Charles University in Prague

Faculty of Science

DOCTORAL THESIS

Jiří Brabec

Parallel Implementation of Multireference Coupled Cluster Methods and Calculations on Large Systems

Department of Physical and Macromolecular Chemistry

Supervisor: Mgr. Jiří Pittner Dr. rer. Nat.

Study programme: Physical Chemistry

Prague, 2012

I would like to thank my supervisor Dr. Jiří Pittner for his kind support, patience and encouragement, Dr. Karol Kowalski for help and friendly support, my girlfriend Radka and family for home atmosphere.

Financial support by the Grant Agency of Charles University (GAUK 43.252494), Grant Agency of the Czech Republic (GACR Project No. 208/11/2222) and Extreme Scale Computing Initiative, a Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory is also appreciated. A large portion of the calculations have been performed using EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory. The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by the Battelle Memorial Institute under Contract DE-AC06.76RLO-1830.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In .... date .... signature of the author

Název práce: Paralelní implementace multireferenčních coupled cluster metod a výpočty na velkých systémech
Autor: Jiří Brabec
Katedra fyzikální a makromolekulární chemie
Vedoucí disertační práce: Mgr. Jiří Pittner Dr. rer. Nat., Ústav fyzikální chemie J. Heyrovského AVČR

Abstrakt: Práce je rozdělena do několika částí. Nejdříve jsme implementovali paralelní kód BW-MRCCSD metody, který je založený na automaticky generovaném jednoreferenčním CC kódu z Tensor Contraction Engine (TCE). Následně jsme navrhli nové schéma nezávislé dvouúrovňové paralelizace, na jejímž základě jsme implementovali BW-MRCCSD a Mk-MRCCSD metody. Nový algoritmus využívá tzv. procesorových skupin, které mohou vykonávat části úloh navzájem nezávisle. V první úrovni paralelizace je řešení pro jednotlivé reference rozděleno mezi procesorové skupiny (reference-level parallelism), v druhé úrovni jsou pak úlohy dané reference distribuovány mezi procesy v odpovídající procesorové skupině (task-level parallelism). Testy škálování byly provedeny až na 24000 jádrech pro β-karoten. Další část je věnována paralelním MRCC výpočtům, kde je zároveň demostrována využitelnost nové implementace na větších a chemicky zajímavých systémech. Poslední část práce se zabývá nově navrženou universal state-selective (USS) korekcí ke stavově specifickým MRCC metodám, její implementací a aplikací. Vlastnosti této korekce jsou ukázány na několika příkladech disociačních křivek a porovnány s FCI.

Title: Parallel Implementation of Multireference Coupled Cluster Methods and Calculations on Large Systems

Author: Jiří Brabec

Department: Department of Physical and Macromolecular Chemistry

Supervisor: Mgr. Jiří Pittner Dr. rer. Nat., J. Heyrovský Institute of Physical Chemistry of the ASCR

Abstract: Firstly, we have developed a Tensor Contraction Engine-based implementation of the BW-MRCCSD approach. The scalability tests have been performed across thousand of cores. We have further developed a novel two-level parallel algorithm for Hilbert-space MRCC methods which uses the processor groups. In this approach, references are distributed among processor groups (reference-level parallelism) and tasks of each reference are distributed inside of a given processor group (task-level parallelism). We have shown that our implementation scales across 24000 cores. The usability of our code was demonstrated on larger systems (dodecane, polycarbenes and naphthyne isomers). Finally, we present novel universal state-selective (USS) corrections to the state-specific MR-CC methods. The USS-corrected MRCC results were compared with the full configuration interaction (FCI) results.

Keywords: multireference, coupled, cluster, parallel

Contents

Introduction

Ab initio description of atoms, molecules, and chemical reactions to a high-accuracy level is a big challenge for quantum chemistry, since it can be a very computationaly demanding task, depending on the size of the system and method employed. First quantum chemical calculations were carried out about 90 years ago when no digital computers existed. Applications of first methods have been performed on smallest systems using just mechanical calculators or paper only. In 1937 the first binary relay-based calculator was invented by Stibitz and during the Second World War John von Neumann designed a new architecture which is used on most current computers. Development of quantum chemical codes went hand in hand with the growing power of available computers. Today, we have powerful parallel computing facilities with petaFLOPS performance formed by hundreds of thousands of processor cores, therefore a well parallelized code is indispensable. Parallel machines provide not only an enormous CPU power, but also a huge operating memory. This allows to keep all intermediates of a job in the fast RAM memory and to use the slow file system in network or on the hard disk just for storage of results or checkpoints. For example, the four-index integral transformation for the β-carotene molecule in the 6-31G basis set requires about 800 GB for intermediates, which substantially exceeds the RAM memory size of current non-parallel machines.

The dawn of peta-scale computer architectures brings a unique chance of validating and applying accurate yet numerically expensive many-body theories to large molecular systems. The multireference Hilbert-space coupled-cluster methods (MRCC) fall into this category. These methods provide means to treat properly static correlation effects for electronic states characterized by strong quasidegeneracy effects. Several problems such as bond-breaking processes, poly-radical species, transition metal compounds, low-spin open-shell states, and reactions involving crossing of potential energy surfaces require by their very nature a multireference level of description.[1]

Multireference coupled cluster formulations based on the Jeziorski-Monkhorst (JM) Ansatz[2], also known as Hilbert-space MRCC methods, can be divided into two major clases: State-Universal MRCC (SU-MRCC) approaches[2–11] and State-Specific (SS-MRCC) methods.[12–25] While the SU-MRCC approaches are based on the Bloch wave operator formalism,[26–29] which furnishes energies of M electronic states, where M is the dimension of the model space ($\mathcal{M}_0$) employed, the SS-MRCC methods, in order to avoid the so-called intruder-state problem,[30–32] were designed to target only one electronic state. However, the recent generalization of the SU-MRCC formalism to General Model Spaces (GMS-SU-MRCC) has shown to alleviate significantly the intruder state problem in the context of the SU-MRCC theory[33–35] and it thus seems that the potential of SU methods has not been exhausted yet.

The SU-MRCCSD implementation of Piecuch and Landman[36] was historically the first parallel implementation of a MRCC theory. Although tests were performed on a relatively small – by today's standards – number of cores, they demonstrated that a significant speedup of the SU-MRCCSD calculations on shared-memory parallel systems having 16-32 CPUs can be achieved. More recent implementation of the Mk-MRCC method[37] allowed to perform MRCC calculations for larger systems but still for rather small model spaces.

The aim of our work is to create an efficient parallel implementation of Hilbert-space MRCC methods for large-scale applications, which will be able to utilise current supercomputers, without limitations to a size of the model space and a number of orbitals.

This work is organized as follows. In Chapter I and II we shortly recall the basics of the multireference coupled cluster methods and parallel programing. In Chapter III and IV we provide details about our reference-based parallel implementation and performance tests. The Chapter V is about large-scale MRCC calculations on the bond-breaking in dodecane, naphthyne isomers and polycarbenes. In Chapter VI we discuss novel USS correction to the MRCC methods. Each of Chapter III, IV and VI has own results and the discussion part. The specific details about theory or computations are presented on the begining of each chapter.

Several chapters of this work are based on results published in the following papers:

Jiří Brabec, Hubertus J. J. van Dam, Karol Kowalski and Jiří Pittner, Chemical Physics Letters, (2011), 514, 347-351

Jiří Brabec, Jiří Pittner, Hubertus J. J. van Dam, Edoardo Aprà, and Karol Kowalski, Journal of Chemical Theory and Computation, (2012), 8, 487-497

Jiří Brabec, Hubertus J. J. van Dam, Jiří Pittner and Karol Kowalski, Journal of Chemical Physics (2012), 136, 124102

Parts of Chapter V have been submitted for publication to Chemical Physics Letters.

1. The Hilbert-space multireference coupled cluster methods

1.1 The single-reference CC theory

In general, a correlated wave function $|\Psi\rangle$ can be written as

$$|\Psi\rangle = \Omega|\Phi\rangle \ , \tag{1.1}$$

where $|\Phi\rangle$ is the reference wave fuction (typically the Hartree-Fock Slater determinant) and Ω is the wave operator. In the coupled cluster method the wave operator is assumed in an exponencial form

$$\Omega = e^{T} \ , \tag{1.2}$$

where T is the cluster operator. We impose the intermediate normalization

$$\langle\Psi|\Phi\rangle = 1 \tag{1.3}$$

while the Slater determinant $|\Phi\rangle$ is normalized too, since it is built from orthonormal spin-orbitals. The cluster operator T is a sum

$$T = T_1 + T_2 + T_3 + \dots \ , \tag{1.4}$$

where T_n is the linear combination of excitation operators, which corresponds to n-tuple excitations. Thus, for single- and double-excitations one can write

$$T_1 = \sum_{i,j} t_i^a a_a^\dagger a_i \tag{1.5}$$

$$T_2 = \sum_{i<j;a<b} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i \tag{1.6}$$

The coefficients $t_{ij..}^{ab..}$ are the cluster amplitudes. We use a standard notation, indices i, j denote occupied, a, b virtual, and p, q general molecular spin-orbitals. $a^\dagger$ and a are fermionic creation or anihilation operators which satisfy the anticommutation relations

$$\begin{aligned} \{a_p^\dagger, a_q\} &= a_p^\dagger a_q + a_q a_p^\dagger = \delta_{pq} \text{ and} \qquad (1.7) \\ \{a_p^\dagger, a_q^\dagger\} &= \{a_p, a_q\} = 0 \; . \qquad (1.8) \end{aligned}$$

The exponential ansatz ensures the size-extensivity of the CC method, but in its most frequently used formulation, the CC method is not variational.

Inserting the CC ansatz to Schrödinger equation and multiplying from the left by e^{-T} yields

$$e^{-T} H e^{T} \Phi = E \Phi \; . \qquad (1.9)$$

We can split the Hamiltonian into the Fermi vacuum expectation value and the normal ordered Hamiltonian H_N as follows

$$H = \langle \Phi | H | \Phi \rangle + H_\mathrm{N} = \langle \Phi | H | \Phi \rangle + \sum_{pq} F_{pq} N\{p^\dagger q\} + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle N\{p^\dagger q^\dagger s r\} \; , \qquad (1.10)$$

and using the Baker-Campbell-Hausdorff formula we will get

$$\begin{aligned} (H_\mathrm{N} e^T)_\mathrm{C} &\equiv e^{-T} H_N e^T = H_N + [H_N, T] + \frac{1}{2}[[H_N, T], T] + \\ &+ \frac{1}{3!}[[[H_N, T], T], T] + \frac{1}{4!}[[[[H_N, T], T], T], T] \; . \qquad (1.11) \end{aligned}$$

The CC equations are thus quartic in T amplitudes, since higher commutators vanish for a two body Hamiltonian (1.10). For the derivation of the correlation energy expression and amplitude equations we project the $e^{-T} H_N e^T |\Phi\rangle$ term to the bra vectors $\langle \Phi |$, $\langle \Phi_i^a |$, $\langle \Phi_{ij}^{ab} |$, etc. :

$$\begin{aligned} \Delta E_\mathrm{corr} &= \langle \Phi | H_N e^T | \Phi \rangle_\mathrm{C} \; , \qquad (1.12) \\ \langle \Phi_{ij\ldots}^{ab\ldots} | H_N e^T | \Phi \rangle_\mathrm{C} &= 0 \qquad (1.13) \end{aligned}$$

where $\langle \Phi_i^a |, \langle \Phi_{ij}^{ab} |, \ldots$ represent single-, double- or higher excitations from $\langle \Phi |$. The notation $(H_\mathrm{N} e^T)_\mathrm{C}$ indicates that the commutator expansion (1.11) corresponds to connected terms in the diagramatic representation.

1.2 Multireference pertubation theory

In the pertubation theory the Hamiltonian is splitted to an unperturbed part H_0 and a perturbation V

$$H = H_0 + V \; , \qquad (1.14)$$

where H_0 has eigenvectors $|\Phi_\alpha\rangle$. Eigenvectors of the perturbed Hamiltonian H are denoted as $|\Psi_\alpha\rangle$ and its eigenvalues E_α.

By $\mathcal{M}_0$ we denote the M-dimensional reference space spanned by $|\Phi_\mu\rangle$, $\mathcal{M}_0^\perp$ its orthogonal complement, and corresponding projection operators P and Q $(P + Q = 1)$. The projection of the α-eigenfunction of H to $\mathcal{M}_0$ is

$$\Psi_\alpha^P \equiv P\Psi_\alpha = \sum_{\mu=1}^{M} c_\alpha^\mu \Phi_\mu \ . \tag{1.15}$$

If $\mathcal{M}_0$ is a complete model space (CMS), i.e. it contains all possible Slater determinants for the selected number of active spin-orbitals and electrons, we can use the intermediate normalization $\langle\Psi_\alpha|\Psi_\alpha^P\rangle = 1$. The wave operator Ω_α builds Ψ_α back from the projection Ψ_α:

$$\Psi_\alpha = \Omega_\alpha \Psi_\alpha^P \tag{1.16}$$

The exact energy E_α is an eigenvalue of the effective Hamiltonian H_α^{eff} and if the model space is complete, we can write

$$H_\alpha^{\text{eff}} = PH\Omega_\alpha P \ , \tag{1.17}$$

and

$$H_\alpha^{\text{eff}} \Psi_\alpha^P = E_\alpha \Psi_\alpha^P \ . \tag{1.18}$$

In state-specific methods only the αth-eigenvalue of the Hamiltonian has a physical meaning. To obtain the exact wave function, we need to determine the wave operator. For this purpose the Bloch equation for Ω was introduced. For the Rayleigh-Schrödinger perturbation theory the Bloch equation reads[38]

$$[\Omega, H_0]P = V\Omega P - \Omega P V \Omega P \tag{1.19}$$

For the state-specific BW theory Hubač derived the form[12]

$$\Omega_\alpha = 1 + B_\alpha V \Omega_\alpha \ , \tag{1.20}$$

where B_α is the BW resolvent

$$B_\alpha = \frac{Q}{E_\alpha - H_0} \tag{1.21}$$

The parametrized Bloch equation for continuous transition from RSPT to BWPT can be written as[39]

$$\begin{aligned}
&\lambda E_\alpha \Omega_\alpha P_\alpha + (1-\lambda)[\Omega_\alpha, H_0]P \\
&= \lambda H \Omega_\alpha P_\alpha + (1-\lambda)[V\Omega_\alpha - \Omega_\alpha P V \Omega_\alpha]P\,,
\end{aligned} \tag{1.22}$$

For $\lambda = 1$, the equation corresponds to the Brillouin-Wigner perturbation theory, whereas $\lambda = 0$ leads to the Rayleigh-Schrödinger theory and the state-universal method.

1.3 Multireference coupled cluster theory

In the Hilbert space multireference coupled cluster methods the wave operator is taken in the form of the Jeziorski-Monkhorst (JM) Ansatz[2]

$$|\Psi_\alpha\rangle = \sum_{\mu=1}^{M} c_\mu^\alpha e^{T^{(\mu)}} |\Phi_\mu\rangle\ . \tag{1.23}$$

The operators $T^{(\mu)}$ $(\mu = 1, \ldots, M)$ are the reference-specific cluster operators, which produce a linear combination of excited configurations from the orthogonal complement of thje model space $\mathcal{M}_0^\perp$ when acting on the corresponding reference function $|\Phi_\mu\rangle$, $(\mu = 1, \ldots, M)$. Further, c_μ^α $(\mu = 1, \ldots, M)$ coefficients are components of the right eigenvector of the effective Hamiltonian H^{eff}, whose matrix elements are defined for the CMS as

$$H_{\nu\mu}^{\text{eff}} = \langle \Phi_\nu | H e^{T^{(\mu)}} | \Phi_\mu \rangle_{\text{C}} \tag{1.24}$$

where the subscript C designates a connected part of the expression. Using the normal ordered form of the Hamiltonian (1.10) with respect to the Fermi vacuum $|\Phi_\mu\rangle$

$$H = H_{\mu\mu} + H_N(\mu) \tag{1.25}$$

the matrix elements can be rewritten as

$$H_{\nu\mu}^{\text{eff}} = H_{\mu\mu}\delta_{\mu\nu} + \langle \Phi_\nu | H_N(\mu) e^{T^{(\mu)}} | \Phi_\mu \rangle_{\text{C}} \tag{1.26}$$

In this work, we will entirely focus on complete model spaces. To maintain the intermediate normalization (1.3), all "internal" amplitudes, i.e. amplitudes which transform one reference $|\Phi_\mu\rangle$ to another $|\Phi_\nu\rangle$, $\mu, \nu \in \mathcal{M}_0$, are set to zero.

1.3.1 Brillouin-Wigner MRCC method

The multireference Brillouin-Wigner CC method[12,13,40–43], introduced by Hubač *et. al.*[12] and further developed in various directions by Pittner et. al.[11,14,15,44–48], was the first state-specific multireference coupled cluster method proposed. Compared to the SUCC, this approach is insensitive to intruder states, its cost scales only linearly with the number of references, and the amplitude equations have a very simple structure as they are coupled only through the exact energy. However, MR BWCC lacks the important advantage of single reference CC theory, the size-extensivity, due to the presence of unlinked terms in the cluster equations. To circumvent this problem, an approximate a posteriori size-extensivity correction[44] and a more rigorous iterative scheme[11,15] have been developed and tested on a variety of diatomics[49–54], and successfully applied to larger chemically interesting systems[55–62].

The equations for the cluster amplitudes can be derived from the generalized Bloch equation (1.22). After insertion of the JM Ansatz into (1.22), the resulting equation is applied to the state $|\Psi_\alpha\rangle$. Since for $\lambda \neq 0$ the method is state-specific, the sufficiency conditions are applied, splitting the equation system to individual equations sets for each reference $|\Phi_\mu\rangle$, in a consistent way with the $\lambda = 0$ limit. Division by c_μ^α and projection onto the Slater determinants $|\Phi_\vartheta^{(\mu)}\rangle$ excited with respect to $|\Phi_\mu\rangle$ yields the cluster equations from which the amplitudes of $T(\mu)$ are determined:

$$\begin{aligned}\lambda(\mathcal{E}_\alpha - H^{\text{eff}}_{\mu\mu})\langle\Phi_\vartheta^{(\mu)}|e^{T(\mu)}|\Phi_\mu\rangle &= \langle\Phi_\vartheta^{(\mu)}|[H_N(\mu)e^{T(\mu)}]_{\text{C}}|\Phi_\mu\rangle \\ &+ \langle\Phi_\vartheta^{(\mu)}|[H_N(\mu)e^{T(\mu)}]_{\text{DC,L}}|\Phi_\mu\rangle \\ &- (1-\lambda)\sum_{\nu\neq\mu}\langle\Phi_\vartheta^{(\mu)}|e^{T(\nu)}|\Phi_\nu\rangle H^{\text{eff}}_{\nu\mu}\,. \end{aligned} \tag{1.27}$$

Concerning evaluation of H^{eff} matrix elements, the expression for the diagonal elements $H^{\text{eff}}_{\mu\mu}$ is the same as for the CC energy in the single reference theory, when the amplitudes of the μ-th reference are used. The off-diagonal elements of the H^{eff} can be viewed as the connected parts of the right hand sides of the cluster equations corresponding to the internal amplitude which transforms the μ-th reference to the ν-th reference. As long as the model space is restricted to mutually mono- and biexcited references, the off-diagonal elements of H^{eff} can be collected from the vector of the respective right hand sides at the CCSD level.

The correlation energy is obtained by diagonalization of the H^{eff} matrix, while the corresponding eigenvector provides the c_μ^α coefficients. Since the theory is state-specific, only the selected eigenvalue, for which we converge the equations, has a physical meaning.

For the purpose of obtaining BWCC size-extensivity corrections, the (1-λ) -scaled coupling terms are neglected and the disconnected linked (DC,L) terms are scaled by λ, giving rise to:

$$\begin{aligned}\lambda(\mathcal{E}_\alpha - H^{\text{eff}}_{\mu\mu})\langle\Phi^{(\mu)}_\vartheta|e^{T(\mu)}|\Phi_\mu\rangle &= \langle\Phi^{(\mu)}_\vartheta|[H_N(\mu)e^{T(\mu)}]_{\text{C}}|\Phi_\mu\rangle \\ &+ \lambda\langle\Phi^{(\mu)}_\vartheta|[H_N(\mu)e^{T(\mu)}]_{\text{DC,L}}|\Phi_\mu\rangle\,. \end{aligned}\tag{1.28}$$

The usually employed a posteriori size-extensivity correction[44] consists in solving this equation iteratively with $\lambda = 1$ until convergence is achieved and subsequently making one additional iteration with $\lambda = 0$, which yields corrected amplitudes, from which the corrected H^{eff} is computed and its diagonalization finally yields the corrected energy.

