	0.0008
12.6	−43.8550038	0.0456	0.0133	0.0041	5.6757		8.5532	0.0260			0.0110	0.0007
12.8	−43.8549919	0.0412	0.0121	0.0038	5.6668		8.5523	0.0259			0.0102	0.0006
13.0	−43.8549812	0.0374	0.0109	0.0035	5.6589		8.5516	0.0258			0.0094	0.0006
16.0	−43.8549040	0.0101	0.0029	0.0009	5.6044		8.5465	0.0247			0.0045	0.0001
20.0	−43.8548828	0.0025	0.0007	0.0002	5.5899		8.5454	0.0244			0.0035	0.0001
25.0	−43.8548776	0.0007	0.0002	0.0001	5.5864		8.5453	0.0244			0.0034	0.0001

a From Ref. [161].
b From Ref. [166].
c Using $\eta = 10^{-4}$ in case of CAS(6,6)$_a$ and $\eta = 10^{-6}$ otherwise.

Table A12: Primary data for Fig. 28 (BeH$_2$ model system), as published in Ref. [P3]. The results for the various ic-MRCCx methods are given as energy differences to FCI in mE_h.

	Simultaneous ansatz		Sequential ansatz			
z / a_0	SD(T)	SDT-1	SD	SD(T)	3	SDT
0	0.0768	0.0765	0.3846	0.0787	0.0781	-0.0010
1	0.0886	0.0859	0.4097	0.0913	0.0881	0.0029
2	0.1133	0.1058	0.6721	0.1132	0.1183	0.0126
2.5	0.2225	0.2326	1.0084	0.2259	0.2384	0.0195
2.625	0.3123	0.3176	1.2285	0.3053	0.3284	0.0158
2.6875	0.3863	0.3836	1.4036	0.3633	0.3913	0.0142
2.725	0.4435	0.4329	1.5421	0.4032	0.4329	0.0143
2.75	0.4872	0.4695	1.6495	0.4306	0.4601	0.0146
2.775	0.5345	0.5084	1.7648	0.4576	0.4845	0.0142
2.8125	0.6078	0.5670	1.9254	0.4940	0.5080	0.0100
2.85	0.6642	0.6100	2.0070	0.5124	0.4959	0.0008
2.875	0.6730	0.6145	1.9900	0.5024	0.4581	-0.0046
2.9	0.6477	0.5905	1.9234	0.4723	0.4027	-0.0064
2.95	0.5200	0.4756	1.7301	0.3838	0.2935	0.0001
3	0.3838	0.3560	1.5533	0.3077	0.2312	0.0114
3.1	0.2488	0.2492	1.3215	0.2370	0.2107	0.0282
3.25	0.2223	0.2431	1.1259	0.2221	0.2375	0.0367
3.5	0.2187	0.2433	0.8994	0.2135	0.2406	0.0331
4	0.1429	0.1584	0.5199	0.1386	0.1566	0.0172

Table A13: Primary data for Fig. 29 (H$_2$O model system). The results for the various ic-MRCCSDx methods (using a simultaneous exponential ansatz) are given as energy differences to FCI in mE_h.

$R_{\text{O-H}}/R_e$	(T){2}[a]	(T){3}[a]	(T){4}[a,b]	(T){3}$_{\text{it}}$	(T){3}$_{\text{simp}}$	T-1{3}[a]	T{3}[a]
1.0	0.5146	0.3521	0.3251	0.3516	0.3637	0.3578	0.1527
1.2	0.5452	0.3439	0.3340	0.3478	0.3490	0.3487	0.1624
1.4	0.5236	0.3227	0.3409	0.3293	0.3229	0.4287	0.1809
1.6	0.4700	0.2919	0.2977	0.2991	0.3029	0.3310	0.1783
1.8	0.4086	0.2538	0.2472	0.2616	0.2810	0.2768	0.1686
2.0	0.3426	0.2162	0.2080	0.2247	0.2565	0.2350	0.1448
2.2	0.2809	0.1850	0.1803	0.1938	0.2357	0.2007	0.1257
2.4	0.2338	0.1638	0.1619	0.1725	0.2214	0.1770	0.1129
2.6	0.2043	0.1520	0.1513	0.1603	0.2138	0.1632	0.1058
2.8	0.1994	0.1577	0.1574	0.1658	0.1903	0.1667	0.1020
3.0	0.1935	0.1579	0.1578	0.1666	0.1923	0.1670	0.1018
3.2	0.2092	0.1755	0.1754	0.1869	0.1755	0.1794	0.0843
3.4	0.2165	0.1853	0.1852	0.1969	0.1681	0.1884	0.0857
3.6	0.2144	0.1846	0.1846	0.1962	0.1675	0.1881	0.0860
3.8	0.2167	0.1876	0.1876	0.2003	0.1716	0.1928	0.0867

[a] Published in Ref. [P3].
[b] Excluding excitation types $\hat{a}_{uij}^{vwx}$ and $\hat{a}_{uvw}^{xab}$, which are redundant for CAS(4,4) when $\eta \to 0$.

Table A14: Primary data for Fig. 30 (N₂ model system). The results for the various ic-MRCCx methods (using a sequential exponential ansatz) are given as energy differences to FCI in mE_h.

$R_{\text{N-N}} / a_0$	$\eta = 10^{-6}$				$\eta = 10^{-4}$						
	SD[a]	SD(T){3}[a]	3{3}[a]	SDT{3}[a]	SD	SD(T){2}	SD(T){3}	SD(T){4}	SD(T){3}$_{\text{it}}$	SD(T){3}$_{\text{simp}}$	SDT{3}
2.05	1.6639	0.1909	0.1178	0.1286	1.6552	0.5939	0.1535	0.1085	0.1234	0.4662	0.139
2.1	1.6642	0.1859	0.1228	0.1376	1.6540	0.6050	0.1475	0.0897	0.1208	0.4665	0.149
2.15	1.6685	0.1802	0.1306	0.1495	1.6533	0.6190	0.1457	0.0765	0.1222	0.4671	0.158
2.3	1.6530	0.1757	0.1754	0.2031	1.5764	0.6027	0.0956	0.0522	0.0804	0.4309	0.150
2.5	1.5447	0.0746	0.0627	0.1974	1.5441	0.6312	0.0877	0.0299	0.0837	0.4186	0.164
2.75	1.2609	−0.1541	−0.0628	0.1956	1.5221	0.6807	0.1206	0.0617	0.1218	0.4132	0.181
3	1.3435	0.0543	0.3048	0.4181	1.4559	0.7038	0.1883	0.0954	0.1902	0.4128	0.166
3.25	1.4301	0.3271	0.3631	0.2583	1.3539	0.6851	0.2472	0.1177	0.2503	0.4032	0.146
3.5	1.4445	0.4026	0.4593	0.2888	1.4539	0.8440	0.3892	0.2503	0.3918	0.4063	0.236
3.75	1.2768	0.4030	0.4666	0.2275	1.3521	0.7965	0.4052	0.3012	0.4081	0.3830	0.203
4	1.1241	0.3828	0.4274	0.1608	1.2427	0.7258	0.4047	0.3239	0.4070	0.3569	
4.5	0.9198	0.3233	0.3229	0.0563	1.2077	0.6867	0.4435	0.3749	0.4489	0.3370	
5	0.8928	0.3400	0.3470	0.0829	1.4918	0.6865	0.4973	0.4548	0.5033	0.3860	
5.5	0.9531	0.3567	0.3441	0.0776	1.6201	0.6345	0.4980	0.4670	0.5018	0.4268	
6	1.2159	0.4368	0.4396	0.0799	1.5646	0.5917	0.4798	0.4585	0.4815	0.4089	
7	1.5263	0.4694	0.4762	0.0770	1.5263	0.5651	0.4694	0.4519	0.4697	0.3989	
8	1.5135	0.4626	0.4694	0.0711	1.5135	0.5553	0.4626	0.4450	0.4626	0.3921	
9	1.5166	0.4669	0.4737	0.0756	1.5166	0.5592	0.4669	0.4488	0.4669	0.3963	
10	1.5166	0.4673	0.4741	0.0760	1.5166	0.5595	0.4673	0.4488	0.4673	0.3967	

[a] Published in Ref. [P3].

Table A15: Primary data for Fig. 31: Equilibrium structure (in Å and degree), harmonic vibrational frequencies (in cm^{-1}) and absolute energies (in E_h) of ozone (published in Ref. [P3]).

Method	Basis	r_e	θ_e	$\omega_1(a_1)$	$\omega_2(a_1)$	$\omega_3(b_2)$	E
CCSD[a]	DZ	1.2590	117.3	1251	753	1237	
	TZ	1.2500	117.6	1278	763	1266	
	QZ	1.2438	117.7	1294	771	1282	
CCSD(T)[b]	DZ	1.284	116.6	1118	704	977	
	TZ	1.275	116.9	1153	716	1054	
	QZ	1.269	117.1	1169	725	1081	
	5Z	1.267	117.1				
	5Z'	1.268	117.1	1169	725	1079	
CCSDT	DZ[a]	1.2834	116.5	1127	706	1065	
	TZ[b]	1.274	116.8	1163	717	1117	
Mk-MRCCSD	DZ[a]	1.2764	115.9	1146	728	1239	
	TZ[a]	1.2663	116.3	1180	739	1289	
	QZ[a]	1.2594	116.5	1198	748	1310	
	5Z[c]	1.2579	116.5	1203	751	1314	
Mk-MRCCSD(T)[c]	TZ	1.2897	116.1	1081	694	1212	
	QZ	1.2829	116.3	1096	702	1141	
	5Z	1.2814	116.3	1097	704	1048	
ic-MRCCSD	DZ	1.2783	116.14	1135	724	1178	-224.89687
	TZ	1.2677	116.48	1172	736	1238	-225.10788
	QZ	1.2607	116.68	1192	746	1261	-225.17398
	5Z'	1.2602	116.72	1192	746	1261	-225.19133
ic-MRCCSD(T)	DZ	1.2870	116.23	1102	702	1051	-224.91239
	TZ	1.2782	116.54	1136	712	1107	-225.13545
	QZ	1.2714	116.73	1153	721	1135	-225.20492
	5Z'	1.2710	116.75	1152	721	1130	-225.22351
ic-MRCCSD(T)$_{it}$	DZ	1.2870	116.23	1102	702	1050	-224.91239
	TZ	1.2782	116.54	1136	712	1106	-225.13546
	QZ	1.2714	116.73	1153	721	1127	-225.20494
ic-MRCCSDT-1	DZ	1.2923	116.24	1076	688	993	-224.91529
	TZ	1.2835	116.55	1111	699	1053	-225.13910
	QZ	1.2767	116.73	1127	708	1074	-225.20884
ic-MRCC3 (seq.)	DZ	1.2922	116.25	1074	688	992	-224.91508
	TZ	1.2836	116.56	1107	699	1050	-225.13902
	QZ	1.2767	116.74	1123	708	1071	-225.20878
ic-MRCCSDT	DZ	1.2914	116.26	1082	690	1005	-224.91499
	TZ	1.2819	116.56	1118	703	1069	-225.13752
Experiment		1.2728[d]	116.75[d]	1133[e]	715[e]	1087[e]	

[a] From Ref. [91].
[b] From Ref. [211].
[c] From Ref. [189].
[d] From Ref. [207].
[e] From Ref. [208].

Table A16: Primary data for Figs. 35a and 35b: Cumulant norms $\|\boldsymbol{\lambda}_n\|$ for the H_2O model.

| $R_{\mathrm{O-H}}/R_e$ | n | | | | | | | |
	1	2	3	4	5	6	7	8
1.0	1.97591	0.22266	0.03323	$4.667 \cdot 10^{-3}$	$9.670 \cdot 10^{-4}$	$1.266 \cdot 10^{-4}$	$2.959 \cdot 10^{-6}$	$1.272 \cdot 10^{-6}$
1.2	1.95017	0.32488	0.06646	$1.201 \cdot 10^{-2}$	$3.006 \cdot 10^{-3}$	$5.163 \cdot 10^{-4}$	$2.986 \cdot 10^{-5}$	$9.432 \cdot 10^{-6}$
1.4	1.89807	0.47740	0.12660	$3.173 \cdot 10^{-2}$	$7.671 \cdot 10^{-3}$	$1.802 \cdot 10^{-3}$	$2.949 \cdot 10^{-4}$	$7.858 \cdot 10^{-6}$
1.6	1.81242	0.67305	0.20398	$8.350 \cdot 10^{-2}$	$1.600 \cdot 10^{-2}$	$5.536 \cdot 10^{-3}$	$3.177 \cdot 10^{-4}$	$9.768 \cdot 10^{-5}$
1.8	1.70233	0.88229	0.26259	$1.915 \cdot 10^{-1}$	$3.194 \cdot 10^{-2}$	$2.505 \cdot 10^{-2}$	$5.037 \cdot 10^{-3}$	$6.021 \cdot 10^{-3}$
2.0	1.59235	1.06450	0.27160	$3.453 \cdot 10^{-1}$	$5.109 \cdot 10^{-2}$	$8.608 \cdot 10^{-2}$	$1.160 \cdot 10^{-2}$	$3.689 \cdot 10^{-2}$
2.2	1.50987	1.18825	0.23842	$4.872 \cdot 10^{-1}$	$6.596 \cdot 10^{-2}$	$1.697 \cdot 10^{-1}$	$8.615 \cdot 10^{-3}$	$8.930 \cdot 10^{-2}$
2.4	1.46201	1.25610	0.19107	$5.776 \cdot 10^{-1}$	$7.074 \cdot 10^{-2}$	$2.316 \cdot 10^{-1}$	$2.796 \cdot 10^{-3}$	$1.302 \cdot 10^{-1}$
2.6	1.43784	1.28973	0.14597	$6.243 \cdot 10^{-1}$	$6.387 \cdot 10^{-2}$	$2.631 \cdot 10^{-1}$	$2.123 \cdot 10^{-3}$	$1.509 \cdot 10^{-1}$
2.8	1.42602	1.30621	0.10844	$6.471 \cdot 10^{-1}$	$5.173 \cdot 10^{-2}$	$2.771 \cdot 10^{-1}$	$2.480 \cdot 10^{-3}$	$1.597 \cdot 10^{-1}$
3.0	1.42017	1.31444	0.07917	$6.582 \cdot 10^{-1}$	$3.946 \cdot 10^{-2}$	$2.832 \cdot 10^{-1}$	$2.201 \cdot 10^{-3}$	$1.635 \cdot 10^{-1}$
3.2	1.41723	1.31860	0.05716	$6.637 \cdot 10^{-1}$	$2.913 \cdot 10^{-2}$	$2.860 \cdot 10^{-1}$	$1.738 \cdot 10^{-3}$	$1.651 \cdot 10^{-1}$
3.4	1.41573	1.32071	0.04092	$6.664 \cdot 10^{-1}$	$2.109 \cdot 10^{-2}$	$2.874 \cdot 10^{-1}$	$1.299 \cdot 10^{-3}$	$1.659 \cdot 10^{-1}$
3.6	1.41497	1.32179	0.02907	$6.679 \cdot 10^{-1}$	$1.507 \cdot 10^{-2}$	$2.880 \cdot 10^{-1}$	$9.428 \cdot 10^{-4}$	$1.663 \cdot 10^{-1}$
3.8	1.41459	1.32234	0.02047	$6.686 \cdot 10^{-1}$	$1.065 \cdot 10^{-2}$	$2.884 \cdot 10^{-1}$	$6.710 \cdot 10^{-4}$	$1.665 \cdot 10^{-1}$
∞	1.41421	1.32288	0	0.66927	0	0.28868	0	0.16667