1.3.2 Mukherjee's MRCC method

The original derivation of the Mk MRCC method[16] is based on the Schrödinger equation and JM Ansatz (1.23).

$$\sum_\mu He^{T(\mu)}|\Phi_\mu\rangle c_\mu = E\sum_\mu e^{T(\mu)}|\Phi_\mu\rangle c_\mu \tag{1.29}$$

Inserting the resolution of identity in the form

$$I = e^{T(\mu)}(P+Q)e^{-T(\mu)} = \sum_\nu e^{T(\mu)}|\Phi_\nu\rangle\langle\Phi_\nu|e^{-T(\mu)} + e^{T(\mu)}Qe^{-T(\mu)} \tag{1.30}$$

to the l.h.s. of (1.29) we obtain

$$\sum_\mu\sum_\nu e^{T(\mu)}|\Phi_\nu\rangle H^{\text{eff}}_{\nu\mu}c^\alpha_\mu + \sum_\mu e^{T(\mu)}Qe^{-T(\mu)}He^{T(\mu)}|\Phi_\mu\rangle c^\alpha_\mu = E\sum_\mu e^{T(\mu)}|\Phi_\mu\rangle c^\alpha_\mu \tag{1.31}$$

Next, the indicies μ and ν are interchanged in the first term on the left-hand side of (1.31)

$$\sum_\mu\sum_\nu e^{T(\nu)}|\Phi_\mu\rangle H^{\text{eff}}_{\nu\mu}c^\alpha_\nu + \sum_\mu e^{T(\mu)}Qe^{-T(\mu)}He^{T(\mu)}|\Phi_\mu\rangle c^\alpha_\mu = E\sum_\mu e^{T(\mu)}|\Phi_\mu\rangle c^\alpha_\mu \tag{1.32}$$

Sufficiency conditions are again introduced as requirement that each term of the summation in (1.32) vanishes individually. Multiplying the resulting equation by $e^{-T(\mu)}$ from the left and projecting to $\langle\Phi_{\vartheta}^{(\mu)}|$ yields working equations for determination of cluster amplitudes

$$\langle\Phi_{\vartheta}^{(\mu)}|e^{-T(\mu)}He^{T(\mu)}|\Phi_\mu\rangle c_\mu^\alpha + \sum_{\nu\neq\mu}\langle\Phi_{\vartheta}^{(\mu)}|e^{-T(\mu)}e^{T(\nu)}|\Phi_\mu\rangle H_{\mu\nu}^{\text{eff}}c_\nu^\alpha = 0 \tag{1.33}$$

The explicit form of the coupling terms in the Mk-MRCCSD approach has been derived by Evangelista *et al.*[19]:

$$\langle\Phi_i^a(\mu)|e^{-T^{(\mu)}}e^{T^{(\nu)}}|\Phi_\mu\rangle = t_a^i(\nu/\mu) - t_a^i(\mu)\ , \tag{1.34}$$

$$\begin{aligned}\langle\Phi_{ij}^{ab}(\mu)|e^{-T^{(\mu)}}e^{T^{(\nu)}}|\Phi_\mu\rangle &= -t_{ab}^{ij}(\mu) + P(ij)t_a^i(\mu)t_b^j(\mu) - P(ij)P(ab)t_a^i(\mu)t_b^j(\nu/\mu)\\ &\quad + t_{ab}^{ij}(\nu/\mu) + P(ij)t_a^i(\nu/\mu)t_b^j(\nu/\mu)\ ,\end{aligned} \tag{1.35}$$

where we adopted the notation of Ref.[19] with the permutation operator $P(ij)$ defined as $P(ij)A^{ij} = A^{ij} - A^{ji}$ and the $t_a^i(\nu/\mu)$ and $t_{ab}^{ij}(\nu/\mu)$ amplitudes referring to the so-called "common amplitudes" of operator $T^{(\nu)}$ with respect to vacuum $|\Phi_\mu\rangle$[19], thus

$$t_{i\ldots}^{a\ldots}(\nu/\mu) = \begin{cases} t_{i'\ldots}^{a'\ldots}(\nu) & \text{if} i,\ldots\in \text{occ}(\mu); a,\ldots\in\text{virt}(\mu)\\ 0 & \text{otherwise.}\end{cases}$$

and primed indices indicate renumbering the spin-orbitals with respect to $|\Phi_\mu\rangle$. In the original Mukherjee's paper[16] the evaluation of new amplitudes is suggested in a form where the c_μ^α coefficient occurs in a denominator

$$t(\text{new}) = \frac{\langle\Phi_{i\ldots}^{a\ldots}(\mu)|V_{\text{N}}(\mu)e^{T^{(\mu)}}|\Phi_\mu\rangle_{\text{C}} + \sum_{\mu\neq\nu}\{t_{i\ldots}^{a\ldots}(\mu) + \langle\Phi_{i\ldots}^{a\ldots}(\mu)|e^{-T^{(\mu)}}e^{T^{(\nu)}}|\Phi_\mu\rangle\}H_{\mu\nu}^{\text{eff}}\frac{c_\nu^\alpha}{c_\mu^\alpha}}{D_{i\ldots}^{a\ldots}(\mu) + E_\alpha - H_{\mu\nu}^{\text{eff}}}\ , \tag{1.36}$$

where $D_{i\ldots}^{a\ldots}(\mu) = F_{ii}^\mu + \ldots - F_{aa}^\mu - \ldots$ denotes the Rayleigh-Schrödinger denominator and F_{pq}^μ are Fock matrix elements with respect to vacuum $|\Phi_\mu\rangle$. When several references have a small weight (i.e. c_μ^α are near zero), this approach may cause numerical problems. Evangelista *at al.* proposed[19] another approach for amplitude iteration,

$$t(\text{new}) = t(\text{old}) + \frac{\langle\Phi_{i\ldots}^{a\ldots}(\mu)|e^{-T^{(\mu)}}He^{T^{(\mu)}}|\Phi_\mu\rangle c_\mu^\alpha + \sum_{\mu\neq\nu}\langle\Phi_{i\ldots}^{a\ldots}(\mu)|e^{-T^{(\mu)}}e^{T^{(\nu)}}|\Phi_\mu\rangle H_{\mu\nu}^{\text{eff}}c_\nu^\alpha}{D_{i\ldots}^{a\ldots}(\mu) + E_\alpha - H_{\mu\nu}^{\text{eff}}} \tag{1.37}$$

which avoids the division by c^{α}_{μ}. Our implementation of the Mk-MRCC method is based on Eq.(1.37).

1.4 General formulation for Hilbert-space MR-CC amplitude equations

In the SU- as well as the SS-MRCCSD case, the amplitude equations can be generally written as an aggregate of reference-specific equations, i.e.,

$$\mathbf{R}^{(\mu)} = \mathbf{F}^{(\mu)}(T^{(\mu)}) + \mathbf{G}^{(\mu)}(T^{(1)}, \ldots, T^{(\mu)}, \ldots, T^{(M)}) = \mathbf{0} \;\; \forall_{\mu=1,\ldots,M} \tag{1.38}$$

where the vectors $\mathbf{F}^{(\mu)}$ and $\mathbf{G}^{(\mu)}$ refer to direct and coupling terms, respectively. For all MRCCSD theories discussed here, the components $\mathrm{F}^{(\mu)}(\vartheta)$ of the direct term $\mathbf{F}^{(\mu)}$ are equal to

$$\mathrm{F}^{(\mu)}_{\vartheta}(T^{(\mu)}) = \langle \Phi_{\vartheta}(\mu)|(He^{T^{(\mu)}})_C|\Phi_{\mu}\rangle \tag{1.39}$$

where $\langle \Phi_{\vartheta}(\mu)|$ corresponds to singly or doubly excited (with respect to $|\Phi_{\mu}\rangle$) configurations from $\mathcal{M}_0^{\perp}$ ($\langle \Phi_i^a(\mu)|$, $\langle \Phi_{ij}^{ab}(\mu)|$, ...). The algebraic form of the coupling terms is method-specific. While the direct term depends only on the cluster operator corresponding to a given reference, the coupling terms may involve all possible cluster operators. When the notational convention employed for the direct term is used, the coupling terms can be expressed as follows:

- SU-MRCCSD method:

$$\mathrm{G}^{(\mu)}_{\vartheta}(T^{(1)}, \ldots, T^{(\mu)}, \ldots, T^{(M)}) = \sum_{\substack{\nu=1 \\ \nu\neq\mu}}^{M} \langle \Phi_{\vartheta}(\mu)|e^{-T^{(\mu)}}e^{T^{(\nu)}}|\Phi_{\nu}\rangle H^{\mathrm{eff}}_{\nu\mu} \tag{1.40}$$

- Mk-MRCCSD method:

$$\mathrm{G}^{(\mu)}_{\vartheta}(T^{(1)}, \ldots, T^{(\mu)}, \ldots, T^{(M)}) = \sum_{\substack{\nu=1 \\ \nu\neq\mu}}^{M} \langle \Phi_{\vartheta}(\mu)|e^{-T^{(\mu)}}e^{T^{(\nu)}}|\Phi_{\mu}\rangle H^{\mathrm{eff}}_{\mu\nu}\frac{c_{\nu}}{c_{\mu}} \tag{1.41}$$

• BW-MRCCSD method:

$$\begin{aligned} &\mathrm{G}_{\vartheta}^{(\mu)}(T^{(1)},\ldots,T^{(\mu)},\ldots,T^{(M)}) = \\ &\langle\Phi_{\vartheta}(\mu)|(H_{\mathrm{N}}(\mu)e^{T^{(\mu)}})_{DC,L}|\Phi_{\mu}\rangle - \Delta E_{\mathrm{corr}}\langle\Phi_{\vartheta}(\mu)|e^{T^{(\mu)}}|\Phi_{\mu}\rangle \end{aligned} \tag{1.42}$$

In Eq.(1.42), ΔE_{corr} corresponds to the eigenvalue of the effective Hamiltonian, whereas 'DC,L' subscript in (1.42) refers to disconnected linked parts of a given operator expression.

The diagrammatic form of the SU-MRCC coupling terms (1.40) has been extensively discussed in the context of the general model space formulation of the SU-MRCC theory.[63] The simplest form of the coupling term is employed by the BW-MRCCSD approach, where the only quantity coupling the equations for various references is the energy E_{α}. For the BW-MRCCSD and Mk-MRCCSD approaches, the numerical effort associated with forming residual vectors $\mathbf{R}^{(\mu)}$ is dominated by calculating the direct terms $\mathbf{F}^{(\mu)}(T^{(\mu)})$ with corresponding N^6 numerical complexity per each reference $|\Phi_{\mu}\rangle$, where N symbolically designates system size. Given a large numerical overhead, proportional to $M \times N^6$, efficient parallel codes are needed in order to make the MRCCSD theories applicable to the large systems.

2. Parallel computing

2.1 Hardware

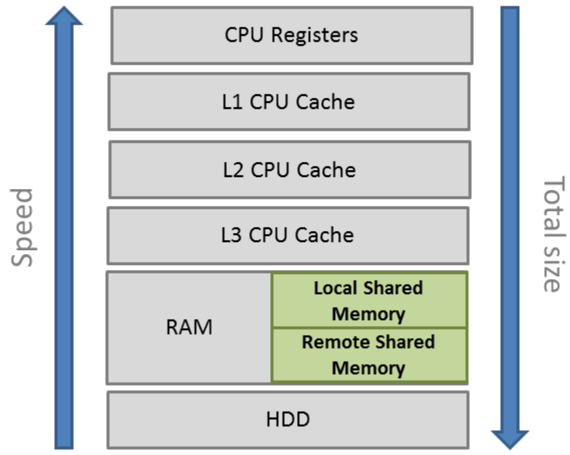

Figure 2.1: Memory hierarchy in a parallel cluster

Each supercomputer consists of individual computing units, so-called nodes. The node has one or more multi-core CPU's, a computing memory (RAM) and possibly a disc. All nodes are interconnected by a network interface. The fundamental problem in the parallel computing is the data distribution and handling. A fast interconnection between two arbitrary nodes is essential to achieve a uniformly high CPU load across the whole supercomputer. Fig.2.1 shows a simplified memory architecture of a parallel computing cluster. In comparison with a single workstation, the computing memory in each node splits to the local and remote memory. The local memory is used for smaller data blocks which are are needed only by the computational procedure on the given node, while the remote memory is used for large shared data files distributed among nodes. The remote memory has unified adress space and the data inside are visible to all nodes or to a given subset of nodes, respectively. The access speed is given by parameters of the network interface. In general, this speed is significantly lower then for the local memory. The network interconnection can be characterized by several parameters: Network bandwidth, latency and topology. The latency t_l can be defined as a startup time required for a message sent from point A to point B. The bandwidth ω is the data transfer rate usualy measured in bytes/s. The total time t required to send one data packet of length l can be expressed as

$$t = t_l + l/\omega \, . \tag{2.1}$$

For small packets the dominant contribution is the latency, while for large packets the determining parameter is the bandwidth.

The network topology can be a simple ring, point-to-point, mesh, tree, bus, star, line or a hybrid topology formed by a combination of two or more standard topologies. The topology of large supercomputers may be much more complicated and may have several subdomains in more levels. The final choice of the network topology in the supercomputer is usualy a compromise between the cost and performance, thus programmers must routinely deal with the cost of the data transfer. An additional discussion about network topologies is beyond the scope of this work.

2.2 Software

In the parallel programming it is neccessary to provide means the data tranfer between individual threads of execution running one the same node or different nodes. When the threads are located on the same node, the simplest way is to use a shared memory. In the latter case, the transport mechanism is provided by special libraries. The most common paradigm is the message passing, where one thread calls the send routine and the other thread calls a matching receive routine. This simple algorithm is used by libraries such as current MPI or TCGMSG, but it has an important limitation. Both sending and receiving processes have to participate in message transfer and be synchronized, thus it can lead to load balancing problems, especially when the demands of different tasks differ significantly.
This issue can be solved by virtual shared memory. In this paradigm, any process can access the remote memory without a direct assistance of the remote process on the node where the data are physically located. An implementation of this approach is provided by libraries such as Global Arrays (GA), Linda, or MPI-2.

The NWChem program uses the GA toolkit. GA forms an abstraction layer that alleviates the task of the developer by isolating most of the complexities involved in parallelizing software that make use of dense matrices. It relies on several components: a message passing library, the ARMCI one-sided communication library, and a dynamic memory allocator (MA library, based on the C language interfaced for the Fortran 77/90 applications). Most of the NWChem modules (including the ones used for this work) make very little use of MPI, since the majority of the communication is managed by the ARMCI one-sided communication library.

2.3 Tensor Contraction Engine

The Tensor Contraction Engine (TCE) software[64] is an important part of the NWChem program. The basic idea of the TCE is, that any method which can be formulated as a series of tensor contractions, can be implemented automatically starting from the algebraic expressions and ending by translation of factorized working equations to Fortran code.[65] Our MRCC implementation was developed by adapting TCE generated implementation of the single reference CC method. To understand the details of the paralell MRCCSD implementations, several general remarks on the structure of TCE generated codes are necessary. In a code generated by TCE, the domains of spin-orbital indices are partitioned into smaller subsets containing the spin-orbitals of the same spatial and spin symmetries. This partitioning into subsets (henceforth called tiles) leads to a block-structure of all tensors representing one- and two-electron integrals, as well as cluster amplitudes and recursive intermediates. In single-reference theories, occupied (unoccupied) tiles are designated as `[I]`,`[J]`,`[K]`,... (`[A]`,`[B]`,`[C]`,...). In the multireference case, we will use separate tiling schemes for each reference, which will significantly improve the performance of the MRCCSD implementation. In order to distinguish between tiling schemes corresponding to various references, we will use the notation $\texttt{[I(}\mu\texttt{)]}$,$\texttt{[J(}\mu\texttt{)]}$,$\texttt{[K(}\mu\texttt{)]}\ldots$ and $\texttt{[A(}\mu\texttt{)]}$,$\texttt{[B(}\mu\texttt{)]}$,$\texttt{[C(}\mu\texttt{)]}\ldots$ for occupied and unoccupied tiles in the reference $|\Phi_\mu\rangle$. For example, the singly and doubly excited cluster amplitudes corresponding to this reference are stored in memory using the following block structure

$$\texttt{t}^{\texttt{[I(}\mu\texttt{)]}}_{\texttt{[A(}\mu\texttt{)]}}(\mu) \equiv \{t^i_a(\mu) \;:\; i \in \texttt{[I(}\mu\texttt{)]}, a \in \texttt{[A(}\mu\texttt{)]}\} \tag{2.2}$$

$$\begin{aligned} \texttt{t}^{\texttt{[I(}\mu\texttt{)]}\texttt{[J(}\mu\texttt{)]}}_{\texttt{[A(}\mu\texttt{)]}\texttt{[B(}\mu\texttt{)]}}(\mu) \equiv \{&t^{ij}_{ab}(\mu) \;:\; i \in \texttt{[I(}\mu\texttt{)]}, j \in \texttt{[J(}\mu\texttt{)]}, \\ &a \in \texttt{[A(}\mu\texttt{)]}, b \in \texttt{[B(}\mu\texttt{)]}\} \end{aligned} \tag{2.3}$$

where the occupied i, j and unoccupied a, b indices (in the sense of reference $\Phi_\mu\rangle$) belong to tiles $\texttt{[I(}\mu\texttt{)]}$,$\texttt{[J(}\mu\texttt{)]}$,$\texttt{[A(}\mu\texttt{)]}$,$\texttt{[B(}\mu\texttt{)]}$, respectively, and $t^i_a(\mu)$ and $t^{ij}_{ab}(\mu)$ are the antisymmetric cluster amplitudes defining $T_1^{(\mu)}$ and $T_2^{(\mu)}$ operators. Amplitudes from subsets (2.2) and (2.3) corresponding to excitations within the model space are set equal to zero. The other tensors corresponding to the integrals and intermediates are represented in a similar way. This data representation provides the necessary granularity ideally suitable for dynamic load balancing.

For example, the parallel calculation of the so-called 4-particle diagram

$$v_{ab}^{ef} t_{ef}^{ij}(\mu) \ \ (i < j \ \ \ a < b \ , \ e < f) \tag{2.4}$$

is performed using the following scheme, where `v` (2-electron integrals) and `t(`μ`)` tensors are assumed to be distributed on GA across all cores used in a given run

```
[A(μ)]=noab(μ)+1,noab(μ)+nvab(μ)
[B(μ)]=[A(μ)],noab(μ)+nvab(μ)
[I(μ)]=1,noab(μ)
[J(μ)]=[I(μ)],noab(μ)
dynamic load balancing:  task is assigned to the first available process
create local memory region r^{[I(μ)][J(μ)]}_{[A(μ)][B(μ)]}
 [E]=noab(μ)+1,noab(μ)+nvab(μ)
 [F]=[E],noab(μ)+nvab(μ)
  ga_get( t^{[I(μ)][J(μ)]}_{[E(μ)][F(μ)]} (μ))
  ga_get( v^{[E(μ)][F(μ)]}_{[A(μ)][B(μ)]} )
  r^{[I(μ)][J(μ)]}_{[A(μ)][B(μ)]} :=+ dgemm(  v^{[E(μ)][F(μ)]}_{[A(μ)][B(μ)]} t^{[I(μ)][J(μ)]}_{[E(μ)][F(μ)]} (μ) )
  enddo
  enddo
ga_acc( r^{[I(μ)][J(μ)]}_{[A(μ)][B(μ)]} ) → MRCCSD residual vector for μ-th reference
                          (or appropriate elements of the effective Hamiltonian)
delete local memory region r^{[I(μ)][J(μ)]}_{[A(μ)][B(μ)]}
end of load balance assigned task
enddo
enddo
enddo
enddo
```

where `ga_get()` and `ga_acc()` are typical one-sided GA calls which fetch from and accumulate to the corresponding global array. Other parts of the code make use also of `ga_put()` calls. The `noab(`μ`)` and `nvab(`μ`)` refer to the total number of occupied and unoccupied tiles in reference $|\Phi_\mu\rangle$. The accumulation to the effective Hamiltonian occurs when excitations from `[I(`μ`)]`, `[J(`μ`)]`, `[A(`μ`)]`, `[B(`μ`)]` tiles correspond to the internal excitations (i.e. within the model space). If the maximum size of tiles is defined by a parameter `tilesize`, then the maximum amount of communication required to fetch the `v` and `t` blocks is proportional to $2\times$`tilesize`^{4}, which in typical situation `tilesize`=30 is equivalent to 12.4 MB. This communication upper bound is also valid for other diagrams in the direct and coupling terms.

3. Parallel implementation of the BW-MRCCSD method

This chapter has been published in Journal of Chemical Theory and Computation: Jiří Brabec, Hubertus J. J. van Dam, Karol Kowalski and Jiří Pittner, Chemical Physics Letters, (2011), 514, 347-351

3.1 Parallel implementation

Our first MR BWCCSD parallel implementation based on the task-level parallelism was developed by adapting the TCE generated implementation of the single reference CCSD method to the MR case. The DC,L and UL terms have also been implemented using the TCE functionality. The spinorbital character of the MR BWCCSD implementations makes it applicable to any type of references, including ones based on restricted, restricted open-shell, and unrestricted Hartree Fock spinorbitals (RHF, ROHF, and UHF respectively). In analogy to the single reference case, the iterative MR BWCCSD approach involves many intermediate steps of various numerical complexity which are conglomerates of the `ga_get()`, `ga_put()`, and `ga_acc()` calls together with BLAS-level 3 `dgemm` calls.

The data representation in the CC code is based on the partitioning of the entire spinorbital domain into smaller subsets (tiles) containing spinorbitals corresponding to the same spin and spatial symmetries. Each reference uses its own partitioning of the spinorbital domain. The occupied and unoccupied tiles with respect to a given reference $|\Phi_\mu\rangle$ are designated as $[i(\mu)]$,$[j(\mu)]$,$[k(\mu)]$,... and $[a(\mu)]$,$[b(\mu)]$, $[c(\mu)]$,..., respectively. This partitioning entails the partitioning of all tensors involved in the MR BWCCSD calculations including cluster amplitudes, recursive intermediates, and integrals. For example, the tensor corresponding to doubly excited amplitudes with respect to a given reference $|\Phi_\mu\rangle$is stored in the block form defined by smaller 4-dimensional tensors

$$t^{[i(\mu)][j(\mu)]}_{[a(\mu)][b(\mu)]}(\mu) \tag{3.1}$$

representing a subset of doubly excited amplitudes defined by the indices belonging to the $[i(\mu)]$,$[j(\mu)]$,$[a(\mu)]$,and $[b(\mu)]$ tiles. As a consequence of the intermediate normalization, the internal excitations are excluded from the $T_1(\mu)$ and $T_2(\mu)$ operators. These excitations, however, are included in the vector storing the direct term $\langle\Phi_\vartheta(\mu)|[H_N(\mu)e^{T(\mu)}]_{\mathrm{C}}|\Phi_\mu\rangle$, which gives us access to the elements of the MR BWCCSD effective Hamiltonian (1.26) . Currently, only mono and biexcited elements of the effective Hamiltonian operator are implemented.

The block structure of the MR BWCCSD tensors also defines the granularity of the code. Parallelization of the code, based on dynamic load balancing, occurs over the `do-loops` over occupied/unoccupied tiles. The units of parallel work are said to constitute a task pool – a collection of tasks that can be executed in parallel. The scalability of each section of the calculation is directly proportional to the size of its task pool. The `tilesize` parameter can be used to define the maximum size of tiles and to tune the granularity to given architecture specifications. The `tilesize` also defines the efficacy of `dgemm` calls and local memory requirements. For the iterative MR BWCCSD the local memory demand is proportional to $(\texttt{tilesize})^4$. The MR BWCCSD code is an aggregate of a large number of (memory- and computation-wise) highly diversified subroutines, which can lead to pronounced problems with the load balancing. For example, for the most computationally intensive part of the MR BWCCSD approach, associated with the inclusion of 4-index integrals, the corresponding task pool can easily be two orders of magnitude larger than the task pool for subroutines calculating one-body intermediates (additionally, all tiles corresponding to active spinorbitals have a size 1, which can also lead to load imbalance issues). Overall, the iterative calculation consists of numerous such task pools. Executing one task pool at a time brings out these disproportions and leads to inefficient utilization of computational resources. This also leads to a large number of synchronization steps and associated load imbalance.

The tensors produced and consumed by the task pools generate producer-consumer relationships across the task pool. We explored the dependencies exposed between the tasks and grouped them into classes that can be characterized as a collective task pools. This was done for the most expensive direct term in such a way as to ensure sufficient parallelism in each class while minimizing the total number of such classes. This procedure enabled us to reduce the number of synchronization steps. In contrast to the original implementation (referred to as the MR BWCCSD(A1) algorithm) the implementation based on the improved task scheduling will be referred to as MR BWCCSD(A2) algorithm.

3.2 Computational details

For the performance test of our implementation we have chosen the 1,7-naphthyne isomer (Figure 3.1. In general, naphthyne isomers are very interesting systems from the viewpoint of MRCC methods, having two σ-radical centers at different atoms. Depending on the mutual distance of these radical centers, the strength of the quasi-degeneracy varies and should grow with the distance. The 1,7 isomer belongs to those which have strong multireference character[66].

The four most important configurations have been used as references — the Hartree-Fock determinant and the HOMO $\rightarrow$ LUMO mono- and doubly-excited configurations from the CAS(2,2) space. Since HOMO and LUMO exhibit the same spatial symmetry, two references are not sufficient.

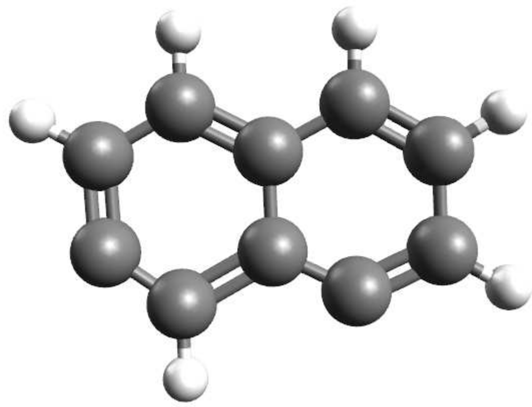

Figure 3.1: 1,7-naphthyne isomer

In our tests we used various size basis sets: cc-pVDZ[67], 6-311G**[68], Roos aug-DZ ANO (Roos aDZ ANO)[69], and cc-pVTZ[67] consisting of 170, 216, 284, and 384 basis functions, respectively, when spherical representation of d (f) functions is used, and 180, 226, 304, 440 basis function, when cartesian representation is involved. We will use notation X(s) and X(c) for basis set X in the former and latter case. In all calculations all electrons were correlated and RHF orbitals were used. The scalability tests were performed on the Chinook system at PNNL consisting of 2310 HP DL185 nodes with dual socket, 64-bit, Quad-core AMD 2.2 GHz Opteron processors. Each node has 32 GB of memory, i.e. 4 GB per core. The nodes are connected with a single rail Infiniband interconnect using Voltaire switches and Melanox network interface cards. Since in our tests we used in-core algorithms and all 2-electron integrals were stored in memory, in order to accommodate these integrals sets (each reference has its own set of integrals) our initial core pool is relatively large (320 cores).

3.3 Results and discussion

Basis set	number of atomic basis functions	MR BWCCSD energy	CCSD energy	CCSD(T) energy
cc-pVDZ (c)	180	-383.409952	-383.370483	-383.482383
6-311G** (s)	216	-383.655572	-383.615180	-383.735768
6-311G** (c)	226	-383.677793	-383.637122	-383.760103
Roos aDZ ANO (s)	284	-383.654192	-383.613623	-383.741081
Roos aDZ ANO (c)	304	-383.702472	-383.661766	-383.789998
cc-pVTZ (s)	384	-383.840329	-383.797020	-383.935954
cc-pVTZ (c)	440	-383.887897	-383.844458	-383.984428

Table 3.1: MR BWCCSD (4-reference) energies of the lowest singlet state of 1,7-naphthyne obtained with various basis sets. In the discussed calculations all electrons were correlated.

The results of our tests are shown in Tables 3.1, 3.2 and 3.3 and in Figures 3.2 and 3.3. The MR BWCCSD, CCSD, and CCSD(T) energies for various basis sets are shown in Table 3.1. All energies were converged using tight convergence criterion set equal to 10^{-7}. All CCSD calculations clearly indicate multiconfigurational character of the ground-state wave functions. The largest doubly excited CCSD amplitude corresponds to HOMO → LUMO excitation, equally the largest singly excited amplitudes correspond to HOMO → LUMO excitations as well. For example, for the cc-pVTZ(c) basis set these amplitudes are equal to -0.42 and 0.11, respectively. Since the $|\Phi\rangle$, $|\Phi_{H\alpha H\beta}^{L\alpha L\beta}\rangle$, $|\Phi_{H\alpha}^{L\alpha}\rangle$, and $|\Phi_{H\beta}^{L\beta}\rangle$ (where L refers to LUMO and H to HOMO) determinants are included in the model space, we expect that the MR BWCCSD method should provide a much better description of the correlation effects than the CCSD approach.

In all calculations shown in Table 3.1, the MR BWCCSD energies are invariably located between the CCSD and CCSD(T) ones. Given the size of the CCSD amplitudes the perturbative CCSD(T) estimates of higher order excitations may not be very accurate in describing the ground state energy. This situation is akin to the common problems encountered in studies of bond breaking processes. The

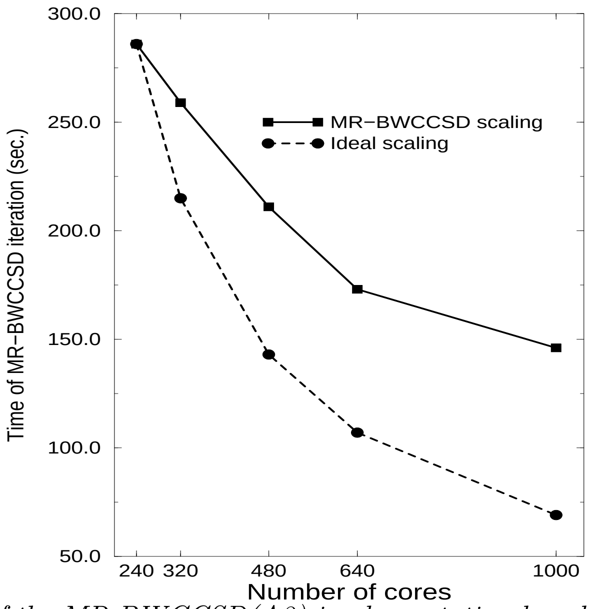

Figure 3.2: Scalability of the MR BWCCSD(A2) implementation based on the improved task scheduling. The Roos aug-DZ ANO(s) basis set was used. Timings were determined from calculations on the Chinook system at EMSL.

largest discrepancies between the CCSD and MR BWCCSD energies can be observed for the larger basis sets. For example, for the cc-pVTZ(c) basis set the difference between CCSD and MR BWCCSD energies is equal to 43.3 milliHartree. The structure of the eigenvectors of the effective Hamiltonian can be used as measure of static correlation included in the model space. For the cc-pVTZ(c) basis set the right eigenvector of $H^{\rm eff}$ is of the form $(0.77, -0.62, 0.08, 0.08)$, where the components ($C_\mu^\alpha$ coefficients describe the contributions from the $|\Phi\rangle$, $|\Phi_{H\alpha H\beta}^{L\alpha L\beta}\rangle$, $|\Phi_{H\alpha}^{L\alpha}\rangle$, and $|\Phi_{H\beta}^{L\beta}\rangle$ determinants. It is also instructive to compare the number of iterations required to converge the CCSD and MR BWCCSD energies using the same convergence criteria. In the calculations reported here the single reference solver was based on the DIIS method while the multi reference solver employed a simple Jacobi-type algorithm. For the 6-311G**(c), Roos aug-DZ ANO(c), and cc-pVTZ(c) basis sets the MR BWCCSD (CCSD) method requires 34 (31), 38 (33), 39 (31) iterations to converge energy to the 10^{-7} threshold.

As discussed earlier, we have two implementations of the MR BWCCSD method. In Table 3.2 we compare the times per iteration of these two approaches for two basis sets: Roos aug-DZ ANO and cc-pVTZ. One can see that in both cases the MR BWCCSD(A2) approach offers a significant performance enhancement over the MR BWCCSD(A1) version. For example, for the Roos aug-DZ ANO (s) basis the MR BWCCSD(A2) is 1.89 time faster than the MR BWCCSD(A1) implementation when 640 cores are used. For the cc-pVTZ(s) basis MR BWCCSD(A2) offers 1.83 time speed-up when 1020 cores are used.

Basis set	number of cores	MR BWCCSD(A1) time (sec.)	MR BWCCSD(A2) time (sec.)	Speed-up
Roos aDZ ANO (s)	640	323	171	1.89
Roos aDZ ANO (c)	640	356	190	1.87
cc-pVTZ (s)	1020	583	319	1.83
cc-pVTZ (c)	1280	851	544	1.56

Table 3.2: Comparison of the MR BWCCSD timings per iteration for standard algorithm (MR BWCCSD(A1)) and the algorithm based on improved task scheduling. In all calculations all electrons were correlated and `tilesize=30` was used.

Number of cores	Time per iteration (sec.)
	Roos aug-DZ ANO(s)
320	228
640	171
1000	134

Table 3.3: Timings per iterations of the MR BWCCSD(A2) algorithm as a function of the number of cores used. In all calculations all electrons were correlated and `tilesize=30` was used.

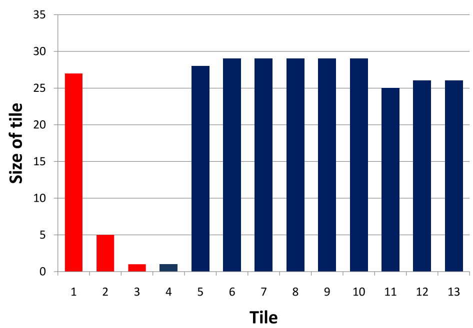

Figure 3.3: The typical sizes of α tiles in the Roos aug-DZ ANO MR BWCCSD calculations (red - occupied α tiles, blue - unoccupied α tiles).

The scalability tests were performed for Roos aug-DZ ANO basis set (see Figure 3.2 and Table 3.3 for timings). One can see that going from 320 cores to 1020 cores results in almost two-fold reduction of time per iteration of the MR BWCCSD(A2) algorithm. The main problem of current implementation is associated with the tile structure used in the calculations, which requires that each active spinorbital is described by a separate tile. A typical situation is shown in Fig. 3.2, where among several large-size tiles there are few very small tiles of size 1, which can cause a problems with load balancing. In the next chapter we will eliminate this requirement by on-the-fly control of the internal excitation in the $T^{(\mu)}$ operator (in each iteration this will require zeroing the corresponding cluster amplitudes). Similar changes will be applied to the residual vectors.

4. Parallel implementation of MRCC methods based on two-level parallelism

This chapter has been published in Journal of Chemical Theory and Computation: Jiří Brabec, Jiří Pittner, Hubertus J. J. van Dam, Edoardo Aprà, and Karol Kowalski, Journal of Chemical Theory and Computation 2012 8 (2), 487-497

4.1 Parallel algorithm

In this section, we will describe an algorithm which employs a two-level parallelism: (1) task-level parallelism, used to calculate given set of reference-specific equations ($\mathbf{R}^{(\mu)}$); (2) reference-level parallelism, where each set of $\mathbf{R}^{(\mu)}$ equations (or their aggregate) is calculated on separate processor group (PG). In this work we will focus on the BW-MRCCSD and Mk-MRCCSD implementations although the same ideas can be naturally extended to the SU-MRCCSD case. Use of processor groups has been shown to be an efficient approach for the parallelization of electronic structure codes[70–72]. By adopting this coarse grain level of parallelism, it is possible to overcome the scaling limit of finer level parallelism that includes all the processing elements. However, use of processor groups is only feasible when the calculation can be easily divided in subunits, each one of them with a similar computational cost; among the example of methods amenable to this kind of parallelization scheme, we can mention methods that require the calculation of energy derivatives by numerical differentiation (needed for geometry optimization, transition state searches, frequencies calculation), reduced scaling methods that divide molecules into fragments, and linear algebra parts of electronic structure periodic codes where "images" are assigned to different groups of processors (e.g., k-point parallelism for diagonalization of a matrix represented in reciprocal space). In our opinion, MRCC methods with their coarse grain level parallelism corresponding to the reference-level parallelization are yet another candidate for an effective use of processor groups. In the context of MRCCSD theories, the use of PGs substantially reduces the amount of communication within a given subgroup. We will demonstrate in the remainder of this thesis that this approach is capable of efficient scaling to a large processor count, overcoming communication bottlenecks present in the task level parallelism.

The MRCC implementations discussed in this work use in-core algorithms; i.e., all electron integrals, cluster amplitudes, and recursive intermediates are stored in GA. As mentioned earlier, task-level parallelism is associated with performing calculations within each PG and requires parallel implementations for the direct and coupling terms.

An integral part of forming MRCCSD equations is related to calculating the coupling terms (see Eqs.(1.40)-(1.42)). While for the BW-MRCCSD approach, this is an inexpensive part of the whole calculation, in the Mk-MRCCSD case this constitutes a more challenging task due to the fact that Mk-MRCCSD coupling terms (1.41) involve amplitudes corresponding to all references, which entails intensive communication. Our implementation utilizes the explicit form of the coupling terms derived for the Mk-MRCCSD approximation by Evangelista *et al.*,[19]. In our implementation, for each $\mathbf{G}^{(\mu)}$ in the loop over references ν we first calculate the auxiliary amplitudes $t_a^i(\nu/\mu)$ and $t_{ab}^{ij}(\nu/\mu)$, and then we use a simple TCE generated code to calculate contributions (1.34) and (1.35).

The essential part of the effort associated with the task-level parallelism has been discussed in chapter III. The main difference between the implementation reported here and the previous chapter is the fact that the active spinorbitals are no longer grouped into small separate tiles, which contributed to significant problems with load balancing. Instead of using separate tiles for active spinorbitals, these spinorbitals are included in larger tiles, which make no distinction between active and inactive spin-orbitals. For example, in the BW-MRCCSD and Mk-MRCSSD codes, the components of the $\mathbf{F}_\mu$ vector form a subset of all components corresponding to a larger vector $\mathbf{E}_\mu$, which contains all elements

$$\langle\Phi_\nu|(He^{T^{(\mu)}})_C|\Phi_\mu\rangle, \langle\Phi_i^a(\mu)\rangle|(He^{T^{(\mu)}})_C|\Phi_\mu\rangle, \langle\Phi_{ij}^{ab}(\mu)\rangle|(He^{T^{(\mu)}})_C|\Phi_\mu\rangle \ . \tag{4.1}$$

aligned in some order defined by tiling scheme, where the $\langle\Phi_\nu|(He^{T^{(\mu)}})_C|\Phi_\mu\rangle$ corresponds to matrix elements of the effective Hamiltonian (1.24). Although this new approach results in a more irregular distribution of the elements of the effective Hamiltonian (see Figure 4.1), the larger sizes of the tiles lead to a better utilization of `BLAS` procedures[73–75] and substantially improve the load balancing. In analogy to chapter III, matrix elements between up to mutually doubly excited references have been included in the effective Hamiltonian.

To facilitate discussion of the reference-level parallelism, let us introduce basic terminology and notation. By processor group (G_i), we will understand a partitioning of the processor domain (D) into smaller pieces, which can be symbolically

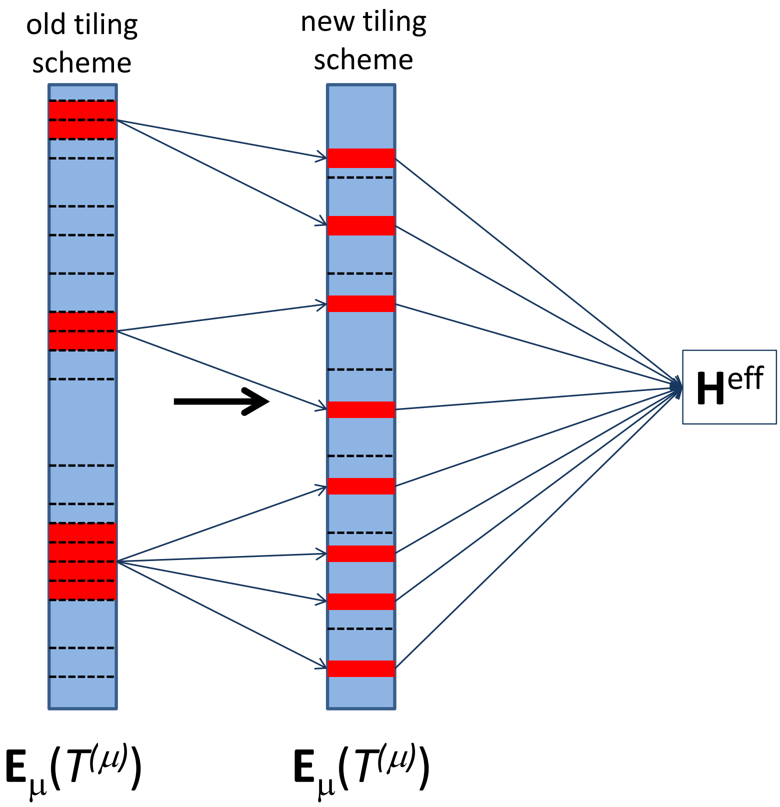

Figure 4.1: Schematic representation of the new tiling algorithms. The red blocks correspond to the elements of the effective Hamiltonian, while the blue ones correspond to the elements of the $\mathbf{F}_\mu(\mathrm{T}^{(\mu)})$ *vector. The dashed lines separate different blocks of the* $\mathbf{E}_\mu(\mathrm{T}^{(\mu)})$ *vector induced by a given tiling scheme. The block structure of* $\mathbf{E}_\mu(\mathrm{T}^{(\mu)})$ *in the old tiling scheme is more irregular. In the new tiling scheme, all* $\mathbf{E}_\mu(\mathrm{T}^{(\mu)})$ *blocks are much less diversified sizewise.*

expressed

$$D = \bigcup_{i=1,\ldots,I} G_i \tag{4.2}$$

where I is the total number of the processor groups. We will assume that the number of processors in each group is the same and is equal to S, i.e.,

$$S_i = S = \frac{N_p}{I} \quad (i = 1, \ldots, I) \tag{4.3}$$

where S_i is the size of the ith group. In the above equation N_p stands for the total number of processors, which is a multiple of PGs number I. The key idea of the discussed algorithm is to distribute processes of forming $\mathbf{R}^{(\mu)}$ equations over various processor groups. In the simplest case, the work organization chart (symbolically designated by W) corresponds to the situation when a single PG is delegated to calculate single set of equations $\mathbf{R}^{(\mu)}$. In this case, the number of PGs coincides with the size of the model space (i.e. $I = M$). A more general situation corresponds to the case when each PG (G_i) forms several ($n_r(i)$) residual vectors $\mathbf{R}^{(\mu)}$. This can be symbolically denoted as

$$W = \bigcup_{i=1,\ldots,I} W_i(n_r(i)) \tag{4.4}$$

where W_i refers to work-load on the corresponding processor group G_i. In order to provide best load balancing between the workloads on each PG, it is natural to assume that

$$n_r = n_r(i) = \frac{M}{I} \ \ (i = 1, \ldots, I)\ . \tag{4.5}$$

In the following, in order to report specific configurations of the processor groups used in the calculation, we will use the $PG(n_r, I, S)$ notation.

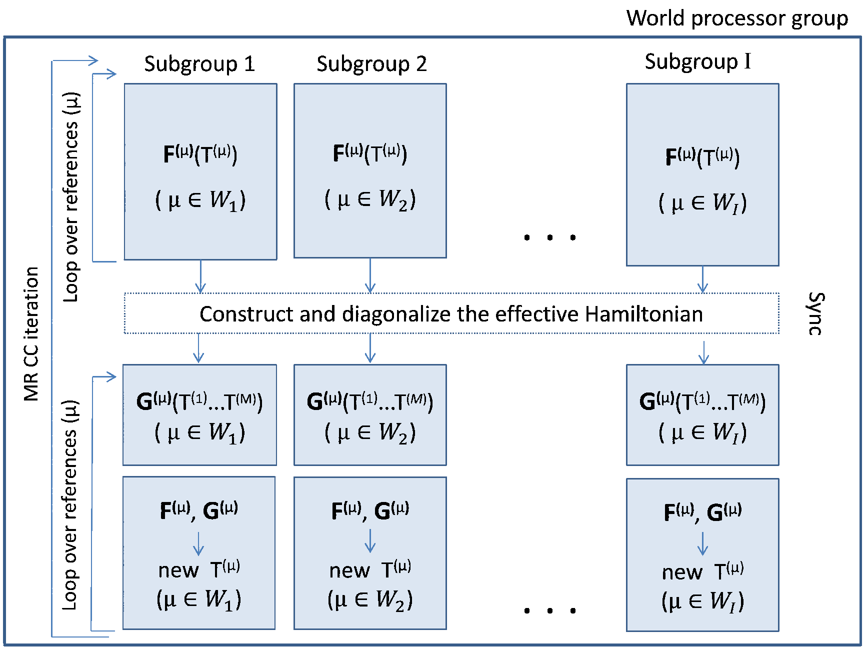

Figure 4.2: Simplified representation of the processor-group-based MRCC algorithm for the BW-MRCCSD and Mk-MRCCSD formulations.