Table A17: Primary data for Fig. 36 (symmetric dissociation of H_2O, ic-MRCCSD), given as energy differences to FCI in mE_h. 'n.c.' stands for 'not converged'.

| $R_{\mathrm{O-H}}/R_e$ | $\eta = 10^{-2}$ | | $\eta = 10^{-3}$ | | $\eta = 10^{-4}$ | |
	$n_{\max} = 3$	$n_{\max} = 2$	$n_{\max} = 3$	$n_{\max} = 2$	$n_{\max} = 3$	$n_{\max} = 2$
1.0	3.0703	3.1971	1.3592	n.c.	1.3353	n.c.
1.2	1.8416	1.9424	1.3736	n.c.	1.2974	n.c.
1.4	1.8212	2.2454	1.3057	n.c.	1.1846	n.c.
1.6	1.9104	2.5266	1.2873	n.c.	1.1057	n.c.
1.8	1.5432	2.0550	1.2466	n.c.	1.0613	n.c.
2.0	1.3438	1.6939	1.1626	n.c.	1.0337	n.c.
2.2	1.1491	1.3683	1.1491	1.3683	0.9823	n.c.
2.4	1.1189	1.1990	1.0146	1.0940	0.9267	n.c.
2.6	0.9879	1.0139	0.9296	0.9514	0.9296	n.c.
2.8	0.9229	0.9320	0.9229	0.9320	0.8816	n.c.
3.0	0.8910	0.8947	0.8910	0.8947	n.c.	n.c.
3.2	0.8745	0.8763	0.8745	0.8763	0.8745	0.8763
3.4	0.8655	0.8664	0.8655	0.8664	0.8655	0.8664
3.6	0.8605	0.8610	0.8605	0.8610	0.8605	0.8610
3.8	0.8577	0.8579	0.8577	0.8579	0.8577	0.8579

Table A18: Primary data for Fig. 35c: Cumulant norms $\|\boldsymbol{\lambda}_n\|$ for the N_2 model system.

$R_{\mathrm{N-N}}$ / a_0	1	2	3	4	5	6	7	8	9	10	11	12
2.1	2.39323	0.38731	0.08321	0.02707	$9.656{\cdot}10^{-3}$	$3.817{\cdot}10^{-3}$	$1.415{\cdot}10^{-3}$	$4.016{\cdot}10^{-4}$	$6.799{\cdot}10^{-5}$	$1.790{\cdot}10^{-5}$	$3.202{\cdot}10^{-6}$	$1.169{\cdot}10^{-6}$
2.5	2.34321	0.54679	0.14886	0.06081	$2.491{\cdot}10^{-2}$	$1.332{\cdot}10^{-2}$	$6.278{\cdot}10^{-3}$	$2.110{\cdot}10^{-3}$	$4.902{\cdot}10^{-4}$	$2.013{\cdot}10^{-4}$	$5.735{\cdot}10^{-5}$	$1.865{\cdot}10^{-5}$
3.0	2.21819	0.84592	0.26574	0.16143	$5.678{\cdot}10^{-2}$	$3.714{\cdot}10^{-2}$	$1.802{\cdot}10^{-2}$	$8.709{\cdot}10^{-3}$	$3.721{\cdot}10^{-3}$	$1.295{\cdot}10^{-3}$	$2.440{\cdot}10^{-4}$	$6.386{\cdot}10^{-5}$
3.3	2.09192	1.09428	0.31786	0.32031	$8.455{\cdot}10^{-2}$	$1.009{\cdot}10^{-1}$	$3.520{\cdot}10^{-2}$	$4.727{\cdot}10^{-2}$	$1.404{\cdot}10^{-2}$	$1.541{\cdot}10^{-2}$	$3.627{\cdot}10^{-3}$	$2.944{\cdot}10^{-3}$
3.6	1.95670	1.33618	0.31556	0.55079	$1.204{\cdot}10^{-1}$	$3.184{\cdot}10^{-1}$	$7.674{\cdot}10^{-2}$	$2.640{\cdot}10^{-1}$	$6.024{\cdot}10^{-2}$	$1.893{\cdot}10^{-1}$	$3.429{\cdot}10^{-2}$	$9.108{\cdot}10^{-2}$
4.2	1.80061	1.60554	0.22463	0.86301	$1.264{\cdot}10^{-1}$	$7.848{\cdot}10^{-1}$	$8.576{\cdot}10^{-2}$	$9.700{\cdot}10^{-1}$	$7.016{\cdot}10^{-2}$	1.10574	$4.762{\cdot}10^{-2}$	$8.702{\cdot}10^{-1}$
∞	1.73205	1.71998	0	1.00580	0	1.00967	0	1.37205	0	1.71941	0	1.49212

Table A19: Primary data for Fig. 37 (N_2 model system, ic-MRCCSD using a sequential exponential ansatz), given as energy differences to FCI in mE_h. 'n.c.' stands for 'not converged'.

$R_{\mathrm{N-N}}$ / a_0	$\eta = 10^{-2}$ ($n_{\max}$)				$\eta = 10^{-3}$ ($n_{\max}$)				$\eta = 10^{-4}$ ($n_{\max}$)			
	5 (full)	4	3	2	5 (full)	4	3	2	5 (full)	4	3	2
2.05	3.0390	3.0390	3.0143	4.5926	1.9689	1.9688	1.9414	n.c.	1.6552	1.6555	1.6272	n.c.
2.1	2.7268	2.7266	2.6977	4.5130	1.9455	1.9455	1.9209	n.c.	1.6540	1.6544	1.6249	n.c.
2.15	2.8402	2.8400	2.8047	4.7925	1.9709	1.9709	1.9364	n.c.	1.6533	1.6537	1.6227	n.c.
2.3	2.9409	2.9404	2.8741	4.7132	1.9882	1.9885	1.9501	n.c.	1.5764	1.6060	n.c.	n.c.
2.5	3.0116	3.0104	2.8853	5.3690	1.9860	1.9870	1.9550	n.c.	1.5441	1.5770	n.c.	n.c.
2.75	3.9044	3.9015	3.6826	6.2672	1.7263	1.7337	1.7695	n.c.	1.5221	1.5987	n.c.	n.c.
3	3.6424	3.6399	3.3643	6.2316	1.8382	1.8599	n.c.	n.c.	1.4559	1.5259	n.c.	n.c.
3.25	3.4982	3.4973	3.3895	6.8742	1.9623	1.9654	n.c.	n.c.	1.3539	1.3460	n.c.	n.c.
3.5	4.1293	4.1290	n.c.	n.c.	1.8972	1.8965	n.c.	n.c.	1.4539	n.c.	n.c.	n.c.
3.75	3.5126	3.5122	3.8237	n.c.	1.7091	1.7088	n.c.	n.c.	1.3521	n.c.	n.c.	n.c.
4	3.1019	3.1016	3.4264	3.6320	1.7313	1.7309	n.c.	n.c.	1.2427	1.2089	n.c.	n.c.
4.5	2.0984	2.0983	2.3259	2.4075	1.7888	1.7888	n.c.	n.c.	1.2077	1.2017	n.c.	n.c.
5	1.7496	1.7496	1.8950	1.9153	1.7496	1.7496	1.8950	1.9153	1.4918	1.4914	n.c.	n.c.
5.5	1.6202	1.6202	1.7218	1.7296	1.6202	1.6202	1.7218	1.7296	1.6201	1.6201	n.c.	n.c.
6	1.5646	1.5646	1.6451	1.6489	1.5646	1.5646	1.6451	1.6489	1.5646	1.5646	1.6451	1.6489
7	1.5263	1.5263	1.5928	1.5936	1.5263	1.5263	1.5928	1.5936	1.5263	1.5263	1.5928	1.5936
8	1.5135	1.5135	1.5775	1.5776	1.5135	1.5135	1.5775	1.5776	1.5135	1.5135	1.5775	1.5776
9	1.5166	1.5166	1.5803	1.5803	1.5166	1.5166	1.5803	1.5803	1.5166	1.5166	1.5803	1.5803
10	1.5166	1.5166	1.5803	1.5803	1.5166	1.5166	1.5803	1.5803	1.5166	1.5166	1.5803	1.5803