The schematic representation of our algorithm is shown in Fig.4.2. In the first step, the domain D (or the world processor group) is partitioned into I-groups $\{G_i\}_{i=1}^{I}$. Each group calculates the corresponding subset of $\mathbf{F}^{(\mu)}$ vectors and the subset of matrix elements of the effective Hamiltonian. In the second step, the effective Hamiltonian is formed and diagonalized to obtain the energy estimate in a given iteration cycle. In the third step, the processor groups are used to calculate coupling terms and to build the updates for the cluster amplitudes. In this process, each processor group works on the cluster amplitudes for the subset of references delegated to a given PG. As a result of the different structure of the sufficiency conditions and the presence of communication-intensive coupling terms in the Mk-MRCCSD approaches, the utilization of the PG and data organization are slightly different in the BW-MRCCSD and Mk-MRCCSD approaches:

- In BW-MRCCSD, all sufficiency conditions involve only cluster amplitudes relevant to a given reference. Therefore, each PG stores corresponding copies of cluster amplitudes, one- and two-electron integrals, residual vectors, and all recursive intermediates required to calculate the direct term.

- Since the reference-specific Mk-MRCCSD equations require all types of cluster amplitudes, in the Mk-MRCCSD approach, all cluster amplitudes are stored in the so-called `world group`; i.e., they can be accessible by all PGs. Other tensors including one- and two-body integrals, residual vectors, and recursive intermediates of the direct terms are stored only on a corresponding processor group.

For relatively small model spaces discussed here, these schemes seem to provide the most effective solution. The data organization described above determines the communication and memory requirements of our code. One should mention that while the most memory consuming contribution (in the sense of a given PG) is related to the two-electron integrals (of the N^4 size, where N stands for the number of spinorbitals), the size of doubly excited cluster amplitudes file and the size of the maximum recursive intermediate used in the direct terms are proportional to $n_o^2 n_u^2$ (n_o and n_u refer to the numbers of occupied and unoccupied spin-orbitals).

We have two types of synchronization steps: one is internal to PG (internal synchronization), and the other is executed by all processes (global synchronization). While the PG internal synchronization steps are required in calculating direct and coupling terms, the global synchronization steps are required before and after calculating the effective Hamiltonian and before and after calling DIIS solver ("after" synchronizations are required to broadcast the results to all processes). The improved implementation of the direct term (the so-call A2 algorithm of chapter III) requires only four internal PG synchronization steps. In the more laborious coupling terms of the Mk-MRCCSD approach, there are several PG internal synchronization steps corresponding to each reference contributing to a given coupling term.

Several techniques can be used to find the solution of highly nonlinear BW-MRCCSD and Mk-MRCCSD equations. For the BW-MRCCSD method, we used a simple Jacobi style solver, which can easily be extended to the DIIS algorithm[76,77]. In general, the BW-MRCCSD equations are much easier to converge than their Mk-MRCCSD counterparts. Achieving the same accuracy with the Mk-MRCCSD approach poses a more challenging task. In our calculations, we employed a simple iterative scheme: We performed several (usually around 10) Jacobi iterations with the BW-MRCCSD method followed by the DIIS algorithm for the Mk-MRCCSD equations. Usually, this hybrid solver is capable of reducing the number of iterations required to converge the Mk-MRCCSD equations (in situations where the convergence is possible to achieve). In general, however, it is hard to design an efficient solver to the Mk-MRCCSD equations capable of reducing the number of required iterations to the level of single reference CC methods or even of the BWMRCCSD approach. Moreover, performing more BW-MRCCSD iterations does not always lead to a systematic reduction in the iteration count of the subsequent Mk-MRCCSD solver. We believe that more sophisticated Newton-Raphson-type forms of the preconditioners used in the amplitudes updates may be more efficient in the reduction of total number of iterations. However, this will introduce additional numerical effort associated with approximating the Jacobian matrix in each iteration.

4.2 Computational details

We performed numerical tests for several benchmark systems: from small N_2 system to larger molecules such a 1,7-naphthyne and β-carotene. The MRCC calculations for N_2, 1,7-naphthyne, and β-carotene have been performed using aug-cc-pVTZ,[67] cc-pVTZ,[67] and 6-31G basis sets[78], respectively. For the 1,7 naphtyne isomer, we used the geometry from the chapter III, while the geometry of the β-carotene was optimized with the DFT B3LYP approach[79] using cc-pVTZ basis set (the geometries of 1,7-naphthyne and β-carotene and the corresponding Hartree-Fock energies can be found in the Supporting Information). For the N_2 systems, the calculations have been performed for stretched internuclear geometry (4.136 Bohr). All calculations were based on the Hartree-Fock orbitals using C_{2v}, C_s, and C_i symmetry groups for the N_2 molecule, 1,7-naphthyne isomer, and β-carotene, respectively. The following complete model spaces have been used for these systems: CMS(6,6) (N_2), CMS(2,2) (1,7-naphthyne isomer) and CMS(4,4) (β-carotene). These model spaces include the following orbitals: $5a_1$ (HOMO-

2), $1b_1$ (HOMO-1), $1b_2$ (HOMO), $2b_2$ (LUMO), $2b_1$ (LUMO+1), $6a_1$ (LUMO+2) for the N_2 molecule, $28a'$ (HOMO) and $29a'$ (LUMO) for 1,7-naphthyne isomer, and $74a_g$ (HOMO-1), $74a_u$ (HOMO), $75a_g$ (LUMO) and $75a_u$ (LUMO+1) for β-carotene. The corresponding complete model spaces include 112 (M = 112), 4 (M = 4), and 20 (M = 20) reference functions, respectively. For N_2 and β-carotene, all core orbitals have been kept frozen. Additionally, for the N_2 system, the highest two virtual orbitals were kept frozen.

Our codes have been integrated with the NWChem software.[80] The scalability tests were performed on two computer architectures: for the N_2 and 1,7-naphthyne systems we used the Chinook system at PNNL consisting of 2310 HP DL185 nodes with dual socket, 64-bit Quad-core AMD 2.2 GHz Opteron processors. Each node has 32 GB of memory, i.e. 4 GB per core. The nodes are connected with a single rail Infiniband interconnect using Voltaire switches and Mellanox network interface cards, whereas for β-carotene we used the Jaguar system at ORNL, which contains 18 688 compute nodes in addition to dedicated login/service nodes. Each XT5 compute node contains dual hex-core AMD Opteron 2435 (Istanbul) processors and 16 GB of memory and uses the Cray proprietary Seastar network.

4.3 Results and discussion

We start our analysis of the parallel performance of the MRCCSD codes with a discussion of the BW-MRCCSD implementation for the N_2 molecule. Although the size of the molecule hardly allows one to categorize it as a "large-molecular systems", several interesting performance-related issues can be illustrated with this example. In particular, we can illustrate the performance of the two-level parallelism. In Fig.4.3, we showed the timings of the BW-MRCCSD iterations as a function of the number of processes used. In these tests, we used our first version of the BW-MRCCSD approach (version A1 of chapter III). The calculations were performed for several configurations corresponding to different uses of processor groups: $PG(7,16,1)$, $PG(4,28,1)$, $PG(2,56,1)$, $PG(1,112,1)$, $PG(1,112,2)$, $PG(1,112,4)$, $PG(1,112,8)$, $PG(1,112,16)$, $PG(1,112,32)$, where 16, 28, 56, 112, 224, 448, 896, 1792, 3584 cores have been used, respectively. As explained earlier, we use the $P(n_r, I, S)$ specification of the computational setup where n_r, I, and S designate the number of equations per processor group, total number of processor groups, and number of processors in each group, respectively. One can notice that the performance of the original version of the code quickly saturates

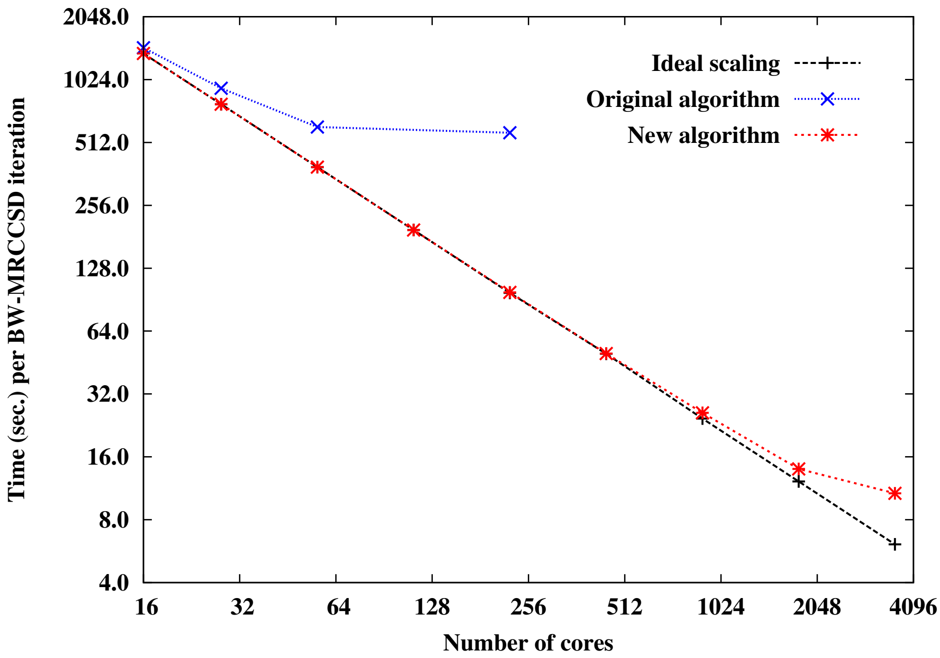

Figure 4.3: The scalability of the BW-MRCCSD/CMS(6,6) (112 reference functions; $M = 112$) calculations for the N_2 molecule in the aug-cc-pVTZ basis set (two lowest occupied and two highest unoccupied orbitals were kept frozen in the calculations). The calculations were performed using 16, 28, 56, 112, 224, 448, 896, 1792, 3584 cores using the following processor groups $PG(7,16,1)$, $PG(4,28,1)$, $PG(2,56,1)$, $PG(1,112,1)$, $PG(1,112,2)$, $PG(1,112,4)$, $PG(1,112,8)$, $PG(1,112,16)$, $PG(1,112,32)$ respectively (see text for details).

around 56 cores, which is a consequence of the small size of the system and relatively small task pool when standard way of executing BW-MRCCSD sequence of equations is invoked (i.e., serial execution of parallel routines corresponding to direct and coupling terms for various reference functions). Compared to the original algorithm, the use of the processor groups can significantly enlarge the processes' pool that can be effectively utilized in parallel BW-MRCCSD calculations for this system, allowing for efficient calculations with 3584 cores. This was possible thanks to the flexible use of two-level parallelism. While up to 112 cores mostly reference-level parallelism is employed, which is associated with calculating various number of reference-specific BW-MRCCSD equations per processor group (see the definition of the n_r parameter), for calculations involving larger number of cores, the task-level parallelism (or the parallelism of the particular procedures used to calculate $\mathbf{E}_\mu$ vector) plays the leading role. From 16 to 112 cores, the calculations were performed with one core per processor group, which means that the number of processor groups is equivalent to the total number of cores employed. From 224 cores, the number of cores per processor group starts

Number of cores	PG configuration	Time (sec.)
1024	$PG(4,1,1024)$	231
2048	$PG(2,2,1024)$	118
4096	$PG(1,4,1024)$	62

Table 4.1: BW-MRCCSD timings (per BW-MRCCSD iteration) for the 1,7-naphthyne as described by the cc-pVTZ basis set (in spherical representation). We used (2,2) CMS (4 reference functions), all electrons were correlated in the course of calculations. The A2 algorithm (improved task scheduling described in chapter III) and improved tiling scheme were used. The scalability tests were performed on the Chinook system at PNNL using four cores per eight-core node. The calculated BW-MRCCSD energy with a posteriori correction of Refs.[43,44] is equal to -383.84033 Hartree.

to increase from 2 cores per processor group (224 processes run) to 36 cores per processor group (3584 processes run). The good performance from 112 to 3584 should be mostly attributed to the good scalability of the original code in the region of 2-32 cores. This example gives us a general idea how these two levels of parallelism can be intertwined to provide an optimum use of large number of processors.

Next, we analyze the code performance for the larger system, the 1,7-naphthyne isomer. In contrast to the N_2 benchmark, the 1,7-naphthyne isomer is a much larger molecule, while the model space considered is much smaller (M = 4 vs M = 112). By having two σ-radical centers at different atoms, the naphthyne isomers are very interesting systems from the viewpoint of MRCC calculations. Depending on the mutual distance of these radical centers, the strength of the quasi-degeneracy varies and should grow with the distance. The 1,7 isomer (Fig.3.1) belongs to those which have strong multireference character.[66] For the lowest singlet state of the 1,7 isomer the CMS(2,2) contains four reference functions including RHF determinant $|\Phi\rangle$, two singly excited determinants $|\Phi_{H\alpha}^{L\alpha}\rangle$ and $|\Phi_{H\beta}^{L\beta}\rangle$, and doubly excited configuration $|\Phi_{H\alpha H\beta}^{L\alpha L\beta}\rangle$ where H and L correspond to the highest occupied and lowest unoccupied orbitals, respectively. In previous chapter we showed that the $|\Phi\rangle$ and $|\Phi_{H\alpha H\beta}^{L\alpha L\beta}\rangle$ determinants are nearly degenerate in the expansion for the ground state. The scalability tests for the 1,7-naphthyne isomer are collected in Tables 4.1 and 4.2, and in Figs. 4.4, 4.5, and 4.6. The BW-MRCCSD and Mk-MRCCSD scalability tests for the cc-pVTZ basis set (in spherical representation) have been performed on 1024, 2048, and 4096 cores of Chinook PNNL system using the $PG(4,1,1024)$, $PG(2,2,1024)$, and $PG(1,4,1024)$ con-

Number of cores	PG configuration	Time (sec.)
	BW-MRCCSD	
1024	$PG(4,1,1024)$	525
2048	$PG(2,2,1024)$	235
4096	$PG(1,4,1024)$	86
	Mk-MRCCSD	
1024	$PG(4,1,1024)$	580
2048	$PG(2,2,1024)$	290
4096	$PG(1,4,1024)$	140

Table 4.2: BW-MRCCSD and Mk-MRCCSD timings (per BW-MRCCSD/Mk-MRCCSD iteration) for the 1,7-naphthyne as described by the cc-pVTZ basis set (in spherical representation). We used (2,2) CMS (4 reference functions), all electrons were correlated in the course of calculations. The A2 algorithm (improved task scheduling described in chapter III) and improved tiling scheme were used. The scalability tests were performed on the Chinook system at PNNL using eight cores per eight-core node.The calculated BW-MRCCSD energy with a posteriori correction of Refs.[43,44] is equal to -383.84033 Hartree, while the Mk-MRCCSD energy is equal to -383.82863 Hartree

figurations, respectively. In these tests eight cores of the eight-cores node have been used (see Table 4.1). These results indicate a very good performance of the iterative BW-MRCCSD calculations when moving from 1024 to 4096 cores and using reference-level parallelism. For the same basis set and for the same computational setup, one can compare the timings of the timing of the BW-MRCCSD and Mk-MRCCSD approaches (see Table 4.2). The additional cost associated with the inclusion of more demanding (from the point of view of communication) coupling terms in the Mk-MRCCSD equation (see Eq.(1.41)) is a function of the number of cores used. It is interesting to notice that for each case (1024, 2048, and 4096 cores) the timing per Mk-MRCCSD iteration is by $\simeq$ 55 s bigger compared to the corresponding timings of the BW-MRCCSD iterations. At the same point, the scalability of the Mk-MRCCSD code for 1,7-naphthyne isomer remains still very good. A different computational setup - four cores per eight-core node on the Chinook system - resulted in a significant reduction of the time per one BW-MRCCSD iteration. For 1024 and 2048 cores, one can observe a two-fold reduction of time per BW-MRCCSD iteration (see Table 4.1 and Fig.4.6). This improvement in performances by using a reduced number of cores per node can be attributed to the following: (i) ARMCI (the low level communication library used by GA) is implemented on several high-performance networks (including Infiniband and Seastar) using a communication helper thread that burns resources[81], (ii) The limited amount of memory available on each core (e.g., 1.3 GB on Jaguar Cray XT5); therefore using fewer cores one obtains a larger amount of memory per core. (iii) When network resources are shared among multiple cores, communication performance might see a degradation, more likely seen in latency (see Figure 4b of reference[82]). However, timings of the BW-MRCCSD runs shown in Tables 4.1 and 4.2 suggest that the use of eight cores per node is preferable from the point of view of efficient utilization of the computational resources since the time to solution is reduced. For example, the 86 s per BW-MRCCSD iteration when 512 nodes are used in an eight-core per node run (for a total of 4096 processors, see Table 4.2) should be compared with the 118 s per BW-MRCCSD iteration when the same number of nodes is used, but only four cores are utilized on each node (2048 cores, see Table 4.1).

The last remark on the 1,7-naphthyne isomer concerns the performance of the iterative solvers for the BW-MRCCSD and Mk-MRCCSD approaches. A typical situation is shown in Figure 4.7 where the discrepancies between exact MRCCSD energy and energies obtained for particular iterations are shown for calculations

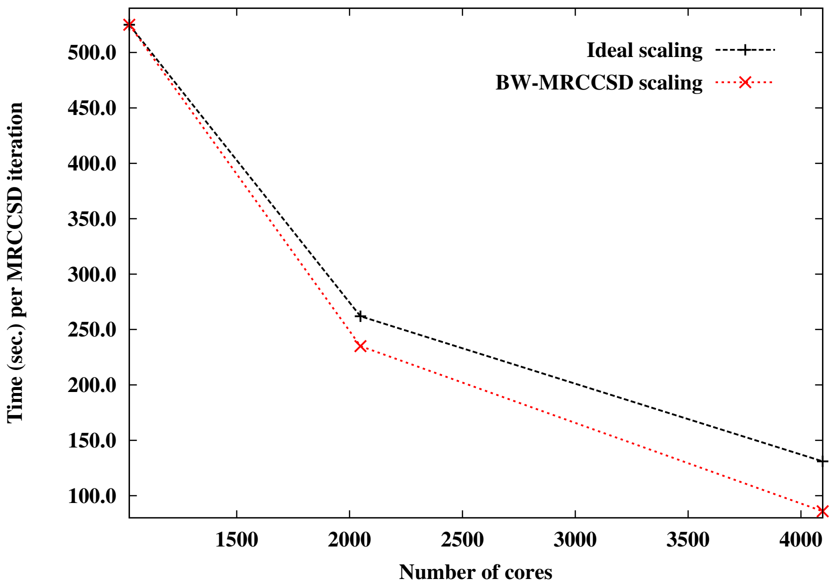

Figure 4.4: The scalability of the BW-MRCCSD/CMS(2,2) (4 reference functions; M = 4) calculations for the 1,7-naphthyne isomer in the spherical cc-pVTZ basis set (all orbitals were correlated in the calculations). The calculations were performed using 1024, 2048, and 4096 cores employing $PG(4,1,1024)$, $PG(2,2,1024)$, *and* $PG(1,4,1024)$ *configurations, respectively (see Table 4.2 for the corresponding timings). In these tests we used 8 cores per 8-core node of Chinook system at PNNL.*

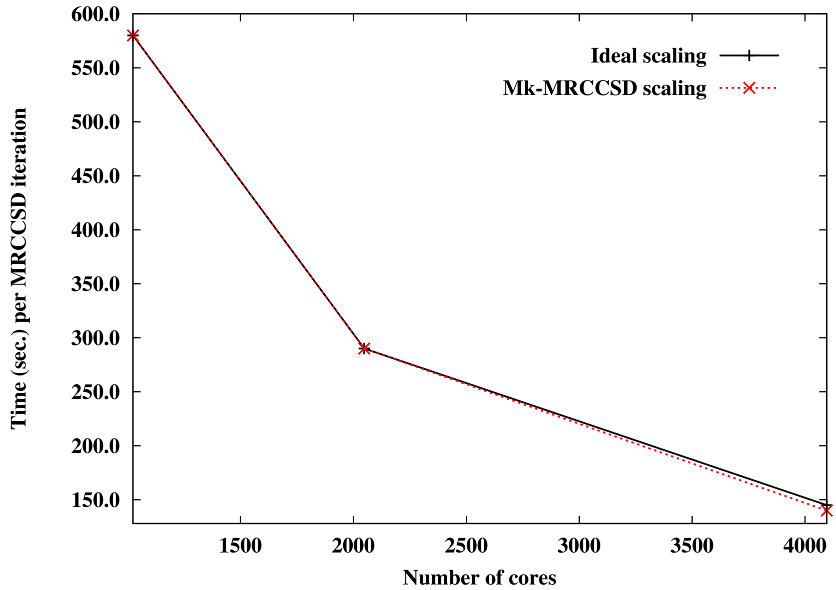

Figure 4.5: The scalability of the Mk-MRCCSD/CMS(2,2) (4 reference functions; $M = 4$) calculations for the 1,7-naphthyne isomer in the spherical cc-pVTZ basis set (all orbitals were correlated in the calculations). The calculations were performed using 1024, 2048, and 4096 cores employing $PG(4, 1, 1024)$, $PG(2, 2, 1024)$, and $PG(1, 4, 1024)$ configurations, respectively (see Table 4.2 for the corresponding timings). In these tests we used eight cores per eight-core node of Chinook system at PNNL.

for the 1,7-naphthyne isomer in the cc-pVTZ basis set (spherical representation). The application of the DIIS procedure to the BW-MRCCSD iterations (DIIS cycle length was set equal to 5) resulted in 13 iterations required to converge energy to the 10^{-7} Hartree threshold. For the Mk- MRCCSD approach, which is notorious for convergence problems, we used a slightly different strategy. First, we let the code perform 10 BW-MRCCSD iterations without DIIS support. Then, the full form of the Mk-MRCCSD equations was switched back on, and the remaining iterations were performed with the help of the DIIS algorithm. In the Mk-MRCCSD case, we found that the use of a longer DIIS cycle may be beneficial for accelerating the convergence. In the discussed calculations, it took 53 iterations to converge the Mk- MRCCSD energy to within 10^{-7} Hartree. When discussing the iterative processes, it is also instructive to compare the structure of the right eigenvectors of the effective Hamiltonians obtained with the BW-MRCCSD (with a posteriori correction) and Mk- MRCCSD methods. Both approaches yield similar eigenvectors (0.774, -0.624, -0.076, -0.076) for BW-MRCCSD and (0.796, -0.593, -0.090, -0.090 for Mk-MRCCSD) where the components corresponds to the $\{|20|, |02|, |\alpha\beta|, |\beta\alpha|\}$ reference determinants (where we show occupancies of active orbitals only; for example, $|20|$ corresponds to the determinant with dou-

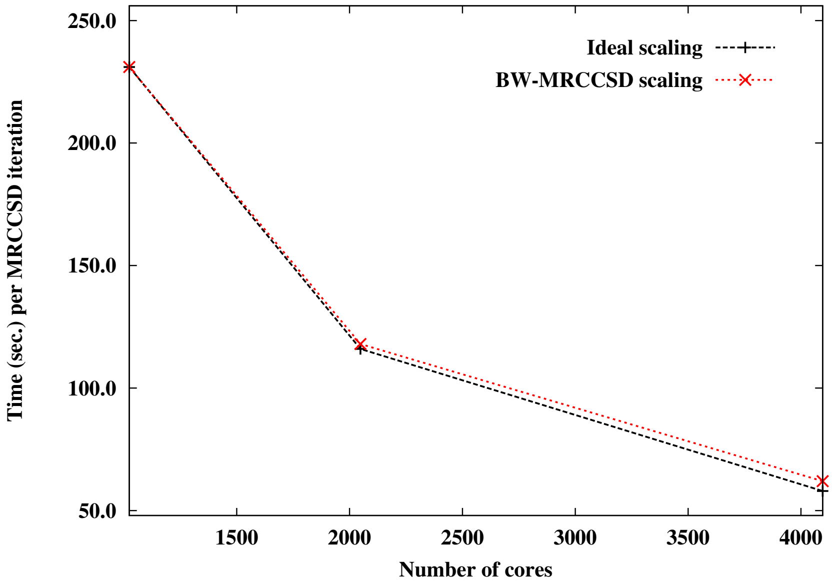

Figure 4.6: The scalability of the BW-MRCCSD/CMS(2,2) (4 reference functions; $M = 4$) calculations for the 1,7-naphthyne isomer in the spherical cc-pVTZ basis set (all orbitals were correlated in the calculations). The calculations were performed using 1024, 2048, and 4096 cores employing $PG(4, 1, 1024)$, $PG(2, 2, 1024)$, and $PG(1, 4, 1024)$ configurations, respectively (see Table 4.1 for the corresponding timings). In these tests we used four cores per eight-core node of Chinook system at PNNL.

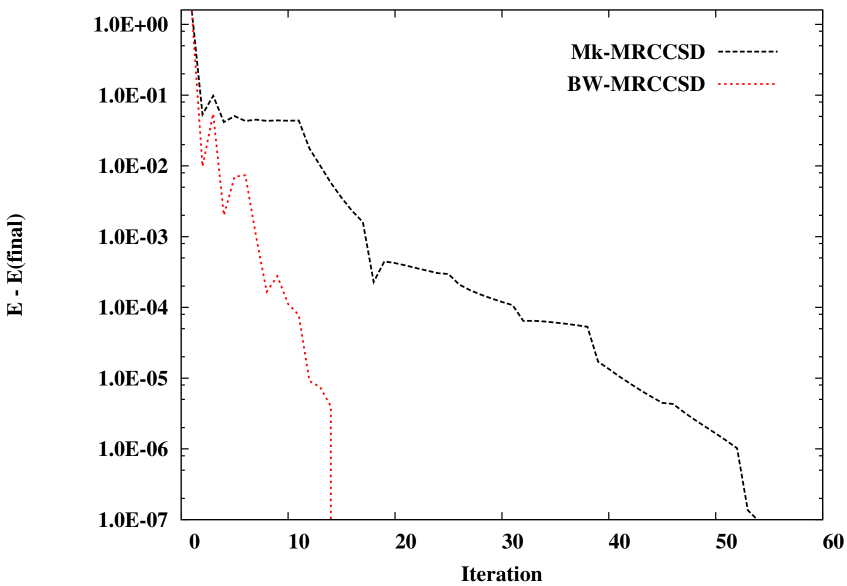

Figure 4.7: Comparison of the energy convergence characteristic during the BW-MRCCSD and Mk-MRCCSD iterations for the 1,7-naphthyne isomer in the cc-pVTZ basis set (in the spherical representation). All core electrons were kept frozen in the calculations.

bly occupied active orbital $28a'$ while $|\alpha\beta|$ corresponds to the determinant with α and β electrons on active orbitals $28a'$ and $29a'$, respectively). The structure of these eigenvectors shows the importance of the "open-shell" configurations $|\alpha\beta|$ and $|\beta\alpha|$ in the wave function expansion.

Our last example is related to the scalability tests for the β-carotene molecule (see Fig.4.8). The choice of this system is dictated by the fact that the proper description of excited states of carotenoids poses a significant challenge for theoretical approaches and requires multireference methods to describe complicated low-lying excited states.[83,84] However, for system of this size, the MRCC calculations are very expensive. In the present work, we would like to demonstrate that the MRCC calculations (in the BW-MRCCSD variant) are possible at least for the ground state (in the future, we are planning to perform excited-state studies using SU-MRCCSD method, which is characterized by a similar numerical complexity as the BW-MRCCSD method).

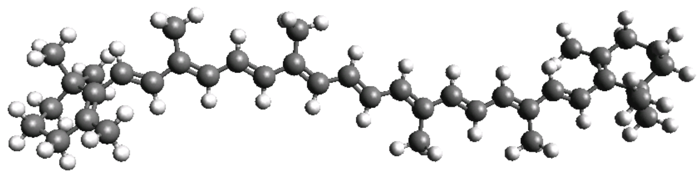

Figure 4.8: The structure of the β-carotene. The geometry optimization was performed using DFT B3LYP approach using cc-pVTZ basis set.

In all our numerical tests, we used processor groups including 1200 processors ($S = 12000$); therefore, these tests, in contrast to the N_2 tests, mainly reflects the reference-level parallelism. The results of our simulations are shown in Table 4.3 and in Fig.4.9, where we included timings for 6000, 12000, and 24000 core runs. In these calculations, four ($n_r = 4$), two ($n_r = 2$), and one ($n_r = 1$) reference-specific equations were generated on each group, respectively.

Moreover, these timings were obtained with the A2 algorithm based on a new

Number of cores	Configuration	Time (sec.)
6000	$PG(4, 5, 1200)$	2710
12000	$PG(2, 10, 1200)$	1216
24000	$PG(1, 20, 1200)$	660

Table 4.3: Timings for β-carotene in 6-31G basis set. The (4,4) CMS was used (20 reference functions) and all core electrons were kept frozen in calculations. In our tests we used version A2 from chapter III based on the improved tiling scheme described previously. All calculations were performed on Cray XT5 system at ORNL. The calculated BW-MRCCSD energy with a posteriori correction of Refs.[43,44] is equal to -1550.6189 Hartree.

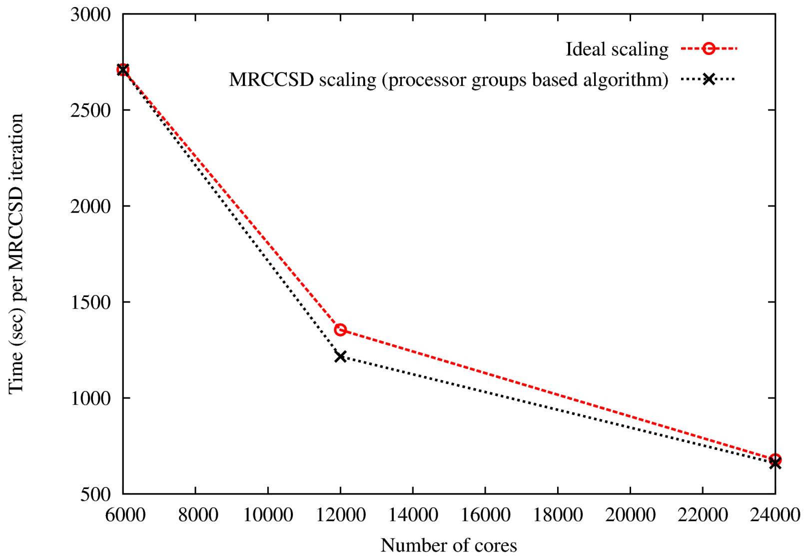

Figure 4.9: The scalability of the BW-MRCCSD/CMS(4,4) (20 reference functions, $M = 20$) calculations for the β-carotene in the 6-31G basis set (all core orbitals were kept frozen in the calculations). The calculations were performed using 6000, 12000, and 24000 cores employing $PG(4, 5, 1200)$, $PG(2, 10, 1200)$, and $PG(1, 20, 1200)$ configurations, respectively. In these tests we used six cores per twelve-core node of Jaguar system at ORNL.

tiling scheme (differences in timings between A2 algorithm using old and new tiling schemes are significant; for example old tiling scheme is characterized by $\simeq$ 1900s per iteration on 32000 cores when 20 processor groups are used). From Table 4.3, one can notice that parallel efficiency reaches superlinear 102 % when 6000 and 24000 jobs are compared. The results of Table 4.3 suggest that this efficiency is reduced when larger jobs are compared. For example, efficiency of 24000 and 12000 core jobs drops down to 92 %, which is more realistic estimate of parallel performance of the BW-MRCCSD method when large number of cores is used. We believe that these results provide convincing arguments for using the processor groups as a efficient tool for building efficient MRCC implementations. The structure of the right eigenvector of the BW-MRCCSD effective Hamiltonian (again, the BW-MRCCSD amplitudes include the a posteriori correction) shows that the dominant contributions stem from the following reference determinants: $|2200|$ (-0.925), $|2002|$ (0.194), $|0202|$ (0.157), $|0220|$ (0.120), $|\alpha\beta\alpha\beta|$ (-0.117), and $|\beta\alpha\beta\alpha|$ (-0.117), which demonstarted mildly multireference character of the ground state of the β-carotene molecule.

5. Large-scale calculations with the SS MRCC methods

Parts of this chapter have recently been submitted for publication to Chemical Physics Letters.

5.1 Dodecane

In this section the stretching of the $CH_3 - C_{11}H_{23}$ bond (in the following text denoted as C1–C2) in the dodecane (see Figure 5.1) is studied. Li *et al.*[85] have shown that CCSD and CCSD(T) methods break down at larger bond length. They resorted to using CR-CC(2,3) approach to describe this dissociation process. Alternatively MRCC should be able to describe this bond breaking process correctly[86]. With 98 electrons this system represents a challenging test of MRCC implementation capabilities.

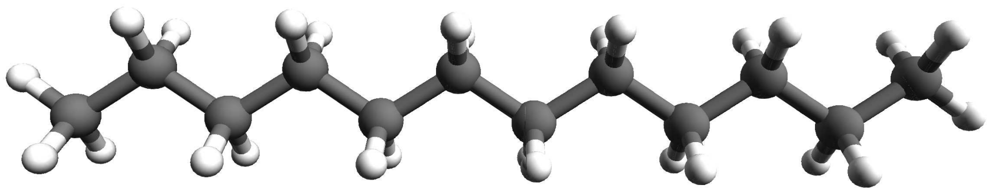

Figure 5.1: The equilibrium geometry of dodecane

In our study we started with a geometry optimized at the Density Functional Theory (DFT) level using the B3LYP[79,87–90] functional and the cc-pVTZ[67,91,92] basis set. Starting from this geometry, MRCCSD calculations were performed for different geometries where the bond lengths between C1 and C2 were set to various multiples of the equilibrium bond length. All other bond lengths and angles were kept fixed at their equilibrium geometry values. The proper dissociating active space for this reaction involves the highest occupied bonding orbital and the lowest unoccupied orbital. This generates a two electrons in two orbitals CMS containing four Slater determinants ($M = 4$) which can be symbolically denoted as $(49a)^2(50a)^0$, $(49a)^0(50a)^2$, $(49a)^\alpha(50a)^\beta$, and $(49a)^\beta(50a)^\alpha$. In all CC calculations the spherical harmonic cc-pVDZ[67] basis set was employed. In our studies core orbitals were dropped from the correlated calculations. Hence 74 electrons were correlated using 286 orbitals.

Method	R_e	$(49a)^2(50a)^0$	$(49a)^0(50a)^2$	$(49a)^1(50a)^1$
MR BWCCSD ap.	1.0	1.00	0.00	0.00
MR MkCCSD	1.0	1.00	0.00	0.00
MR BWCCSD ap.	1.5	0.98	0.18	0.01
MR MkCCSD	1.5	0.98	0.17	0.00
MR BWCCSD ap.	2.0	0.89	0.45	0.03
MR MkCCSD	2.0	0.89	0.45	0.04
MR BWCCSD ap.	3.0	0.73	0.68	0.07
MR MkCCSD	3.0	0.73	0.68	0.07

Table 5.1: The absolute values of right eigenvectors obtained from MRCC calculations on dodecane at different C1 - C2 bond lengths. The weights of the $(49a)^1(50a)^1$ configuration correspond to equal weights of two configurations $(49a)^\alpha(50a)^\beta$ and $(49a)^\beta(50a)^\alpha$.

The results of the MR MkCCSD and MR BWCCSD calculations are shown in Tables 5.1 and 5.2. In Table 5.1 the eigenvectors of the MR BWCCSD effective Hamiltonian are given to illustrate the increasing multireference character of the system with increasing bond length. The data show that the configuration with doubly occupied anti-bonding orbital becomes more important as the bond length increases. This effect is significant for any bond lengths beyond twice the equilibrium bond length. The configuration with one electron in both the bonding and anti-bonding orbital is, as expected, not contributing significantly at any stage. This is a typical situation which is encountered in a breaking of a single bond.

In Table 5.2 the total energies obtained for a number of multireference and single-reference methods are compared. The last column is based on data from Li et al.[85]. They used the 6-31G(d)[78] basis set and geometries optimized with RHF. Nevertheless we have taken their energies and shifted them to match our CCSD

R_e	MR BWCCSD ap.	MkCCSD	CCSD	CCSD(T)	CIM CR(2,3)[85]
1.0	-471.50740	-471.5074	-471.50740	-471.56499	-471.50740
1.5	-471.40811	-471.4072	-471.40537	-471.46731	-471.40684
2.0	-471.34214	-471.3426	-471.33030	-471.40519	-471.34244
3.0	-471.32605	-471.3256	-471.29759	-471.40876	-471.32809

Table 5.2: The total energies of dodecane for MR BWCCSD with a posteriori (MR BWCCSD ap.) size extensivity corrections and MR MkCCSD using a 4 configuration model space and traditional CCSD and CCSD(T) methods at different C1 - C2 bond lengths (see text for details)

results at the equilibrium bond distance. The scale of the problems experienced by the perturbative SRCC methods is best epitomized by the CCSD(T) approach[86]. It has been documented that the CR-CC(2,3) method[86] is capable of providing very accurate results for single bond breaking processes. Several numerical studies have clearly showed that in such situations the CR-CC(2,3) approach produces nearly CCSDT results[86,93,94]. From the results it is clear that CR-CC(2,3) closely follows MRCC methods.

The results show that both MR BWCCSD with a posteriori correction and MR MkCCSD approaches give very similar results at all bond lengths. The largest difference in the energy occurs at 3.0 R_e and is 0.4 milli-Hartree. These differences between MR MkCCSD and MR BWCCSD can be attributed to the fact that a posteriori correction may in some cases overestimate exact energies.

5.2 Relative energies of naphthyne isomers

Naphthynes and their derivatives are transient species of considerable interest because they participate in a variety of organic reactions. They also play a role as antibiotics and anti-tumor agents[95]. As these diradical species are very short lived experimental data is scarce and computation is a relatively practical avenue to gain insight into the properties of these systems. In a previous study[66] Li and Paldus considered ten naphthyne isomers. They found that the singlet-triplet splitting in these systems depends on the distance between the radical centers. If the radical centers are located next to each other the singlet state forms a triple bond between the two carbons involved which stabilizes the state significantly. As a result the singlet-triplet splitting is found to be high. As the distance between the radical centers increases the singlet state is much less stabilized making the singlet-triplet gap smaller.

Li and Paldus used the RMR CCSD method for their calculations to account for near degeneracies. However, they had to reduce their model space based on perturbation theory and reduce the cc-pVDZ basis set by removing the hydrogen p-functions to limit the amount of computation needed. They also optimized the molecular structures for each of the spin states calculating adiabatic singlet-triplet splittings.

In this work we are interested in assessing the capabilities of multireference coupled cluster methods. Hence we selected the naphthyne isomers for which the singlet states have the strongest multireference character. From Li and Paldus it is clear that these isomers are 1,8-, 1,7-, 2,7- and 2,6-naphthynes in increasing order of multireference character. The triplet states were calculated using single reference calculations and the singlet states using a model space based on 2-electrons in 2-orbitals leading to either 2 or 4 spin-orbital reference configurations. We used the singlet geometries from Li and Paldus, and we computed the singlet and triplet states using different basis sets. We used cc-pVDZ[67] with Cartesian basis functions, cc-pVTZ[67] with spherical harmonic basis functions, and aug-cc-pVTZ*[67] where the p- and d-functions on hydrogen were removed. Of the resulting basis sets 180, 374 and 530 basis functions, respectively, were correlated for all naphthynes isomers. In all calculations the orbitals we used were obtained from Hartree-Fock calculations on the singlet state. In the calculations with the cc-pVTZ and aug-cc-pVTZ* basis sets the 20 core-electrons from carbon were not correlated, otherwise all electrons were correlated.

Our results are presented in Tables 5.3 and 5.4 and in Figure 5.2. Table 5.3 contains the singlet state total energies for all three basis sets. Table 5.4 lists the singlet-triplet splittings for the cc-pVDZ and cc-pVTZ basis sets. These splittings are also plotted as a function of different isomers in Figure 5.2. In the figure we have also included the CCSD and RMR CCSD results from Li and Paldus.

The data in Table 5.3 show that the total energy is lowered by 0.2 to 0.3 Hartree going from the cc-pVDZ to the cc-pVTZ basis set. The singlet-triplet splittings in Table 5.4 do not change by more than 3 kcal/mol. This is also clearly reflected in Figure 5.2. The figure also shows that our results are in good agreement with the results by Li and Paldus for the 2,7- and 2,6-naphthynes, differing by at most 2.5 kcal/mol. Our results differ from those by Li and Paldus by at most 7.6 kcal/mol for the 1,8- and 1,7-naphthynes. This is a consequence of the fact that we calculated vertical singlet-triplet splittings whereas Li and Paldus calculated adiabatic splittings. Comparing the bond lengths of the singlet and triplet equilibrium structures there are big differences for the 1,8- and 1,7-

Method	Basis	1,8-naphthyne	1,7-naphthyne	2,7-naphthyne	2,6-naphthyne
CCSD	cc-pVDZ	-383.37867	5.14	1.74	8.21
CCSD(T)	cc-pVDZ	-383.47775	-2.91	-7.93	-14.03
MR BWCCSD ap.	cc-pVDZ	-383.41124	0.81	-2.80	-2.39
MR MkCCSD	cc-pVDZ	-383.40292	2.04	-0.99	0.07
CCSD	cc-pVTZ	-383.64961	5.53	2.32	8.45
CCSD(T)	cc-pVTZ	-383.77324	-2.77	-7.59	-13.94
MR BWCCSD ap.	cc-pVTZ	-383.68415	0.40	-3.06	-4.22
MR MkCCSD	cc-pVTZ	-383.67476	1.79	-1.01	-1.37
MR BWCCSD ap.	cc-pVTZ*	-383.6988	-0.06	-3.45	-4.77
MR MkCCSD	cc-pVTZ*	-383.6886	n.c.	-1.57	-2.01

Table 5.3: Energies of the singlet state of naphthyne isomers with strong multireference character computed with different methods and different basis sets. The energies for 1,8-naphthyne are in Hartree, the energies of the other isomers are shown as $E(m, n\text{−naphthyne}) - E(1, 8\text{−naphthyne})$ *in kcal/mol.*

naphthynes and small differences for the 2,7- and 2,6-naphthynes. Hence the vertical excitations lead to larger differences compared to Li and Paldus for the 1,8- and 1,7-naphthynes than for the 2,7- and 2,6-naphthynes.

From Table 5.4 one can notice that the discrepancies in singlet-triplet splitting obtained with various SRCC methods are substantial. For example, for the 2,6-naphthyne isomer the difference between CCSD and CCSD(T) results reaches 44 kcal/mol when the cc-pVTZ basis set is used. It should be also emphasized that the signs of the CCSD(T) singlet-triplet splittings for all isomers considered in the Table 5.4 are consistently opposite when comparing to the CCSD estimates. This problem should be attributed to the problems with proper balancing of correlation effects for the singlet states of naphthyne isomers by the perturbative triples corrections. Usually, this is a manifestation of diverging nature of underlying SR perturbative expansion for the electronic energy.