Table A20: Primary data for Fig. 38 (dissociation of Cr_2), given as total energies in E_h. 'n.c.' indicates calculations that either showed no or very slow convergence.

R_{Cr-Cr}		$\eta = 10^{-3}$, $n_{max} = 4$		$\eta = 10^{-2}$, $n_{max} = 4$	
/ Å	CASSCF	ic-MRCCSD	ic-MRCCSD(T){3}	ic-MRCCSD	ic-MRCCSD(T){3}
1.5	-2099.226982	-2100.317734	-2100.343681	-2100.311148	-2100.338353
1.58	-2099.244657	-2100.336510	-2100.362800	-2100.328546	-2100.356301
1.63	-2099.249103	-2100.341277	-2100.367683	-2100.333898	-2100.361714
1.6788	-2099.250963	-2100.342715	-2100.369261		
1.68	-2099.250991			-2100.334995	-2100.362990
1.73	-2099.251802	-2100.342265	-2100.368881	-2100.334567	-2100.362663
1.78	-2099.252512	-2100.341161	-2100.367694	-2100.333178	-2100.361335
1.83	-2099.253640	-2100.339718	-2100.366153	n.c.	
1.9	-2099.256266			-2100.328919	-2100.357179
2.0	-2099.262005	-2100.335341	-2100.361609	-2100.326012	-2100.354453
2.1	-2099.269181			n.c.	
2.2	-2099.276604	n.c.		-2100.323922	-2100.353437
2.3	-2099.283417			-2100.325602	-2100.355114
2.4	-2099.289204	n.c.		-2100.327221	-2100.356497
2.5	-2099.293860			-2100.328149	-2100.357166
2.7	-2099.300089	n.c.		-2100.327901	-2100.356397
2.9	-2099.303177			-2100.325496	-2100.353488
3.1	-2099.304271	n.c.		-2100.321916	-2100.349442
3.3	-2099.304249			-2100.317988	-2100.345106
∞	-2099.299459	-2100.299717	-2100.325380	-2100.299717	-2100.325380

A.3   Publications

[P1] *Pilot applications of internally contracted multireference coupled cluster theory, and how to choose the cluster operator properly*
M. Hanauer and A. Köhn, J. Chem. Phys. **134**, 204111 (2011).

[P2] *Meaning and magnitude of the reduced density matrix cumulants*
M. Hanauer and A. Köhn, Chem. Phys. **401**, 50 (2012).

[P3] *Perturbative treatment of triple excitations in internally contracted multireference coupled cluster theory*
M. Hanauer and A. Köhn, J. Chem. Phys. **136**, 204107 (2012).

[P4] *Communication: Restoring full size extensivity in internally contracted multireference coupled cluster theory*
M. Hanauer and A. Köhn, J. Chem. Phys. **137**, 131103 (2012).

[P5] *A sequential transformation approach to the internally contracted multireference coupled cluster method*
F. A. Evangelista, M. Hanauer, A. Köhn and J. Gauss, J. Chem. Phys. **136**, 204108 (2012).

[P6] *State-specific multireference coupled-cluster theory*
A. Köhn, M. Hanauer, L. A. Mück, T.-C. Jagau and J. Gauss, WIREs Comput. Mol. Sci. **3**, 176 (2013).

[P7] *Explicitly correlated internally contracted multireference coupled-cluster singles and doubles theory: ic-MRCCSD-F12*
W. Liu, M. Hanauer and A. Köhn, Chem. Phys. Lett. **565**, 122 (2013).

Bibliography

[1] E. B. Wilson, Pure Appl. Chem. **47**, 41 (1976).

[2] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* Dover Publications, Mineola, 1989.

[3] T. Helgaker, P. Jørgensen and J. Olsen. *Molecular Electronic-Structure Theory.* John Wiley & Sons Ltd., Chichester, 2000.

[4] J. Gauss. *Molecular Properties*, in *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotendorst, pp. 509–560. John von Neumann Institute for Computing, Jülich, 2000.

[5] C. J. Barden and H. F. Schaefer III, Pure Appl. Chem. **72**, 1405 (2000).

[6] M. Head-Gordon and E. Artacho, Phys. Today **61**, 58 (2008).

[7] K. Fukui, T. Yonezawa and H. Shingu, J. Chem. Phys. **20**, 722 (1952).

[8] R. B. Woodward and R. Hoffmann, Angew. Chem. Int. Ed. **8**, 781 (1969).

[9] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).

[10] T. D. Crawford and H. F. Schaefer III. *An Introduction to Coupled Cluster Theory for Computational Chemists*, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd, volume 14, pp. 33–136. Wiley, New York, 2000.

[11] I. Shavitt and R. J. Bartlett. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory.* Cambridge University Press, New York, 2009.

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

[13] J. D. Watts, J. Gauss and R. J. Bartlett, J. Chem. Phys. **98**, 8718 (1993).

[14] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez and J. F. Stanton, J. Chem. Phys. **121**, 11599 (2004).

[15] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[16] H. J. Silverstone and O. Sinanoğlu, J. Chem. Phys. **44**, 1899 (1966).

[17] F. Bernardi, A. Bottoni, J. J. W. McDouall, M. A. Robb and H. B. Schlegel, Faraday Symp. Chem. Soc. **19**, 137 (1984).

[18] H. H. Wenk, M. Winkler and W. Sander, Angew. Chem. Int. Ed. **42**, 502 (2003).

[19] K. Pierloot, Int. J. Quantum Chem. **111**, 3291 (2011).

[20] Y. Gong, M. Zhou and L. Andrews, Chem. Rev. **109**, 6765 (2009).

[21] W. Jiang, N. J. DeYonker and A. K. Wilson, J. Chem. Theory Comput. **8**, 460 (2012).

[22] H.-J. Werner, Adv. Chem. Phys. **69**, 1 (1987).

[23] B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).

[24] M. W. Schmidt and M. S. Gordon, Annu. Rev. Phys. Chem. **49**, 233 (1998).

[25] P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka and R. Shepard, Chem. Rev. **112**, 108 (2012).

[26] C. D. Sherrill, J. Chem. Phys. **132**, 110902 (2010).

[27] B. Jeziorski, Mol. Phys. **108**, 3043 (2010).