Despite the fact that the static correlation effects for singlet states of naphthyne isomers are well described by the CMS(2,2), the MR BWCCSD and MR MkCCSD estimates lead to different predictions for the singlet-triplet gap. Although the differences are smaller than in the SR case, for the 2,6-naphthyne isomer the difference between cc-pVTZ MR BWCCSD and MR MkCCSD results amounts to 9 kcal/mol. For this isomer, in contrast to other isomers of Table 5.4, the calculated gaps are of the same sign. It is also interesting to notice that the

Method	Basis	1,8-naphthyne	1,7-naphthyne	2,7-naphthyne	2,6-naphthyne
CCSD	cc-pVDZ	19.04	24.43	24.52	32.43
CCSD(T)	cc-pVDZ	-0.97	-3.82	-5.51	-10.85
MR BWCCSD ap.	cc-pVDZ	-1.40	-0.34	-0.46	1.39
MR MkCCSD	cc-pVDZ	3.82	7.29	6.57	8.48
CCSD	cc-pVTZ	20.21	26.31	26.27	35.40
CCSD(T)	cc-pVTZ	-0.33	-2.80	-4.52	-8.67
MR BWCCSD ap.	cc-pVTZ	-1.47	-0.50	-0.78	1.05
MR MkCCSD	cc-pVTZ	4.43	6.79	7.16	9.79

Table 5.4: The singlet-triplet energy differences of naphthyne isomers given $^1E - ^3E$ in kcal/mol.

MR MkCCSD predictions, although much smaller, have the same signs as the CCSD results. We attribute the difference between the MR BWCCSD and MR MkCCSD results to the perturbative nature of the a posteriori correction.

5.3 Polycarbenes

Polycarbenes are polymers that contain one carbon atom with two unpaired electrons per monomer. These polymers are candidates for organic magnetic materials[96]. In a recent study by Mizukami et al.[97] the DMRG-SCF method[98] was used which includes orbital optimization. They chose an active space that is equivalent to CAS($8n+6$,$8n+6$) for the DMRG calculations. Here CAS(a,b) refers to a distinct electrons in b spatial orbitals, and n refers to the number of unsaturated carbon atoms. They compared their results against those from single reference methods as well as other multi reference methods with a CAS($2n$,$2n$) model space. They reported a distinctive difference in behavior of the total energy of polycarbenes obtained with single reference methods including DFT-B3LYP and CCSD(T) and multireference configuration methods such as DMRG-SCF, CASSCF and CASPT2. The former class of methods predicted that the spin-gap between the singlet state and the high spin state would increase with the system size, whereas the latter class predicted a decrease.

In our study we performed MRCC calculations on polycarbenes for $n = 1$ and 2 using C_{2v} symmetry. We used a CAS($2n$,$2n$) model space, with the 6-31G basis set[78] and the spherical harmonic cc-pVDZ[67] basis set, yielding 2 and 12 configurations in the model spaces for n equals 1 and 2 respectively. The 6-31G

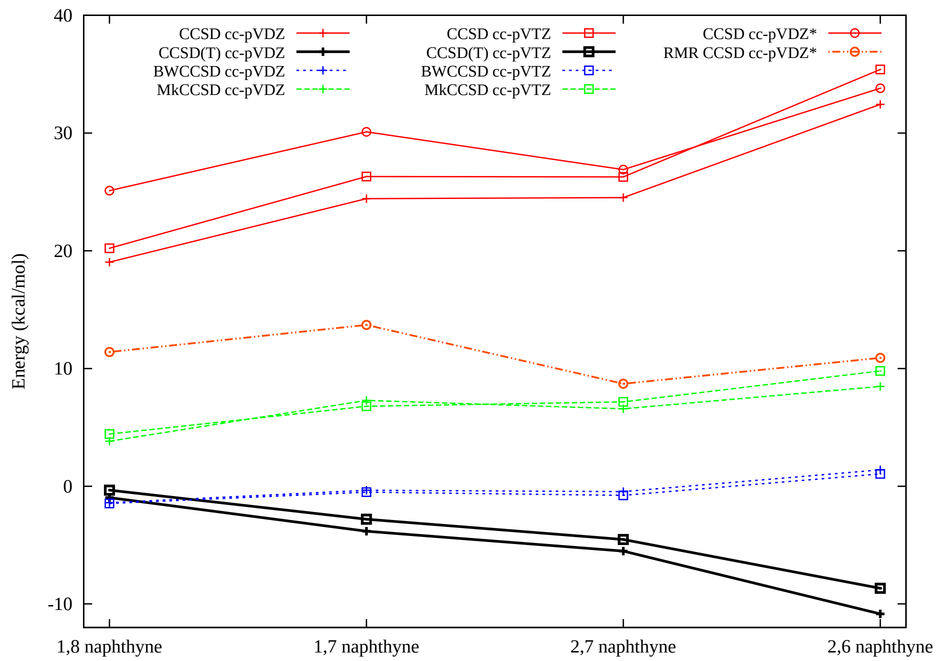

Figure 5.2: The singlet-triplet splittings for the 1,8-, 1,7-, 2,7- and 2,6-naphthynes computed using different basis sets and different methods.

basis set provides 124 and 188 correlated basis functions for the polycarbenes with n=1 and 2 respectively. The cc-pVDZ basis set provides 219 and 330 correlated basis functions. Except the core-electrons on carbon, all other electrons were included in the correlation treatment, i.e. 62 electrons for $n = 1$ and 94 electrons for $n = 2$ were correlated.

The results are presented in Tables 5.5 and 5.6. On the whole, our results confirm that single reference methods predict increasing spin gaps. Interestingly, these gaps are not so much affected by the level of theory used. For example the CCSD and CCSD(T) results are very similar. The basis set used, however, does

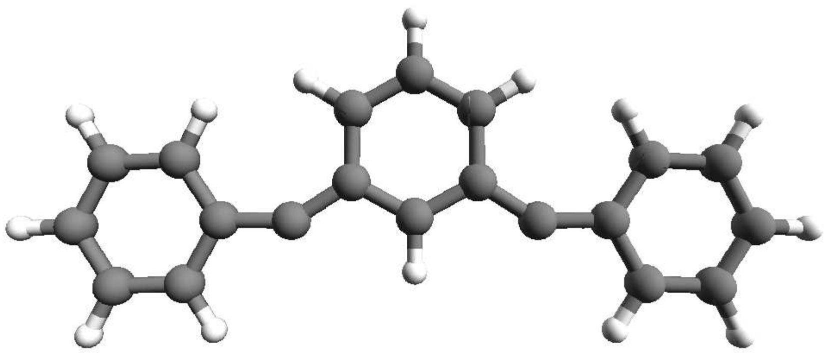

Figure 5.3: Structure of the polycarbene $n = 2$ molecule with two radical centers.

Method	Singlet-Triplet (n=1)		Singlet-Quintet (n=2)	
	6-31g	cc-pVDZ	6-31g	cc-pVDZ
CASSCF[97]	41.4		4.67	
DMRG-CASSCF[97]	34.5		11.08	
CASPT2[97]	30.2		5.24	
CCSD	52.62	41.80	109.87	88.10
CCSD(T)	50.79	40.20	105.60	84.41
MR BWCCSD ap.[a]			111.61	80.00
MR BWCCSD ap.	48.67	36.93	11.68	5.01
MR MkCCSD	50.5	39.3	12.5	n.c.

Table 5.5: Energies of the lowest singlet state and lowest high-spin state according to the given method of the polycarbene n=1 and n=2 in different basis sets. All core electrons have been frozen. Total energy is in Hartree, singlet - quintet gap in miliHartree.

[a] - The MR BWCCSD ap. energies for closed shell singlet state.

References	MR-BWCCSD ap. 6-31g	MR-MkCCSD 6-31g	MR-BWCCSD ap. cc-pVDZ	MR-MkCCSD cc-pVDZ
20	0.96	0.98	0.95	0.98
02	0.29	0.21	0.30	0.21

Table 5.6: The absolute values of the eigenvectors obtained from MRCC calculations in polycarbene($n = 1$).

have a significant impact. The spin gaps are much smaller if the cc-pVDZ basis set is used instead of the 6-31G basis set. The absolute value of eigenvectors obtained for polycarbene $n = 1$ is shown in Table 5.6. Regardless of basis set used, MR MkCCSD has the same values for eigen vectors and it is very close to that of MR BWCCSD.

The “closed-shell” singlet obtained using MR BWCCSD ap. for $n = 2$ are very close to the results obtained from CCSD. The “open-shell” singlet obtained using MR BWCCSD and MR MkCCSD clearly shows it is much lower in energy than the “closed-shell” ones. Our results also show that the spin gaps decrease with chain length if multireference methods are used. The reduction in gap is 31.92 milli-Hartrees from $n = 1$ to $n = 2$ for MR BWCCSD method using cc-pVDZ basis set. We also see that the MR MkCCSD and MR BWCCSD results are very similar. Furthermore, for the multireference calculations we also see that the basis set has a significant impact on the spin gaps. Another interesting aspect to note is that the multireference coupled cluster calculations predict a steeper reduction in the spin gap than DMRG-SCF or CASPT2. However the coupled cluster results are reasonably well aligned with the CASSCF results.

6. Universal state-selective correction

This chapter has been published in Journal of Chemical Physics: Jiří Brabec, Hubertus J. J. van Dam, Jiří Pittner and Karol Kowalski, Journal of Chemical Physics (2012), 136, 12410

6.1 Basics of the USS-MRCC approach

The universal state-selective (USS) functional has been proposed by K. Kowalski[99]. It can be used to define the non-iterative corrections to approximate MR-CC approaches. In particular, he has shown that in the proposed framework the essential part of dynamic correlation can be encapsulated in the so-called correlation Hamiltonian, which in analogy to the effective Hamiltonian, is defined in the model space ($\mathcal{M}_0$).

In this chapter we will discuss the simplest USS corrections to the SS-MRCCSD approaches. The so-called USS(2) approximation, which is derived from the USS-MRCC expansion utilizing the Various Trial Wave functions (VTW) strategy (see Ref.[99]), uses the one- and two-body components of the reference-specific similarity transformed Hamiltonians in defining the correction. Our goal is to explore mainly two questions: (1) to what extent the USS(2) can minimize the energy differences caused by a different choice of the sufficiency conditions for models employing single and double excitations (this feature will be referred to as the universality of the USS approach) and (2) to what extent the USS(2) formalism can eliminate the unwanted features of the BW-MRCCSD formalism associated with its lack of size-extensivity.

The asymmetric USS-MRCC functional[99] is defined by the following expression:

$$E(\bar{\Psi}^{(1)},\ldots,\bar{\Psi}^{(M)}) = \sum_{\mu=1}^{M} \frac{\langle\bar{\Psi}^{(\mu)}|He^{T^{(\mu)}}|\Phi_\mu\rangle}{\langle\bar{\Psi}^{(\mu)}|e^{T^{(\mu)}}|\Phi_\mu\rangle} c_\mu^2 \,, \tag{6.1}$$

where $\langle\bar{\Psi}_\mu|$ are reference-specific trial wavefunctions. Once the trial wavefunctions are replaced by the exact wavefunction then the value of the USS-MRCC is equal to the exact (FCI) energy. As discussed in Ref.[99] the form of the USS-MRCC functional can be greatly simplified by using a reference-dependent parametrization of the trial wave function $\langle\bar{\Psi}_{\mu}|$, i.e.,

$$\langle\bar{\Psi}^{(\mu)}| = \sum_{\nu=1}^{M} d_{\nu}^{(\mu)} \langle\Phi_{\nu}|(1+\Lambda_{\nu}^{(\mu)})e^{-T^{(\mu)}} \ , \tag{6.2}$$

where $\Lambda_{\nu}^{(\mu)} = P_{\nu}\Lambda^{(\mu)}Q$ are "de-excitation" correlation operators, and employing normalization conditions for the $d_{\nu}^{(\mu)}$ coefficients,

$$d_{\mu}^{(\mu)} = c_{\mu} \ . \tag{6.3}$$

Using (6.2) and (6.3) the USS-MRCC functional takes a simple form

$$E = E^{(A)} + \sum_{\mu,\nu=1}^{M} \Delta_{\nu}^{(\mu)} H_{\nu\mu}^{\text{eff}} c_{\mu} + \sum_{\mu,\nu=1}^{M} d_{\nu}^{(\mu)} \langle\Phi_{\nu}|\Lambda_{\nu}^{(\mu)}\bar{H}^{(\mu)}|\Phi_{\mu}\rangle c_{\mu} \ . \tag{6.4}$$

The normalization condition (6.3) fully determines the normalization of the $\mathbf{d}^{(\mu)}$ vector defined as $(d_1^{(\mu)},\ldots,d_{\nu}^{(\mu)},\ldots,d_M^{(\mu)})^T$. In the above equation $E^{(A)}$ stands for the approximate MRCC energy obtained by diagonalizing the effective Hamiltonian (H^{eff}), the $\{c_{\mu}\}_{\mu=1}^{M}$ coefficients are the components of the right eigenvectors of H^{eff} corresponding to the $E^{(A)}$ energy, while the $\Delta_{\nu}^{(\mu)}$ coefficients are defined as

$$d_{\nu}^{(\mu)} = \bar{c}_{\nu} + \Delta_{\nu}^{(\mu)} \ , \tag{6.5}$$

where $\{\bar{c}_{\mu}\}_{\mu=1}^{M}$ coefficients correspond to the components of the left eigenvector of H^{eff} to energy $E^{(A)}$. The $\bar{H}^{(\mu)}$ operators in (6.4) are the reference-specific similarity transformed Hamiltonians, i.e.,

$$\bar{H}^{(\mu)} = e^{-T^{(\mu)}} H e^{T^{(\mu)}} \ . \tag{6.6}$$

In deriving (6.4) two additional normalization conditions have been assumed:

$$\sum_{\mu=1}^{M} c_{\mu}^2 = 1 \tag{6.7}$$

and

$$\sum_{\mu=1}^{M} c_{\mu}\bar{c}_{\mu} = 1 \ . \tag{6.8}$$

The matrix elements $\langle\Phi_{\nu}|\Lambda_{\nu}^{(\mu)}\bar{H}^{(\mu)}|\Phi_{\mu}\rangle$ define the matrix representation of the correlation Hamiltonian ($H^{\rm corr}$)

$$H^{\rm corr} = \sum_{\mu,\nu=1}^{M} P_\nu \Lambda_\nu^{(\mu)} \bar{H}^{(\mu)} P_\mu \; . \tag{6.9}$$

Using the above notation the energy can be re-expressed in a compact form

$$E = E^{(A)} + \sum_{\mu,\nu=1}^{M} \Delta_\nu^{(\mu)} H_{\nu\mu}^{\rm eff} c_\mu + \sum_{\mu,\nu=1}^{M} d_\nu^{(\mu)} H_{\nu\mu}^{\rm corr} c_\mu \; . \tag{6.10}$$

The specific parameterizations of the trial wave functions can lead to desired properties of the USS-MRCC functional. For example, for the complete model space (CMS) the double-exponential parameterizations of the trial wavefunctions

$$\langle \bar{\Psi}^{(\mu)} | = \sum_{\nu=1}^{M} d_\nu^{(\mu)} \langle \Phi_\nu | e^{S_\nu^{(\mu)}} e^{-T^{(\mu)}} \tag{6.11}$$

result in a connected form of the correlation Hamiltonian (assuming that the $T^{(\mu)}$ and $S_\nu^{(\mu)}$ operators are represented by connected diagrams only). Once the model space properly dissociates these conditions assure additive separability of the USS-MRCC functional in the non-interacting subsystem limit (NSL) (see Ref.[99] for details).

In the single reference limit the Λ correction given by equations (6.9) and (6.10) can be easily related to the very efficient single-reference (SR) Λ corrections derived from the unsymmetrical CC functional[100–108] or from the method of moments of CC equations (MMCC).[86,93] For example, when T and Λ operators are obtained in the single-reference CCSD calculations (CC with singles and doubles)[109], then the simplest correction which includes the effect of triply excited configurations is given by the $\langle \Phi | \tilde{\Lambda}_3 \bar{H} | \Phi \rangle$ formula (see[99]) where $\tilde{\Lambda}_3$ is a perturbative approximation to the exact Λ_3 operator. This form is identical with the form of several very efficient SR non-iterative corrections due to triples.[86,93,100–108]

6.2 Non-iterative USS-MRCC corrections

Expansion (6.10) can be utilized in building the non-iterative corrections to the approximate MRCC energy $E^{(A)}$. The two main features of this approach should be highlighted here: (1) If reference-specific trial wave functions are determined without invoking the sufficiency conditions, then in calculating energy correction (second and third terms on the right hand side of Eq.(6.10)) the whole correction can be obtained without calculating cumbersome coupling terms present in some of the MRCC approaches. This is a consequence of the fact that the correction is expressed only in terms of elements of similarity transformed Hamiltonians ($\bar{H}^{(\mu)}$) and $\Lambda_\nu^{(\mu)}$ operators. (2) In order to calculate many-body components ($\bar{H}_i^{(\mu)}$), single-reference implementations can be exploited. In particular, these terms can be used to express higher-order components of the $\Lambda_\nu^{(\mu)}$ (or $S_\nu^{(\mu)}$) operators (see Ref. [99] for details). (3) Simple hierarchy of approximations based on incorporating higher order terms in $\Lambda_\nu^{(\mu)}/S_\nu^{(\mu)}$ and $\bar{H}^{(\mu)}$ operators can be designed within the USS-MRCC approach.

Various approximate strategies have been extensively discussed in Ref. [99], including corrections to the MRCCSD theories. These corrections usually involve the inclusion of higher order effects in the correlation Hamiltonian, where we assume equal rank approximations for the $\Lambda_\nu^{(\mu)}$ and $\bar{H}^{(\mu)}$ operators, i.e,

$$\begin{aligned} H^{\text{corr}} \simeq H^{\text{corr}}(m_A) \;\; &= \textstyle\sum_{\nu,\mu=1}^{M} P_\nu(\Lambda_{\nu,1}^{(\mu)} + \ldots + \Lambda_{\nu,m_A}^{(\mu)}) \\ &\times(\bar{H}_1^{(\mu)} + \ldots + \bar{H}_{m_A}^{(\mu)})P_\mu \;, \end{aligned} \tag{6.12}$$

where m_A is predefined excitation level. In this work we will apply the simplest form of the USS-MRCC correction to the BW-MRCCSD and Mk-MRCCSD approaches defined by m_A=2 condition. One of our motivations is to explore if, and to what extent, the USS corrections employing the same excitation level (m_A=2) as the approximate MRCC methods (MRCCSD) can minimize the energy differences due to a different form of the sufficiency conditions used to determine singly and doubly excited amplitudes. For the BW-MRCCSD approach several algorithms have been proposed in order to eliminate/reduce the inherent size-extensivity errors. In particular, we will relate the $H^{\text{corr}}(m_A = 2)$-based expansion to the a posteriori corrections to the BW-MRCCSD approach. [15,45,110] For $m_A = 2$ the USS-MRCC energies ($E^{\text{USS(2)-BW-MRCCSD}}$ and $E^{\text{USS(2)-Mk-MRCCSD}}$ or $E^{\text{USS(2)-BW}}$ and $E^{\text{USS(2)-Mk}}$ for short) are given by the following expressions:

$$\begin{aligned} E^{\text{USS(2)-X}} \;\; &= E^{\text{X}} + \textstyle\sum_{\mu,\nu=1}^{M} \Delta_\nu^{(\mu)} H_{\nu\mu}^{\text{eff}} c_\mu + \sum_{\mu,\nu=1}^{M} d_\nu^{(\mu)} \langle\Phi_\nu|(\Lambda_{\nu,1}^{(\mu)} + \Lambda_{\nu,2}^{(\mu)}) \\ &\times(\bar{H}_1^{(\mu)} + \bar{H}_2^{(\mu)})|\Phi_\mu\rangle c_\mu \;,\; (\text{X} = \text{BW}, \text{Mk}) \;, \end{aligned} \tag{6.13}$$

where $H_{\nu\mu}^{\text{eff}}$ are the matrix elements of the corresponding effective Hamiltonian. The $\Lambda_{\nu,i}^{(\mu)}$ (or $S_{\nu,i}^{(\mu)}$) amplitudes can be obtained be solving left variants of the Bloch equations (LBEs; see Ref. [99]) for the $\bar{H}^{(\mu)}$ Hamiltonian.

It is important to realize that once the Λ operators are obtained from solving LBEs, the USS(2) $-$ X corrections should be equal to zero for two-electron systems. This results from (1) the definition of the USS functional (6.1), (2) the exactness of the Λ-parametrized wavefunction (6.2) in two-electron limit, and (3) the fact that the BW-MRCCSD and Mk-MRCCSD approaches provide exact energies in two-electron case. It is also interesting to analyze the form of the USS correction when the exponential parametrization of the trial wavefunction (6.11) is based on singly and doubly excited cluster amplitudes. The USS(2) energy takes then the form

$$\begin{aligned} E^{\mathrm{USS(2)_S-X}} &= E^{\mathrm{X}} + \textstyle\sum_{\mu,\nu=1}^{M} \Delta_\nu^{(\mu)} H_{\nu\mu}^{\mathrm{eff}} c_\mu + \sum_{\mu,\nu=1}^{M} d_\nu^{(\mu)} \langle \Phi_\nu | (e^{S_{\nu,1}^{(\mu)}+S_{\nu,2}^{(\mu)}} - 1) \\ &\times (\bar{H}_1^{(\mu)} + \bar{H}_2^{(\mu)}) |\Phi_\mu\rangle c_\mu \ , \ (\mathrm{X{=}BW,\ Mk}) \ , \end{aligned} \tag{6.14}$$

where in analogy to Eq.(6.13) only $\bar{H}_1^{(\mu)}$ and $\bar{H}_2^{(\mu)}$ parts of $\bar{H}^{(\mu)}$ are retained. We will assume that all $S_{\nu,i}^{(\mu)}$ $(i = 1, 2)$ operators are represented by connected diagrams only. Although this expansion has a different form than the expansion (6.13) (the excitation level of $(e^{S_{\nu,1}^{(\mu)}+S_{\nu,2}^{(\mu)}} - 1)$ operators may be higher than the maximum excitation level of $\Lambda_\nu^{(\mu)}$'s (m_A)), one can show that when ensemble of G two-electron systems (for example H_2 molecules) are considered and once the model space for the composite system properly dissociates into the tensor product of the model spaces for individual H_2 molecules $(\mathcal{M}_0(H_2(i)),\ i = 1, ..., G)$

$$\mathcal{M}_0 \stackrel{\mathrm{NSL}}{\rightarrow} \bigotimes_{i=1}^{G} \mathcal{M}_0(H_2(i)) \ , \tag{6.15}$$

then the energy of the $(H_2)_G$ system in the NSL is equal to the sum of exact energies of the non-interacting H_2 molecules, i.e.,

$$E^{\mathrm{USS(2)_S-X}} \stackrel{\mathrm{NSL}}{\rightarrow} G \times E_{H_2} \ . \tag{6.16}$$

This feature is identical with the separability properties of the single-reference MMCC energies.[111] We can also utilize the relations between the Λ and S operators for the exact limit (all excitations are included) to approximate the Λ operators through S ones. The simplest approximation (consistent with the form of expansion (6.13)) can be obtained by approximating Λ's by linear S terms, i.e., $\Lambda_{\nu,i}^{(\mu)} = S_{\nu,i}^{(\mu)}$ $(i = 1, 2)$.

6.3 Working equations for the USS(2) approximations

The algebraic form of the USS corrections requires solving the set of reference-specific LBEs. As mentioned earlier, in the single-reference limit these equations reduce to the standard single-reference Λ equations. A strong argument in favor of utilizing the Λ operators in the context of building single-reference non-iterative corrections have been made by Taube and Bartlett in Ref.[108](a), where it was demonstrated that as a consequence of smaller magnitudes of the Λ amplitudes compared to the corresponding T ones, the non-iterative corrections based on the former set of amplitudes provided more balanced description of correlation effects especially in the strong quasidegeneracy region.

Although solving LBEs is possible for low-rank Λ operators, in this work, in order to explore general features of the USS(2) correction, we will eliminate this step by assuming that $\Lambda^{(\mu)}_{\nu,i}$ $(i=1,2)$ (or $S^{(\mu)}_{\nu,i}$) operators can be approximated by the $T^{(\nu)\,+}_i$ operators, i.e.,

$$\langle\Phi_\nu|\Lambda^{(\mu)}_{\nu,i} \simeq \langle\Phi_\nu|T^{(\nu)\;+}_i\;. \tag{6.17}$$

A similar approximation has been used in calculating MRCC non-iterative correction of Ref.[112]. To justify the form of approximation (6.17) let us assume that the $\Lambda^{(\mu)}_{\nu,i}$ operators can be estimated perturbatively from the direct terms defining the $\bar{H}^{(\mu)}$-based LBE-Λ approach (see Eqs.(35) and (36) of Ref.[99]), i.e.,

$$\langle\Phi_\nu|\Lambda^{(\mu)}_{\nu,1} \;\simeq\; \langle\Phi_\nu|\bar{H}^{(\mu)}R^{(0)}_1(\nu) \tag{6.18}$$

$$\langle\Phi_\nu|\Lambda^{(\mu)}_{\nu,2} \;\simeq\; \langle\Phi_\nu|\bar{H}^{(\mu)}R^{(0)}_2(\nu) \tag{6.19}$$

$$\tag{6.20}$$

where the resolvents $R^{(0)}_n(\mu)$ are defined as

$$R^{(0)}_n(\nu) = \sum_{i_1<\ldots<i_n;a_1<\ldots<a_n} \frac{|\Phi^{a_1\ldots a_n}_{\nu,i_1\ldots i_n}\rangle\langle\Phi^{a_1\ldots a_n}_{\nu,i_1\ldots i_n}|}{\epsilon(\nu)_{i_1}+\ldots+\epsilon(\nu)_{i_n}-\epsilon(\nu)_{a_1}\ldots-\epsilon(\nu)_{a_n}}\;, \tag{6.21}$$

and the summation over occupied $(i_1,i_2,\ldots)$ and unoccupied $(a_1,a_2,\ldots)$ spinorbitals defined with respect to the $|\Phi_\nu\rangle$ reference is limited only to those configurations $|\Phi^{a_1\ldots a_n}_{\nu,i_1\ldots i_n}\rangle$ which belong to the orthogonal complement of the model space. The $\epsilon(\nu)$'s can be considered as the diagonal elements of the Fock matrix corresponding to the ν-th reference (although other choices are possible). To simplify

Eqs.(6.18-6.19) even further, we will approximate the $\bar{H}^{(\mu)}$ operator as

$$\bar{H}^{(\mu)} \simeq H + [H, T^{(\mu)}] \ , \tag{6.22}$$

and replace $\langle\Phi_\nu|HR_i^{(0)}(\nu)$ $(i=1,2)$ by $T_i^{(\nu)(1)\,+}$ (the $\langle\Phi_\nu|HR_i^{(0)}(\nu)$ $(i=1,2)$ terms represent the lowest order estimates of the $T^{(\nu)\,+}$ operators). This leads to the following form for the $\Lambda_{\nu,i}^{(\mu)}$ $(i=1,2)$ operators

$$\langle\Phi_\nu|\Lambda_{\nu,1}^{(\mu)} \ \simeq \ \langle\Phi_\nu|T_1^{(\nu)\,+} + \langle\Phi_\nu|[H, T^{(\mu)}]R_1^{(0)}(\nu) \ , \tag{6.23}$$

$$\langle\Phi_\nu|\Lambda_{\nu,2}^{(\mu)} \ \simeq \ \langle\Phi_\nu|T_2^{(\nu)\,+} + \langle\Phi_\nu|[H, T^{(\mu)}]R_2^{(0)}(\nu) \ , \tag{6.24}$$

which means that in the low-order of theory, the Λ operators can be approximated by the corresponding T^+ operators. In a similar way, if in the effective Hamiltonian for LBE corresponding to the μ-th reference (see Ref.[99])

$$H_{LBE}^{eff\,(\mu)} = P\sum_{\nu=1}^{M}(1 + \Lambda_{\nu,1}^{(\mu)} + \Lambda_{\nu,2}^{(\mu)})\bar{H}^{(\mu)}P \tag{6.25}$$

we approximate $\Lambda_{\nu,i}^{(\nu)}$ by $T_i^{(\nu)\,+}$ $(i=1,2)$ operators and if we use Eq.(6.22) then one can see the lowest order contributions

$$P\sum_{\nu=1}^{M}(1 + T_{\nu,1}^{+} + T_{\nu,2})^{+}HP \tag{6.26}$$

can be viewed as the lowest order approximation to the Hermitian conjugate of the effective Hamiltonian corresponding to a given MRCCSD approach. This means that in the lowest order, the left eigenvectors to the $H_{LBE}^{eff\,(\mu)}$ can be approximated by the right eigenvectors of $H^{\rm eff}$, i.e.,

$$d_\nu^{(\mu)} = c_\nu \ . \tag{6.27}$$

On the basis of the normalization conditions (6.7) and (6.8) this choice of $d_\nu^{(\mu)}$s leads to zeroing $H^{\rm eff}$ dependent term in the USS(2) correction. Now one can rewrite $E^{\rm USS(2)-X}$ (X=BW,Mk) in a simple form

$$E^{\rm USS(2)-X} = E^{\rm X} + \sum_{\mu,\nu=1}^{M} c_\nu\langle\Phi_\nu|(T_1^{(\nu)} + T_2^{(\nu)})^{+}(\bar{H}_1^{(\mu)} + \bar{H}_2^{(\mu)})|\Phi_\mu\rangle c_\mu \ (\text{X=BW,Mk}). \tag{6.28}$$

Following the diagonal corrections of Ref.[99] we will also consider the "diagonal" part of the formula (6.28)

$$E^{\mathrm{D-USS(2)-X}} = E^{\mathrm{X}} + \sum_{\mu=1}^{M} \langle \Phi_\mu | (T_1^{(\mu)} + T_2^{(\mu)})^+ (\bar{H}_1^{(\mu)} + \bar{H}_2^{(\mu)}) | \Phi_\mu \rangle c_\mu^2 \text{ (X=BW,Mk)}. \tag{6.29}$$

Assuming the availability of the $\bar{H}_1^{(\mu)}$ and $\bar{H}_2^{(\mu)}$ amplitudes (for example from the last BW-MRCCSD or Mk-MRCCSD iteration) the cost of forming the USS(2) correction is proportional to $M^2 \times n_o^2 n_u^2$, where n_o and n_u represent the number of occupied and unoccupied spinorbitals in reference $|\Phi_\mu\rangle$. In the D-USS(2) approach this numerical overhead is reduced to $M \times n_o^2 n_u^2$. In the following we will estimate the accuracy of USS(2) and D-USS(2) approaches when applying to BW-MRCCSD and Mk-MRCCSD methods.

6.4 Relationships between USS(2) and a posteriori corrections

Historically, the BW-MRCC methods have been the first MRCC formulations of the State-Specific type. The working equations for the BW-MRCC theory are obtained by introducing the JM Ansatz (1.23) into the Schrödinger equation and postulating the following form of the sufficiency conditions for the cluster amplitudes

$$(E^{(A)} - H^{\mathrm{eff}}_{\mu\mu}) \langle \Phi_\vartheta(\mu) | e^{T^{(\mu)}} | \Phi_\mu \rangle = \langle \Phi_\vartheta(\mu) | H_N(\mu) e^{T^{(\mu)}} | \Phi_\mu \rangle_{C+DC,L} \, , \tag{6.30}$$

where $\langle \Phi_\vartheta(\mu)|$ are excited configurations corresponding to the excitations used to define cluster operator $T^{(\mu)}$. The subscript $C + DC, L$ designates all connected diagrams plus all linked, but disconnected diagrams. It is well known that the approximate BW-MRCC approaches are not size-extensive. In order to address this issue several corrections to the genuine BW-MRCC energies have been proposed.[15,45,110] In the approach proposed by Hubač et al.[110] the converged cluster operator $T^{(\mu)}$ is corrected by performing another iteration involving only connected diagrams on the right hand side of Eq.(6.30). This procedure can be equivalently represented by adding the correction $\Delta T^{(\mu)}$ to the converged $T^{(\mu)}$ operator, i.e.,

$$\bar{T}^{\mu} = T^{(\mu)} + \Delta T^{(\mu)} \, , \tag{6.31}$$

where $\bar{T}^{(\mu)}$ is the corrected cluster operator. The formula for $\Delta T^{(\mu)}$ is as follows

$$\Delta T^{(\mu)}|\Phi_\mu\rangle = R_0^{(\mu)}(He^{T^{(\mu)}})_C|\Phi_\mu\rangle = R_0^{(\mu)}\bar{H}^{(\mu)}|\Phi_\mu\rangle \ . \tag{6.32}$$

The a posteriori corrected BW-MRCCSD (BW-MRCCSD-ap) energies are obtained by diagonalizing the effective Hamiltonian employing the corrected cluster operators $\bar{T}^{(\mu)}$. The matrix elements of this a posteriori corrected effective Hamiltonian ($H^{\mathrm{eff-ap}}$) for the CMS are given by the expression:

$$\begin{aligned} H_{\nu\mu}^{\mathrm{eff-ap}} &= \langle\Phi_\nu|He^{\bar{T}^{(\mu)}}|\Phi_\mu\rangle & (6.33)\\ &= \langle\Phi_\nu|He^{T^{(\mu)}}|\Phi_\mu\rangle + \langle\Phi_\nu|He^{T^{(\mu)}}(e^{\Delta T^{(\mu)}}-1)|\Phi_\mu\rangle \ . & (6.34)\end{aligned}$$

The first element on the right hand side of Eq.(6.34) refers to the matrix element of the genuine effective Hamiltonian matrix utilized in Eq.(6.13), with $E^{\mathrm{BW-MRCCSD}}$ energy being one of its eigenvalues. The second term can be considered as a small correction striving to reduce the size-extensivity error. With this assumption we can approximate the eigenvalue ($E^{\mathrm{BW-MRCCSD-ap}}$) of the $H^{\mathrm{eff-ap}}$ matrix as $\sum_{\nu\mu}\bar{c}_\nu H_{\nu\mu}^{\mathrm{eff-ap}}c_\mu$, i.e.,

$$\begin{aligned} E^{\mathrm{BW-MRCCSD-ap}} &\simeq \sum_{\nu\mu}\bar{c}_\nu(H_{\nu\mu}^{\mathrm{eff}} + \langle\Phi_\nu|He^{T^{(\mu)}}(e^{\Delta T^{(\mu)}}-1)|\Phi_\mu\rangle)c_\mu & (6.35)\\ &\simeq E^{\mathrm{BW-MRCCSD}} + \sum_{\nu\mu}c_\nu\langle\Phi_\nu|H\Delta T^{(\mu)}|\Phi_\mu\rangle c_\mu & (6.36)\\ &= E^{\mathrm{BW-MRCCSD}} + \sum_{\nu\mu}c_\nu\langle\Phi_\nu|HR_0^{(\mu)}\bar{H}^{(\mu)}|\Phi_\mu\rangle c_\mu & (6.37)\\ &= E^{\mathrm{BW-MRCCSD}} + \sum_{\nu\mu}c_\nu\langle\Phi_\nu|(\tilde{T}_1^{(\nu),[1]}(\mu)+\tilde{T}_2^{(\nu),[1]}(\mu))^+ \\ &\quad \times(\bar{H}_1^{(\mu)}+\bar{H}_2^{(\mu)})|\Phi_\mu\rangle c_\mu \ , & (6.38)\end{aligned}$$

where

$$(\tilde{T}_1^{(\nu),[1]}(\mu)+\tilde{T}_2^{(\nu),[1]}(\mu))|\Phi_\nu\rangle = R_0^{(\mu)}H|\Phi_\nu\rangle \ . \tag{6.39}$$

In Eq.(6.36) we approximated the components of the left eigenvector of H^{eff} by the corresponding components of the right eigenvector (i.e., $\bar{c}_\nu = c_\nu$). Additionally we simplified the form of the $\langle\Phi_\nu|He^{T^{(\mu)}}(e^{\Delta T^{(\mu)}}-1)|\Phi_\mu\rangle$ term, approximating it by $\langle\Phi_\nu|H\Delta T^{(\mu)}|\Phi_\mu\rangle$. The first-order $\tilde{T}_i^{(\nu),[1]}(\mu)$ ($i=1,2$) cluster operators contain those single and double excitations with respect to the $|\Phi_\nu\rangle$ reference which also belongs to the manifold of singly and doubly excitations with respect to the $|\Phi_\mu\rangle$ reference. One should notice that the energy denominators are pertinent to the

$|\Phi_\mu\rangle$ reference configuration and expression (6.39) can be considered as a approximation to the true first-order estimates of the $T^{(\nu)}$ amplitudes. The Eq.(6.38) bears a great resemblance to the USS(2) expression given by Eq.(6.28) (although the $T^{(\nu)}$ amplitudes in the USS formula are not μ-dependent). The main difference, however, lies in the perturbative estimates of the $T_1^{(\nu)}$ and $T_2^{(\nu)}$ operators entering Eq.(6.38). This is a main difference between a posteriori corrected BW-MRCCSD and USS(2)-BW-MRCCSD energies. The presence of the perturbative estimates $\tilde{T}_i^{(\nu),[1]}(\mu)$ $(i = 1, 2)$ in the a posteriori corrected BW-MRCCSD theory can lead to significant problems in the presence of intruder states and accidental degeneracies of the perturbative denominators. On the other hand, we expect that the USS(2)-BW-MRCCSD correction will be free of these drawbacks. This issue will be illustrated and discussed in the forthcoming sections.

6.5 Computational details

The USS(2) and D-USS(2) corrections (6.28) and (6.29) have been implemented using TINYMRCC suite of MRCC codes developed in Dr. Pittner's group.[113] The D-USS(2) corrections have also been implemented using Tensor Contraction Engine (TCE).[114,115] Both D-USS(2) implementations have been cross-tested for the correctness of the resulting energies. The larger studies with the D-USS(2) approach have been performed with the TCE implementation. In our calculations we included up to two-body terms in the BW-MRCCSD and Mk-MRCCSD effective Hamiltonian operators. The BW-MRCCSD and Mk-MRCCSD implementations are available in TINYMRCC and in the development version of NWChem.[80,116] Since the inclusion of the off-diagonal terms in the correlation Hamiltonian using TCE capabilities although possible is quite challenging, the full USS(2) calculations have been performed with TINYMRCC implementations. Due to the fact that TINYMRCC has been developed rather for testing various MRCC theories with high-rank cluster operators than to enable calculations for large systems, the USS(2) calculations were performed only for small system (H8 and BeH_2). Nevertheless, we believe that the available USS(2) results are capable of providing valuable insigths into the performance of the USS(2) correction.

In the following sections we use the $(n_{act.el.}, n_{act.orb.})$ notation to define complete model spaces. The $n_{act.el.}$ and $n_{act.orb.}$ refer to the number of active electrons and active orbitals, respectively. In our calculations, except for the H8 case where we used CASSCF orbitals, RHF orbitals were used. The discussion of the USS(2) performance is preceded by the numerical size-consistency tests for the CH_2-Ne and C_2H_4-Ne systems. In this chapter we discuss the results obtained for several benchmark systems: (1) H8 model[4] in the 6-31G basis set,[78] (2) BeH_2 molecule in the DZ basis set,[117] (3) water molecule in the spherical cc-pVDZ basis set,[67] (4) nitrogen molecule in the DZ basis set and spherical cc-pVDZ basis sets, and (5) Be_3 molecule in the ANO basis set.[118,119] We used (2,2) model spaces for H8 and BeH_2 models. For the H_2O molecule (2,2), (4,4), and (6,6) complete model spaces were used. For N_2 (4,4) and (6,6) CMSs were used, while for the Be_3 case we employed (2,2), (4,4), and (6,6) CMSs. Unfortunately, due to its size we could not use the most justified (6,12) model space that correlates $2s$ and $2P$ shells in the Be atoms.

The calculations for H8 with TINYMRCC have been performed with the C_1 symmetry (however, the RHF calculations with D_{2h} show that the model spaces include $1b_{1g}$ and $2a_g$ orbitals). For BeH_2 active orbitals correspond to the $1b_1$ and $3a_1$ orbitals (for larger values of the geometry x parameter we used the $3a_1$ and $1b_2$ orbitals). For the water molecule $3a_1$, $1b_2$, $4a_1$, and $4a_1$ were assumed to be active for (4,4) model space calculations (the (2,2) model space used $1b_2$ and $4a_1$ active orbitals). In the calculations for N_2 active orbitals correspond to $1b_{2u}$, $1b_{1u}$, $3a_g$, $1b_{1g}$, $1b_{2g}$, and $3b_{3u}$ orbitals. The Be_3 system employed the following active orbitals: $3a_g$, $2b_{3u}$, $4a_g$, $3b_{3u}$, $1b_{1u}$, $1b_{2u}$. In our calculations molecular orbitals have not been explicitly localized. Our implementation includes only scalar, one-, and two-body components of the effective Hamiltonian.

6.6 Results and discussion

Since the approximate formulations of the BW-MRCC theory are not size-consistent it is important to estimate the order of the size-consistency errors made in the corresponding calculations. For this purpose we used the CH_2Ne and C_2H_4 systems (see Table 6.1), where the CH_2/C_2H_4 and Ne components are at large separation. In Table 6.1 we collected the results obtained with the D-USS(2) and BW-MRCCSD-ap methods for the cc-pVQZ basis sets.[67] As a measure of the size-consistency error we used Δ parameter defined as $\Delta = E(CH_2 + Ne) - E(CH_2) - E(Ne)$. The separation between the CH_2 and Ne system was chosen to be equal 20 Å. For the cc-pVQZ basis set the size-consistency error of the uncorrected BW-MRCCSD approach: 0.797 milliHartree, is practically eliminated by the D-USS(2) and BW-MRCCSD-ap methods. The size-consistency errors for the D-USS(2) and BW-MRCCSD-ap approaches are smaller than 1 microHartree. For the singlet state of twisted ethylene($R_C - C$=1.451 Å, R_{C-H}=1.080

Å, $\alpha_{H-C-H} = 117.1°$-Ne system (see Table 6.1) the size-consistency error of the D-USS(2) method remains small (0.009 milliHartree). At the same time the BW-MRCCSD-ap and uncorrected BW-MRCCSD errors are equal to 0.031 and 1.753 milliHartree.

System	D-USS(2)-BW	BWSD ap.	BWSD
$E(CH_2)$CMS(2,2)	-39.067720	-39.066922	-39.064033
E(Ne)	-128.878546	-128.878546	-128.878546
Total $E(CH_2)$CMS(2,2) +E(Ne)	-167.946266	-167.945468	-167.942579
$E(CH_2$+Ne)CMS(2,2)	-167.946266	-167.945468	-167.941782
Δ	0.000	0.000	0.797
$E(C_2H_4)$CMS(2,2)	-78.250516	-78.248395	-78.240848
E(Ne)	-128.679637	-128.679637	-128.679637
Total $E(C_2H_4)$CMS(2,2)+E(Ne)	-206.930153	-206.928032	-206.920485
$E(C_2H_4$+Ne)CMS(2,2)	-206.930144	-206.928001	-206.918732
Δ	0.009	0.031	1.753

Table 6.1: MRCC size-consistency tests for the CH_2Ne system in the cc-pVQZ basis set,[67] and for the singlet state of the C_2H_4Ne system in the cc-pVDZ basis set.[67] The Δ parameters are defined as $\Delta = E(\mathrm{CH_2 + Ne}) - E(\mathrm{CH_2}) - E(\mathrm{Ne})$ for the CH_2Ne system and $\Delta = E(\mathrm{C_2H_4 + Ne}) - E(\mathrm{C_2H_4}) - E(\mathrm{Ne})$ for the C_2H_4Ne molecule. The distance between linear CH_2 molecule (R_{C-H}=1.07 Å) and Ne atom (in the CH_2Ne system) was set equal to 20 Åand the Ne atom was located on the CH_2 symmetry axis perpendicular to the C-H-C line. The distance between ethylene molecule and Ne atom was set equal to 20 Å. Total energies are reported in Hartree while Δs in milliHartree. Energy for the Ne atom corresponds to the single reference CCSD energy.