[28] J. Paldus, J. Pittner and P. Čársky. *Multireference Coupled-Cluster Methods: Recent Developments*, in *Recent Progress in Coupled Cluster Methods*, edited by J. Leszczynski, P. Čársky, J. Paldus and J. Pittner, pp. 455–489. Springer, London, 2010.

[29] D. I. Lyakh, M. Musiał, V. F. Lotrich and R. J. Bartlett, Chem. Rev. **112**, 182 (2012).

[30] D. Mukherjee. *A Coupled Cluster Approach to the Electron Correlation Problem Using a Correlated Reference State*, in *Recent Progress in Many-Body Theories*, edited by E. Schachinger, H. Mitter and H. Sormann, volume 4, pp. 127–133. Plenum Press, New York, 1995.

[31] U. S. Mahapatra, B. Datta, B. Bandyopadhyay and D. Mukherjee, Adv. Quantum Chem. **30**, 163 (1998).

[32] M. Nooijen and V. Lotrich, J. Mol. Struct. (Theochem) **547**, 253 (2001).

[33] L. Kong, Int. J. Quantum Chem. **110**, 2603 (2010).

[34] A. Banerjee and J. Simons, Int. J. Quantum Chem. **19**, 207 (1981).

[35] A. Banerjee and J. Simons, J. Chem. Phys. **76**, 4548 (1982).

[36] A. Banerjee and J. Simons, Chem. Phys. **81**, 297 (1983).

[37] A. Banerjee and J. Simons, Chem. Phys. **87**, 215 (1984).

[38] F. A. Evangelista and J. Gauss, J. Chem. Phys. **134**, 114102 (2011).

[39] E. Neuscamman, T. Yanai and G. K.-L. Chan, J. Chem. Phys. **130**, 124102 (2009).

[40] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).

[41] T. Yanai and G. K.-L. Chan, J. Chem. Phys. **127**, 104107 (2007).

[42] E. Neuscamman, T. Yanai and G. K.-L. Chan, Int. Rev. Phys. Chem. **29**, 231 (2010).

[43] F. A. Evangelista and J. Gauss, Chem. Phys. **401**, 27 (2012).

[44] A. Köhn, J. Chem. Phys. **130**, 104104 (2009).

[45] M. Hanauer and A. Köhn, J. Chem. Phys. **131**, 124118 (2009).

[46] M. R. Hoffmann and J. Simons, J. Chem. Phys. **88**, 993 (1988).

[47] Z. Chen and M. R. Hoffmann, J. Chem. Phys. **137**, 014108 (2012).

[48] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).

[49] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[50] R. Fink and V. Staemmler, Theor. Chim. Acta **87**, 129 (1993).

[51] L. Kong, K. R. Shamasundar, O. Demel and M. Nooijen, J. Chem. Phys. **130**, 114101 (2009).

[52] T. Yanai, Y. Kurashige, E. Neuscamman and G. K.-L. Chan, Phys. Chem. Chem. Phys. **14**, 7809 (2012).

[53] W. Kutzelnigg. *Theory of Electron Correlation*, in *Explicitly Correlated Wave Functions in Chemistry and Physics: Theory and Applications*, edited by J. Rychlewski, pp. 3–90. Kluwer Academic Publishers, London, 2003.

[54] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **110**, 2800 (1999).

[55] W. Kutzelnigg, J. Chem. Phys. **80**, 822 (1984).

[56] G. C. Wick, Phys. Rev. **80**, 268 (1950).

[57] L. Kong and E. F. Valeev, J. Chem. Phys. **134**, 214109 (2011).

[58] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Nat. Chem. **2**, 106 (2010).

[59] R. J. Bartlett and J. F. Stanton. *Applications of Post-Hartree-Fock Methods: A Tutorial*, in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd, volume 5, pp. 65–169. Wiley, New York, 1994.

[60] X. Li and J. Paldus, J. Chem. Phys. **129**, 174101 (2008).

[61] A. T. Amos and G. G. Hall, Proc. Roy. Soc. London Ser. A **263**, 483 (1961).

[62] J. E. Harriman, J. Chem. Phys. **40**, 2827 (1964).

[63] P. Pulay and T. P. Hamilton, J. Chem. Phys. **88**, 4926 (1988).

[64] J. W. Hollett and P. M. W. Gill, J. Chem. Phys. **134**, 114111 (2011).

[65] T. J. Lee and P. R. Taylor, Int. J. Quantum Chem. Symp. **23**, 199 (1989).

[66] C. L. Janssen and I. M. B. Nielsen, Chem. Phys. Lett. **290**, 423 (1998).

[67] K. Boguslawski, P. Tecmer, Ö. Legeza and M. Reiher, J. Phys. Chem. Lett. **3**, 3129 (2012).

[68] B. Levy and G. Berthier, Int. J. Quantum Chem. **2**, 307 (1968).

[69] B. O. Roos and P. R. Taylor, Chem. Phys. **48**, 157 (1980).

[70] W. Kutzelnigg, Theor. Chim. Acta **83**, 263 (1992).

[71] T. Kinoshita, O. Hino and R. J. Bartlett, J. Chem. Phys. **123**, 074106 (2005).

[72] X. Li and J. Paldus, J. Chem. Phys. **107**, 6257 (1997).

[73] P. Piecuch, Mol. Phys. **108**, 2987 (2010).

[74] N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).

[75] V. V. Ivanov, D. I. Lyakh and L. Adamowicz, Phys. Chem. Chem. Phys. **11**, 2355 (2009).

[76] M. Kállay, P. G. Szalay and P. R. Surján, J. Chem. Phys. **117**, 980 (2002).

[77] D. I. Lyakh and R. J. Bartlett, J. Chem. Phys. **133**, 244112 (2010).

[78] K. Jankowski, J. Paldus and P. Piecuch, Theor. Chim. Acta **80**, 223 (1991).

[79] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).

[80] J. Shen and P. Piecuch, Chem. Phys. **401**, 180 (2012).

[81] A. I. Krylov, Chem. Phys. Lett. **338**, 375 (2001).

[82] S. V. Levchenko and A. I. Krylov, J. Chem. Phys. **120**, 175 (2004).

[83] M. Musiał, A. Perera and R. J. Bartlett, J. Chem. Phys. **134**, 114108 (2011).

[84] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).

[85] P. Piecuch and K. Kowalski, Int. J. Mol. Sci. **3**, 676 (2002).

[86] X. Li and J. Paldus, J. Chem. Phys. **119**, 5320 (2003).

[87] L. Kong, Int. J. Quantum Chem. **109**, 441 (2009).

[88] I. Hubač and P. Neogrády, Phys. Rev. A **50**, 4558 (1994).

[89] J. Mášik and I. Hubač, Adv. Quantum Chem. **31**, 75 (1998).

[90] U. S. Mahapatra, B. Datta and D. Mukherjee, J. Chem. Phys. **110**, 6171 (1999).

[91] F. A. Evangelista, W. D. Allen and H. F. Schaefer III, J. Chem. Phys. **127**, 024102 (2007).

[92] E. Prochnow, F. A. Evangelista, H. F. Schaefer III, W. D. Allen and J. Gauss, J. Chem. Phys. **131**, 064109 (2009).

[93] S. Das, M. Kállay and D. Mukherjee, Chem. Phys. **392**, 83 (2012).

[94] R. Maitra, D. Sinha and D. Mukherjee, J. Chem. Phys. **137**, 024105 (2012).

[95] S. Sen, A. Shee and D. Mukherjee, J. Chem. Phys. **137**, 074104 (2012).

[96] T.-C. Jagau and J. Gauss, Chem. Phys. **401**, 73 (2012).

[97] T.-C. Jagau and J. Gauss, J. Chem. Phys. **137**, 044116 (2012).

[98] T.-C. Jagau and J. Gauss, J. Chem. Phys. **137**, 044115 (2012).

[99] L. A. Mück and J. Gauss, J. Chem. Phys. **136**, 111103 (2012).

[100] U. S. Mahapatra and S. Chattopadhyay, J. Chem. Phys. **133**, 074102 (2010).

[101] M. Hanrath, J. Chem. Phys. **123**, 084102 (2005).

[102] M. Hanrath. *A Possibility For a Multi-Reference Coupled-Cluster: The MRexpT Ansatz*, in *Recent Progress in Coupled Cluster Methods*, edited by J. Leszczynski, P. Čársky, J. Paldus and J. Pittner, pp. 175–190. Springer, London, 2010.

[103] F. A. Evangelista and J. Gauss, J. Chem. Phys. **133**, 044101 (2010).

[104] S. Das, S. Pathak, R. Maitra and D. Mukherjee. *Development and Applications of Non-Perturbative Approximants to the State-Specific Multi-Reference Coupled Cluster Theory: The Two Distinct Variants*, in *Recent Progress in Coupled Cluster Methods*, edited by J. Leszczynski, P. Čársky, J. Paldus and J. Pittner, pp. 57–77. Springer, London, 2010.

[105] S. Das, S. Pathak, D. Datta and D. Mukherjee, J. Chem. Phys. **136**, 164104 (2012).

[106] D. Sinha, R. Maitra and D. Mukherjee, J. Chem. Phys. **137**, 094104 (2012).

[107] K. R. Shamasundar, G. Knizia and H.-J. Werner, J. Chem. Phys. **135**, 054101 (2011).

[108] D. Mukherjee, Mol. Phys. **30**, 1861 (1975).

[109] L. Z. Stolarczyk and H. J. Monkhorst, Phys. Rev. A **32**, 725 (1985).

[110] L. Z. Stolarczyk and H. J. Monkhorst, Mol. Phys. **108**, 3067 (2010).

[111] D. Mukherjee, Chem. Phys. Lett. **274**, 561 (1997).

[112] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **107**, 432 (1997).

[113] W. Kutzelnigg, K. R. Shamasundar and D. Mukherjee, Mol. Phys. **108**, 433 (2010).

[114] D. Datta, L. Kong and M. Nooijen, J. Chem. Phys. **134**, 214116 (2011).

[115] D. Datta and M. Nooijen, J. Chem. Phys. **137**, 204107 (2012).

[116] H. Baker and M. A. Robb, Mol. Phys. **50**, 1077 (1983).

[117] S. Li, J. Chem. Phys. **120**, 5017 (2004).

[118] T. Fang, J. Shen and S. Li. *Block Correlated Coupled Cluster Theory With a Complete Active-Space Self-Consistent-Field Reference Function: The General Formalism and Applications*, in *Recent Progress in Coupled Cluster Methods*, edited by J. Leszczynski, P. Čársky, J. Paldus and J. Pittner, pp. 145–174. Springer, London, 2010.

[119] B. H. Brandow, Rev. Mod. Phys. **39**, 771 (1967).

[120] M. B. Ruskai, J. Math. Phys. **11**, 3218 (1970).

[121] P.-O. Löwdin, Adv. Quantum Chem. **5**, 185 (1970).

[122] T. Shiozaki, G. Knizia and H.-J. Werner, J. Chem. Phys. **134**, 034113 (2011).

[123] D. Mukherjee and S. Pal, Adv. Quantum Chem. **20**, 291 (1989).

[124] D. Mukhopadhyay, S. Mukhopadhyay, R. Chaudhuri and D. Mukherjee, Theor. Chim. Acta **80**, 441 (1991).

[125] M. Nooijen, K. R. Shamasundar and D. Mukherjee, Mol. Phys. **103**, 2277 (2005).

[126] R. J. Bartlett and G. D. Purvis III, Int. J. Quantum Chem. **14**, 561 (1978).

[127] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).

[128] D. I. Lyakh, Mol. Phys. **110**, 1469 (2012).

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

[130] G. D. Purvis III, R. Shepard, F. B. Brown and R. J. Bartlett, Int. J. Quantum Chem. **23**, 835 (1983).

[131] F. A. Evangelista, W. D. Allen and H. F. Schaefer III, J. Chem. Phys. **125**, 154113 (2006).

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

[133] R. J. Gdanitz and R. Ahlrichs, Chem. Phys. Lett. **143**, 413 (1988).

[134] P. G. Szalay and R. J. Bartlett, J. Chem. Phys. **103**, 3600 (1995).

[135] J. Pittner, H. V. Gonzalez, R. J. Gdanitz and P. Čársky, Chem. Phys. Lett. **386**, 211 (2004).

[136] P. J. A. Ruttink, J. H. van Lenthe and P. Todorov, Mol. Phys. **103**, 2497 (2005).

[137] D. I. Lyakh, V. V. Ivanov and L. Adamowicz, Theor. Chem. Acc. **116**, 427 (2006).

[138] J. Olsen, H. Larsen, M. Fülscher and A. Köhn. Lucia, a general active space program.

[139] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).

[140] A. Engels-Putzka and M. Hanrath, J. Mol. Struct. (Theochem) **902**, 59 (2009).

[141] S. Das, D. Mukherjee and M. Kállay, J. Chem. Phys. **132**, 074103 (2010).

[142] S. Das, M. Kállay and D. Mukherjee, J. Chem. Phys. **133**, 234110 (2010).

[143] E. Dalgaard and P. Jørgensen, J. Chem. Phys. **69**, 3833 (1978).

[144] R. S. Berry, J. Chem. Phys. **27**, 1288 (1957).

[145] A. Devaquet, Pure Appl. Chem. **41**, 455 (1975).

[146] H.-J. Werner and W. Meyer, J. Chem. Phys. **74**, 5802 (1981).

[147] C. W. Bauschlicher Jr. and S. R. Langhoff, J. Chem. Phys. **89**, 4246 (1988).

[148] A. J. C. Varandas, J. Chem. Phys. **131**, 124128 (2009).

[149] D. E. Woon and T. H. Dunning Jr., J. Chem. Phys. **100**, 2975 (1994).

[150] C. W. Bauschlicher Jr. and P. R. Taylor, J. Chem. Phys. **85**, 2779 (1986).

[151] J. Olsen, P. Jørgensen, H. Koch, A. Balkova and R. J. Bartlett, J. Chem. Phys. **104**, 8007 (1996).

[152] S. Li, J. Ma and Y. Jiang, J. Chem. Phys. **118**, 5736 (2003).

[153] U. S. Mahapatra, S. Chattopadhyay and R. K. Chaudhuri, J. Chem. Phys. **129**, 024108 (2008).

[154] S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).

[155] E. R. Davidson and D. W. Silver, Chem. Phys. Lett. **52**, 403 (1977).

[156] P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993).

[157] J. Paldus, P. E. S. Wormer, F. Visser and A. van der Avoird, J. Chem. Phys. **76**, 2458 (1982).

[158] J. Wisborg Krogh and J. Olsen, Chem. Phys. Lett. **344**, 578 (2001).

[159] X. Li and J. Paldus, J. Chem. Phys. **129**, 054104 (2008).

[160] D. I. Lyakh, V. V. Ivanov and L. Adamowicz, Mol. Phys. **105**, 1335 (2007).

[161] J. M. Junquera-Hernández, J. Sánchez-Marín, G. L. Bendazzoli and S. Evangelisti, J. Chem. Phys. **120**, 8405 (2004).

[162] A. P. Rendell, T. J. Lee and P. R. Taylor, J. Chem. Phys. **92**, 7050 (1990).

[163] M. C. Heaven, J. M. Merritt and V. E. Bondybey, Annu. Rev. Phys. Chem. **62**, 375 (2011).

[164] P.-O. Widmark, P.-Å. Malmqvist and B. O. Roos, Theor. Chim. Acta **77**, 291 (1990).

[165] T. J. Lee, A. P. Rendell and P. R. Taylor, J. Chem. Phys. **92**, 489 (1990).

[166] J. Brabec, H. J. J. van Dam, J. Pittner and K. Kowalski, J. Chem. Phys. **136**, 124102 (2012).

[167] K. Kowalski, J. Chem. Phys. **134**, 194107 (2011).

[168] A. G. Taube, Mol. Phys. **108**, 2951 (2010).

[169] W. D. Laidig and R. J. Bartlett, Chem. Phys. Lett. **104**, 424 (1984).

[170] W. D. Laidig, P. Saxe and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1987).

[171] P. G. Szalay. *Towards State-Specific Formulation of Multireference Coupled-Cluster Theory: Coupled Electron Pair Approximations (CEPA) Leading to Multireference Configuration Interaction (MR-CI) Type Equations*, in *Recent Advances in Coupled-Cluster Methods*, edited by R. J. Bartlett, volume 3, pp. 81–123. World Scientific, London, 1997.

[172] Y. S. Lee, S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **81**, 5906 (1984).

[173] M. Urban, J. Noga, S. J. Cole and R. J. Bartlett, J. Chem. Phys. **83**, 4041 (1985).

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

[175] J. F. Stanton, Chem. Phys. Lett. **281**, 130 (1997).

[176] M. Kállay and J. Gauss, J. Chem. Phys. **123**, 214105 (2005).

[177] M. Kállay and J. Gauss, J. Chem. Phys. **129**, 144101 (2008).

[178] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 5243 (1998).

[179] T. D. Crawford and J. F. Stanton, Int. J. Quantum Chem. **70**, 601 (1998).

[180] S. R. Gwaltney and M. Head-Gordon, J. Chem. Phys. **115**, 2014 (2001).

[181] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).

[182] S. Hirata, M. Nooijen, I. Grabowski and R. J. Bartlett, J. Chem. Phys. **114**, 3919 (2001).

[183] D. I. Lyakh, V. F. Lotrich and R. J. Bartlett, Chem. Phys. Lett. **501**, 166 (2011).

[184] X. Li and J. Paldus, J. Chem. Phys. **124**, 174101 (2006).

[185] P. U. Manohar and A. I. Krylov, J. Chem. Phys. **129**, 194105 (2008).

[186] X. Li and J. Paldus, J. Chem. Phys. **124**, 034112 (2006).

[187] O. Demel and J. Pittner, J. Chem. Phys. **128**, 104108 (2008).

[188] K. Bhaskaran-Nair, O. Demel and J. Pittner, J. Chem. Phys. **129**, 184105 (2008).

[189] F. A. Evangelista, E. Prochnow, J. Gauss and H. F. Schaefer III, J. Chem. Phys. **132**, 074107 (2010).

[190] K. Bhaskaran-Nair, O. Demel, J. Šmydke and J. Pittner, J. Chem. Phys. **134**, 154106 (2011).

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

[192] K. G. Dyall, J. Chem. Phys. **102**, 4909 (1995).

[193] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger and J.-P. Malrieu, J. Chem. Phys. **114**, 10252 (2001).

[194] C. Angeli, R. Cimiraglia and J.-P. Malrieu, J. Chem. Phys. **117**, 9138 (2002).

[195] E. Neuscamman, T. Yanai and G. K.-L. Chan, J. Chem. Phys. **132**, 024106 (2010).

[196] D. P. Tew, W. Klopper and C. Hättig, Chem. Phys. Lett. **452**, 326 (2008).

[197] A. Köhn and D. P. Tew, J. Chem. Phys. **133**, 174117 (2010).

[198] M. Nooijen, J. Chem. Phys. **104**, 2638 (1996).

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

[200] R. Schinke, S. Y. Grebenshchikov, M. V. Ivanov and P. Fleurat-Lessard, Annu. Rev. Phys. Chem. **57**, 625 (2006).

[201] F. Holka, P. G. Szalay, T. Müller and V. G. Tyuterev, J. Phys. Chem. A **114**, 9927 (2010).

[202] R. Dawes, P. Lolur, J. Ma and H. Guo, J. Chem. Phys. **135**, 081102 (2011).

[203] X. Li and J. Paldus, J. Chem. Phys. **110**, 2844 (1999).

[204] I. Hubač, J. Pittner and P. Čársky, J. Chem. Phys. **112**, 8779 (2000).

[205] O. Hino, T. Kinoshita, G. K.-L. Chan and R. J. Bartlett, J. Chem. Phys. **124**, 114311 (2006).

[206] U. S. Mahapatra and S. Chattopadhyay, J. Chem. Phys. **134**, 044113 (2011).

[207] V. G. Tyuterev, S. Tashkun, P. Jensen, A. Barbe and T. Cours, J. Mol. Spectrosc. **198**, 57 (1999).

[208] A. Barbe, A. Chichery, T. Cours, V. G. Tyuterev and J. J. Plateaux, J. Mol. Struct. **616**, 55 (2002).

[209] J. Gauss, J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **95**, 2623 (1991).

[210] P. R. Bunker and P. Jensen. *Molecular Symmetry and Spectroscopy*. NRC Research Press, Ottawa, 2nd edition, 1998.

[211] J. D. Watts and R. J. Bartlett, J. Chem. Phys. **108**, 2511 (1998).

[212] K. A. Peterson and T. H. Dunning Jr., J. Chem. Phys. **117**, 10548 (2002).

[213] C. Wentrup, Aust. J. Chem. **63**, 979 (2010).

[214] M. Winkler and W. Sander, Aust. J. Chem. **63**, 1013 (2010).

[215] P. G. Wenthold, Aust. J. Chem. **63**, 1091 (2010).

[216] D. G. Leopold, A. E. S. Miller and W. C. Lineberger, J. Am. Chem. Soc. **108**, 1379 (1986).

[217] P. G. Wenthold, R. R. Squires and W. C. Lineberger, J. Am. Chem. Soc. **120**, 5279 (1998).

[218] R. Lindh, A. Bernhardsson and M. Schütz, J. Phys. Chem. A **103**, 9913 (1999).

[219] C. J. Cramer, J. J. Nash and R. R. Squires, Chem. Phys. Lett. **277**, 311 (1997).

[220] S. P. de Visser, M. Filatov and S. Shaik, Phys. Chem. Chem. Phys. **2**, 5046 (2000).

[221] L. V. Slipchenko and A. I. Krylov, J. Chem. Phys. **117**, 4694 (2002).

[222] H. Li, S.-Y. Yu, M.-B. Huang and Z.-X. Wang, Chem. Phys. Lett. **450**, 12 (2007).

[223] T. D. Crawford, E. Kraka, J. F. Stanton and D. Cremer, J. Chem. Phys. **114**, 10638 (2001).

[224] Y. Shao, M. Head-Gordon and A. I. Krylov, J. Chem. Phys. **118**, 4807 (2003).

[225] J. F. Stanton, J. Chem. Phys. **101**, 371 (1994).

[226] C. L. Janssen and H. F. Schaefer III, Theor. Chim. Acta **79**, 1 (1991).

[227] B. Jeziorski, J. Paldus and P. Jankowski, Int. J. Quantum Chem. **56**, 129 (1995).

[228] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **104**, 2652 (1996).

[229] D. Datta and D. Mukherjee, J. Chem. Phys. **134**, 054122 (2011).

[230] W. Jiang, N. J. DeYonker, J. J. Determan and A. K. Wilson, J. Phys. Chem. A **116**, 870 (2012).

[231] F. Allouti, L. Manceron and M. E. Alikhani, Phys. Chem. Chem. Phys. **8**, 3715 (2006).

[232] O. Hübner and H.-J. Himmel, Phys. Chem. Chem. Phys. **11**, 2241 (2009).

[233] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov and P.-O. Widmark, J. Phys. Chem. A **108**, 2851 (2004).

[234] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov and P.-O. Widmark, J. Phys. Chem. A **109**, 6575 (2005).

[235] M. Douglas and N. M. Kroll, Ann. Phys. **82**, 89 (1974).

[236] B. A. Hess, Phys. Rev. A **33**, 3742 (1986).

[237] K. Banert, B. Meier, E. Penk, B. Saha, E.-U. Würthwein, S. Grimme, T. Rüffer, D. Schaarschmidt and H. Lang, Chem. Eur. J. **17**, 1128 (2011).

[238] W. J. Hehre, R. Ditchfield and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[239] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).

[240] C. Lee, W. Yang and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[241] F. Colmenero and C. Valdemoro, Int. J. Quantum Chem. **51**, 369 (1994).

[242] H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996).

[243] D. A. Mazziotti, Phys. Rev. A **57**, 4219 (1998).

[244] D. A. Mazziotti, Phys. Rev. A **60**, 3618 (1999).

[245] D. Mukherjee and W. Kutzelnigg, J. Chem. Phys. **114**, 2047 (2001).

[246] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **116**, 4787 (2002).

[247] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **120**, 7340 (2004).

[248] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **120**, 7350 (2004).

[249] W. Kutzelnigg, J. Chem. Phys. **125**, 171101 (2006).

[250] A. C. Simmonett, J. J. Wilke, H. F. Schaefer III and W. Kutzelnigg, J. Chem. Phys. **133**, 174122 (2010).

[251] M. Kendall, A. Stuart and J. K. Ord. *Kendall's Advanced Theory of Statistics*. Charles Griffin, London, 5th edition, 1987.

[252] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer, Berlin, 2nd edition, 1985.

[253] R. Kubo, J. Phys. Soc. Jpn. **17**, 1100 (1962).

[254] D. A. Mazziotti, Chem. Phys. Lett. **289**, 419 (1998).

[255] D. A. Mazziotti, Int. J. Quantum Chem. **70**, 557 (1998).

[256] D. A. Mazziotti, Phys. Rev. A **60**, 4396 (1999).

[257] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik.* Springer, Berlin, 1932.

[258] R. C. Tolman. *The Principles of Statistical Mechanics.* Oxford University Press, Oxford, 1st edition, 1962.

[259] J. E. Harriman, Phys. Rev. A **65**, 052507 (2002).

[260] E. J. Thomas III, J. S. Murray, C. J. O'Connor and P. Politzer, J. Mol. Struct. (Theochem) **487**, 177 (1999).

[261] M. Brynda, L. Gagliardi and B. O. Roos, Chem. Phys. Lett. **471**, 1 (2009).

[262] K. Andersson, B. O. Roos, P.-Å. Malmqvist and P.-O. Widmark, Chem. Phys. Lett. **230**, 391 (1994).

[263] V. E. Bondybey and J. H. English, Chem. Phys. Lett. **94**, 443 (1983).

[264] S. M. Casey and D. G. Leopold, J. Phys. Chem. **97**, 816 (1993).

[265] B. Simard, M.-A. Lebeault-Dorget, A. Marijnissen and J. J. ter Meulen, J. Chem. Phys. **108**, 9668 (1998).

[266] M. Moskovits, W. Limm and T. Mejean, J. Chem. Phys. **82**, 4875 (1985).

[267] M. Moskovits, W. Limm and T. Mejean, J. Phys. Chem. **89**, 3886 (1985).

[268] K. Hilpert and K. Ruthardt, Ber. Bunsenges. Phys. Chem. **91**, 724 (1987).

[269] C.-X. Su, D. A. Hales and P. B. Armentrout, Chem. Phys. Lett. **201**, 199 (1993).

[270] G. A. Ozin and W. E. Klotzbücher, Inorg. Chem. **18**, 2101 (1979).

[271] M. M. Goodgame and W. A. Goddard III, Phys. Rev. Lett. **54**, 661 (1985).

[272] C. W. Bauschlicher Jr. and H. Partridge, Chem. Phys. Lett. **231**, 277 (1994).

[273] K. Andersson, Chem. Phys. Lett. **237**, 212 (1995).

[274] B. O. Roos, Collect. Czech. Chem. Commun. **68**, 265 (2003).

[275] P. Celani, H. Stoll, H.-J. Werner and P. Knowles, Mol. Phys. **102**, 2369 (2004).

[276] C. Angeli, B. Bories, A. Cavallini and R. Cimiraglia, J. Chem. Phys. **124**, 054108 (2006).

[277] T. Müller, J. Phys. Chem. A **113**, 12729 (2009).

[278] F. Ruipérez, F. Aquilante, J. M. Ugalde and I. Infante, J. Chem. Theory Comput. **7**, 1640 (2011).

[279] Y. Kurashige and T. Yanai, J. Chem. Phys. **135**, 094104 (2011).

[280] K. Hongo and R. Maezono, Int. J. Quantum Chem. **112**, 1243 (2012).

[281] P. K. Tamukong, D. Theis, Y. G. Khait and M. R. Hoffmann, J. Phys. Chem. A **116**, 4590 (2012).

[282] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[283] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[284] G. K.-L. Chan, M. Kállay and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).

[285] N. B. Balabanov and K. A. Peterson, J. Chem. Phys. **123**, 064107 (2005).

[286] P. K. Samanta, M. Hanauer, A. Köhn and D. Mukherjee, Manuscript in preparation.

[287] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).

[288] H.-J. Werner and M. Schütz, J. Chem. Phys. **135**, 144116 (2011).

[289] J. E. Subotnik, A. Sodt and M. Head-Gordon, J. Chem. Phys. **128**, 034103 (2008).

[290] F. Neese, F. Wennmohs and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).

[291] J. Friedrich, D. P. Tew, W. Klopper and M. Dolg, J. Chem. Phys. **132**, 164114 (2010).

[292] J. D. Goodpaster, T. A. Barnes, F. R. Manby and T. F. Miller III, J. Chem. Phys. **137**, 224113 (2012).

[293] Dalton, a molecular electronic structure program, release 2.0 (2005), see http://www.kjemi.uio.no/software/dalton/dalton.html.

[294] M. Kállay and P. R. Surján, J. Chem. Phys. **113**, 1359 (2000).

[295] J. Olsen, J. Chem. Phys. **113**, 7140 (2000).

[296] S. Hirata and R. J. Bartlett, Chem. Phys. Lett. **321**, 216 (2000).

[297] W. Duch, J. Phys. A: Math. Gen. **18**, 3283 (1985).

[298] J. Olsen, B. O. Roos, P. Jørgensen and H. J. Aa. Jensen, J. Chem. Phys. **89**, 2185 (1988).

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

[300] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[301] P. Pulay, J. Comput. Chem. **3**, 556 (1982).

[302] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955).