We begin our discussion with the analysis of the MRCC results for the H8 and BeH_2 model systems, which was used in extensive studies of various aspects of the single- and multi-reference CC formulations (see Refs.[4,120,121] and[18,122] for geometry specifications). The H8 model consists of eight hydrogen atoms in an octagonal configuration, where a single parameter α defines the displacement of two pairs of hydrogen atoms from the D_{8h} configuration. The dissociation of BeH_2 towards Be+H_2 along the C_{2v} path can be described by a single parameter x, which corresponds to the distance between Be atom and center of the mass of the H_2 molecule. The results of our calculations are collected in Table 6.2 (H8), Table 6.3 (BeH_2) and in Fig.6.1 (H8). For these systems we were able to perform the USS(2)-BW-MRCCSD and USS(2)-Mk-MRCCSD calculations. From Table 6.2 and Fig.6.1 one can notice that the diagonal D-USS(2)-BW and D-USS(2)-Mk approaches has a tendency to overshoot the FCI energies, especially in the

α	FCI	BWSD	BWSD ap.	D-USS(2)-BW	USS(2)-BW	MkSD	D-USS(2)-Mk	USS(2)-Mk
1.0	-4.115319	0.026	-1.342	-1.124	0.232	-0.226	-1.185	0.186
1.2	-4.219796	0.248	-0.547	-0.336	0.423	0.109	-0.379	0.387
1.4	-4.291245	0.417	-0.215	-0.018	0.567	0.309	-0.057	0.533
1.6	-4.334756	0.558	-0.073	0.122	0.704	0.448	0.080	0.665
1.8	-4.354281	0.676	-0.081	0.115	0.843	0.538	0.062	0.794
2.0	-4.353977	0.709	-0.367	-0.199	0.950	0.502	-0.275	0.884
2.2	-4.339477	0.366	-1.492	-1.527	0.824	-0.022	-1.642	0.732
2.3	-4.330284	-0.199	-2.811	-3.162	0.482	-0.776	-3.295	0.365
2.4	-4.325267	-0.733	-4.087	-4.874	0.127	-1.533	-5.010	-0.039
2.5	-4.330475	-0.383	-3.557	-4.102	0.274	-1.106	-4.264	0.115
2.6	-4.342262	0.311	-2.283	-2.405	0.679	-0.249	-2.586	0.522
2.7	-4.355412	0.716	-1.453	-1.298	0.896	0.266	-1.479	0.734
2.8	-4.368123	0.903	-1.001	-0.665	0.971	0.522	-0.841	0.802
2.9	-4.379875	0.974	-0.765	-0.298	0.972	0.638	-0.469	0.798
3.0	-4.390532	0.983	-0.654	-0.084	0.934	0.678	-0.250	0.757
3.2	-4.408627	0.912	-0.623	0.109	0.812	0.650	-0.047	0.631
3.4	-4.422903	0.797	-0.718	0.151	0.672	0.561	0.004	0.492
3.6	-4.434034	0.676	-0.868	0.132	0.540	0.459	-0.006	0.363
3.8	-4.442638	0.561	-1.046	0.086	0.420	0.358	-0.045	0.247
4.0	-4.449239	0.458	-1.243	0.031	0.313	0.265	-0.093	0.145
4.2	-4.454266	0.370	-1.454	-0.023	0.222	0.185	-0.141	0.058
4.5	-4.459592	0.264	-1.803	-0.096	0.109	0.086	-0.208	-0.050
5.0	-4.464600	0.304	-4.244	-0.024	0.006	0.171	-0.093	-0.111

Table 6.2: The errors of the MRCCSD energies (in milliHartree) with respect to the FCI results for the H8 as described by the 6-31G basis set. All electrons were correlated and (2,2) model space based on the RHF orbitals was used. The FCI energies are given in Hartree. Geometry of the H8 systems is defined by α parameter (see Ref.[4] for details). The abbreviations BWSD, BWSD ap., and MkSD correspond to the BW-MRCCSD, BW-MRCCSD ap., and Mk-MRCCSD approaches, respectively.

vicinity of the $\alpha = 2.4$ geometry. However, using the full form of the of the USS(2) correlation Hamiltonian (USS(2)-BW and USS(2)-Mk approaches) significantly improves the quality of the diagonal USS approximations. For example, the 5.010 milliHartree error of the D-USS(2)-Mk method is reduced to 0.039 milliliHartree by the USS(2)-Mk approach. Similar behavior can be observed for the D-USS(2)-BW and USS(2)-BW methods.

Overall, the USS(2)-BW and USS(2)-Mk approaches provide the most balanced description of the ground-state potential energy surface. Moreover, the maximum difference between USS(2)-BW and USS(2)-Mk energies never exceeds 0.2 millilHartree (similar differences can be observer for the diagonal USS(2) formulations), which should be compared with 2.6 and 3.3 milliHartree differences between the Mk-MRCCSD and BW-MRCCSD and BW-MRCCSD ap. energies, respectively. These results give us a strong argument in favor of the universal-

x	FCI	BWSD	BWSD ap.	USS(2)-BW	MkSD	USS(2)-Mk
0.00	-15.799590	0.506	6.467	0.530	0.412	0.468
0.20	-15.795886	0.491	4.504	0.507	0.402	0.448
0.40	-15.788882	0.493	3.955	0.487	0.406	0.429
0.60	-15.779504	0.509	3.579	0.464	0.424	0.407
0.80	-15.768882	0.535	3.412	0.434	0.453	0.376
1.00	-15.757969	0.571	3.631	0.389	0.492	0.329
1.20	-15.747188	0.612	3.886	0.323	0.537	0.258
1.40	-15.736328	0.653	-705.396	0.231	0.580	0.162
1.60	-15.724675	0.683	-399.191	0.130	0.612	0.058
1.80	-15.711278	0.694	-0.928	0.067	0.623	-0.003
2.00	-15.695277	0.682	-1.606	0.092	0.609	0.032
2.20	-15.676239	0.646	-0.377	0.156	0.564	0.106
2.40	-15.654507	0.558	-0.335	0.096	0.451	0.051
2.50	-15.643186	0.454	-0.551	-0.051	0.320	-0.098
2.60	-15.632515	0.244	-1.015	-0.357	0.066	-0.409
2.70	-15.624771	-0.143	-1.774	-0.841	-0.376	-0.905
2.80	-15.624615	-0.543	-2.305	-1.069	-0.791	-1.152
2.90	-15.633163	-1.228	-3.759	-2.015	-1.575	-2.097
3.00	-15.645845	-0.687	-3.745	-1.979	-1.140	-2.054
3.10	-15.659778	0.004	-2.284	-1.015	-0.466	-1.092
3.20	-15.673913	0.408	-1.032	-0.342	-0.062	-0.419
3.30	-15.687819	0.603	-0.195	0.044	0.141	-0.029
3.40	-15.701252	0.675	0.303	0.251	0.229	0.185
3.60	-15.725880	0.653	0.705	0.409	0.254	0.360
3.80	-15.745765	0.558	0.767	0.439	0.216	0.407
4.00	-15.757450	0.458	0.736	0.435	0.173	0.416

Table 6.3: The errors of the MRCCSD energies (in milliHartree) for the BeH_2 molecule with respect to the FCI energies. In our calculations all electrons were correlated and the DZ basis set was used. The (2,2) CMS was employed based on the CASSCF(2,2) orbitals. The abbreviations BWSD, BWSD ap., and MkSD correspond to the BW-MRCCSD, BW-MRCCSD ap., and Mk-MRCCSD approaches, respectively.

ity of the USS methods. For the BeH_2 system similar features of the USS(2) formalism can be observed. Additionally, the USS(2)-BW approach is capable of eliminating the singular behavior of the BW-MRCCSD ap. approach in the vicinity of $x = 1.40$ geometry. The breakdown of the BW-MRCCSD ap. method should be attributed to the presence of the intruder state and perturbative nature of the a posteriori correction.

Next, we compare the BW-MRCCSD-ap and D-USS(2) energies against the FCI ones for the water molecule at various geometries corresponding to symmetric stretch of the O-H bonds. In Table 6.4 we consider three geometries: $1R_e$, 1.5 R_e, and $2R_e$. The values of R_e (R_e=1.84345 bohr) and the equilibrium value of the H-O-H angle (α_{H-O-H}=110.6°) were taken from Ref.[123]. The results collected in Table 6.4 point to the interesting features of the D-USS(2) approach. First, for

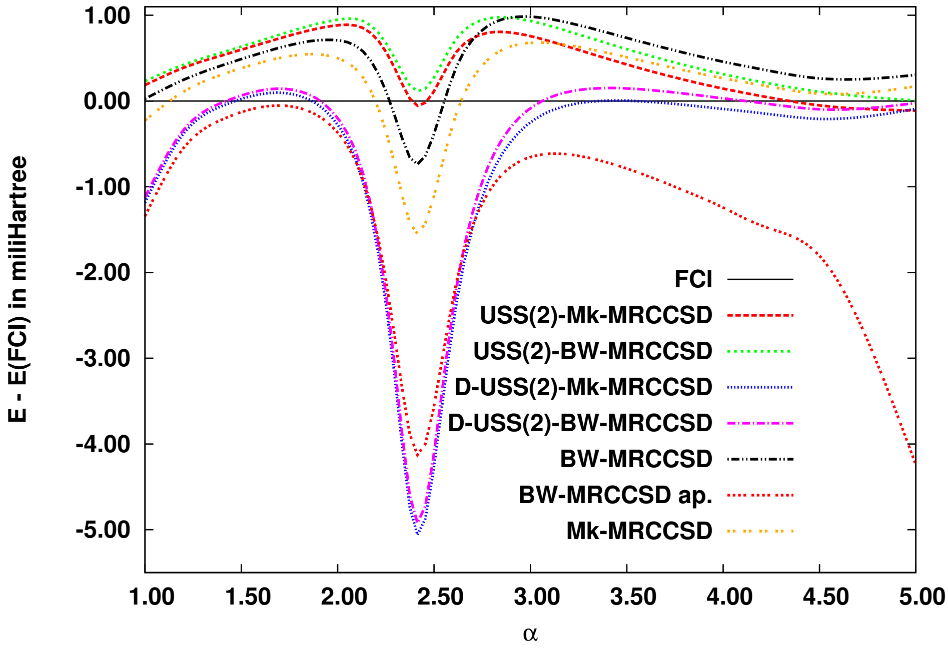

Figure 6.1: Errors of the MRCC energies (with respect to the FCI energies) for various geometries of H8 model. All calculations were performed for the 6-31G basis set.

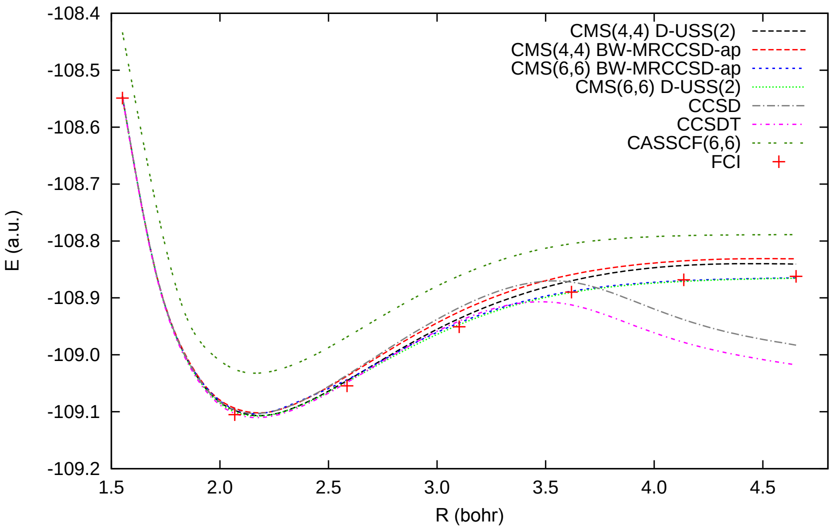

Figure 6.2: Comparison of the CCSD, CCSDT, BW-MRCCSD-ap, D-USS(2), and FCI ground-state potential energy surfaces for the N_2 molecule in DZ basis set.[117] MRCC calculations were performed for two complete model spaces (4,4) and (6,6).

all geometries considered the D-USS(2) results based on the largest (6,6) CMS are consistently better than the corresponding BW-MRCCSD-ap energies. The absolute D-USS(2) errors with respect to the FCI energies are equal to 2.816, 2.234, and 0.101 milliHartree for $1R_e$, $1.5R_e$, and $2R_e$, respectively. This should be compared with the corresponding errors of the BW-MRCCSD-ap approach:

6.679, 2.777, and 0.164 milliHartree. Second, the discrepancies between the results obtained for small (2,2) and large (6,6) model spaces are also smaller for the D-USS(2) approach. For example, the absolute values of differences between D-USS(2) results obtained with the (2,2) and (6,6) model spaces are equal 0.863, 4.786, 3.286 milliHartree, for 1R$_e$, 1.5R$_e$, and 2R$_e$, respectively. Analogous results for the BW-MRCCSD-ap approach amounts to 3.059, 10.812, and 6.957 milliHartree. These results indicate that the D-USS(2) method is less sensitive to the choice of the model space than the BW-MRCCSD-ap approach in spite of the fact that in both cases the same values of converged amplitudes were used in constructing the corrections.

It is well documented that breaking triple bond in the N_2 molecule poses a significant challenge for the single reference methods[124] (see References[125–127] for the detailed discussion of the role of particular higher-order cluster amplitudes in describing N_2 ground-state potential energy surface). From Fig.6.2, which shows ground-state PESs obtained with single- and multi-reference CC methods, one can see that the single reference CCSD and CCSDT methods break down at large internuclear geometries. For the multi-reference calculations the (6,6) model space contains all required static correlation effects in order to provide an accurate description of the ground-state PES. However, to explore how the correlated method can cope with poorer choice of the model spaces we also considered in our studies the smaller (4,4) model space. The corresponding energies are shown in Table 6.5 for representative set of internuclear distances from 1.551 to 4.653 bohr. In Fig.6.2 we have also included the CASSCF(6,6) curve, which has correct shape (the CASSCF(6,6) curve is almost parallel to the FCI one) showing at the same time that a substantial part of the dynamical correlation effects is missing. For both choices of model spaces ((4,4) and (6,6)) the MRCC methods provide substantial improvements over the single reference CC methods. In particular, D-USS(2) and BW-MRCCSD-ap methods entirely eliminate unphysical energy bumps characteristic for the CCSD and CCSDT methods. For (4,4) CMS the D-USS(2) results are more accurate than the ones obtained with the BW-MRCCSD-ap approach. This is especially true for smaller N-N distances (up to $R \simeq 3.0$ bohr). For the larger (6,6) model space, up to $R \simeq 3.6$ bohr the D-USS(2) is slightly better than the BW-MRCCSD-ap, while for larger stretches the BW-MRCCSD-ap approach provides energies, which are marginally better than the D-USS(2) ones (for $R = 4.653$ bohr the errors of the D-USS(2) and BW-MRCCSD-ap energies with respect to the FCI one are 3.310 and 2.641 milliHartree, respectively). With maximum errors not exceeding 7 (D-USS(2)) and 11

(BW-MRCCSD-ap) milliHartree the overall performance of the discussed methods should be considered as satisfactory. In contrast to the results for the water molecule (see Table 6.4), the differences between D-USS(2) results for the (4,4) and (6,6) model spaces can be as large as 28 milliHartree (35 milliHartree in the BW-MRCCSD-ap case).

Method	R_e	$1.5R_e$	$2.0R_e$
CMS(2,2) BW-MRCCSD ap.	3.620	8.035	6.793
CMS(2,2) D-USS(2)-BW	3.679	7.020	3.185
CMS(4,4) BW-MRCCSD ap.	3.096	7.115	1.788
CMS(4,4) D-USS(2)-BW	3.585	7.023	-0.104
CMS(6,6) BW-MRCCSD ap.	6.679	-2.777	-0.164
CMS(6,6) D-USS(2)-BW	2.816	2.234	-0.101
FCI	-76.241860	-76.072348	-75.951665

Table 6.4: The differences between MRCC and FCI energies (in milliHartree) for selected geometries of H_2O molecule in cc-pVDZ basis set.[67] The FCI energies (in Hartree) were taken from Ref.[123].

$R_{\mathrm{N-N}}$	CMS(4,4) D-USS(2)-BW	CMS(4,4) BWSD ap.	CMS(6,6) D-USS(2)-BW	CMS(6,6) BWSD ap.	CCSD	CCSDT	FCI
1.551	2.449	2.269	2.292	3.252	3.132	0.580	-108.549027
2.068	7.355	11.268	4.782	8.769	8.289	2.107	-109.105115
2.585	9.262	17.251	6.844	10.871	19.061	6.064	-109.054626
3.102	14.434	26.133	3.690	6.959	33.545	10.158	-108.950728
3.619	19.975	30.744	-0.462	1.620	17.715	-22.467	-108.889906
4.136	25.166	33.270	-2.820	-1.459	-69.937	-109.767	-108.868239
4.653	21.513	30.844	-3.310	-2.641	-120.836	-155.656	-108.862125

Table 6.5: Comparison of the MRCC energy errors (in milliHartree) with respect to the FCI energies (in Hartree) for selected geometries of N_2 molecule with a DZ basis set.[117] The R_{N-N} distances are given in bohr. The FCI results were taken from Ref.[124]. The abbreviation BWSD ap. corresponds to the BW-MRCCSD ap. approach.

In the cases when the FCI results are available, the quality of approximate ground-state PESs can be estimated based on the calculated non-parallelity errors (NPEs). In Ref.[24], NPEs were calculated for several multi-reference methods for N_2 molecule in cc-pVDZ basis set, which provides an excellent test-bed for assessing D-USS(2) accuracies. In Table 6.6 we collected the NPEs for the D-USS(2), BW-MRCCSD-ap, and uncorrected BW-MRCCSD methods . The uncorrected BW-MRCCSD NPE (23.966 miliHartree) is slightly different from the srMRBWCCSD result of Ref.[24]. The reason lies in the representation of the effective Hamiltonian (see the Computational Details section). The D-USS(2) NPE (5.249 milliHartree) is by 2 milliHartree smaller than the BW-MRCCSD-ap one (7.265

R_{N-N}	D-USS(2)-BW	BW-MRCCSD ap.	BW-MRCCSD	FCI
2.118	7.318	10.406	14.284	-109.278339
2.4	8.263	9.487	19.769	-109.238397
2.7	7.836	9.916	25.825	-109.160305
3.0	6.658	8.725	31.375	-109.086209
3.6	3.930	5.149	38.250	-108.994906
4.2	3.014	3.141	38.195	-108.966950
NPE	5.249	7.265	23.966	

Table 6.6: The MRCC errors (in milliHartree) with respect to FCI energies for the N_2 molecule in spherical cc-pVDZ basis set obtained for several internuclear distances R_{N-N}. In all MRCC calculations CMS(6,6) was used and core orbitals were kept frozen. The NPE errors reported in Ref.[24] using the same basis set and the same set of geometries are equal to 12.033, 0.526, 1.115, 1.667, 19.383, 0.731, 17.628, and 4.724 milliHartree for the SS-MRCCSD(deloc), MRCI, MRAQCC, MRACPF, srM-RBWCCSD, SRMRCC, ⟨SS-MRCEPA(0)⟩(deloc), and ⟨SS-MRCEPA(0)⟩(loc) methods (using nomenclature of Ref.[24]), respectively. The R_{N-N} distances are given in bohr and the FCI energies are given in Hartree.

milliHartree). It is also more than two times smaller than the NPE corresponding to the SS-MRCCSD (or equivalently Mk-MRCCSD) approach. The quality of the D-USS(2) curve is similar to those of the ⟨ SS-MRCEPA(0) ⟩ method.

$R_{\mathrm{Be-Be}}$	CMS(4,4) D-USS(2)-BW	CMS(4,4) BW-MRCCSD ap.	CMS(6,6) D-USS(2)-BW	CMS(6,6) BW-MRCCSD ap.	FCI
3.800	0.383	3.968	0.129	-1.093	-43.857591
3.900	0.338	3.684	0.069	-1.331	-43.862513
4.000	0.309	3.391	0.019	-1.606	-43.865829
4.250	0.341	2.626	-0.034	-2.426	-43.868952
4.500	0.542	1.836	0.042	-3.452	-43.867566
4.750	0.859	1.042	0.216	-4.665	-43.864406
5.000	1.164	0.178	0.374	-6.118	-43.861211
5.200	1.302	-0.618	0.415	-7.510	-43.859187
6.000	0.948	-5.138	-0.197	-15.349	-43.855880
7.000	0.104	-19.587	-1.160	-33.009	-43.855745
8.000	-0.387	-644.828	-1.640	-150.310	-43.855776
9.000	-0.622	-152.013	-1.826	-1554.636	-43.855572
10.000	-0.741	-4734.243	-1.899	-5633.172	-43.855336

Table 6.7: The MRCC energy errors (in milliHartree) with respect to the FCI results for selected geometries of Be_3 molecule with a ANO basis set[119] (see text for details). The FCI energies (in Hartree) were taken from Ref.[119]. The R_{Be-Be} distances are given in bohr.

According to the FCI calculations of Ref.[118] , an interesting situation occurs when the linear dissociation process of the linear Be_3 molecule is considered. The

FCI ground-state potential-energy curve, as a function of the distance R between the central and each of the lateral Be atoms, reveals the deeper primary minimum at $R = 4.26968$ bohr followed by the maximum (barrier) at $R = 6.70089$ bohr, and the shallow van der Waals minimum at $R = 7.88700$ bohr located 13.171 milliHartree above the primary minimum. All the PESs obtained with the D-USS(2) reproduce the basic features of the FCI curve with errors not exceeding 2 milliHartree for all model spaces discussed here. The location of the barrier and the depth of the van der Waals minimum are strongly dependent on the model space employed. The location of the van der Waals minimum is predicted by the D-USS(2) approaches at longer internuclear separations compared to the FCI results (see Fig.6.3, Table 6.7, and Supplementary Material[128]).

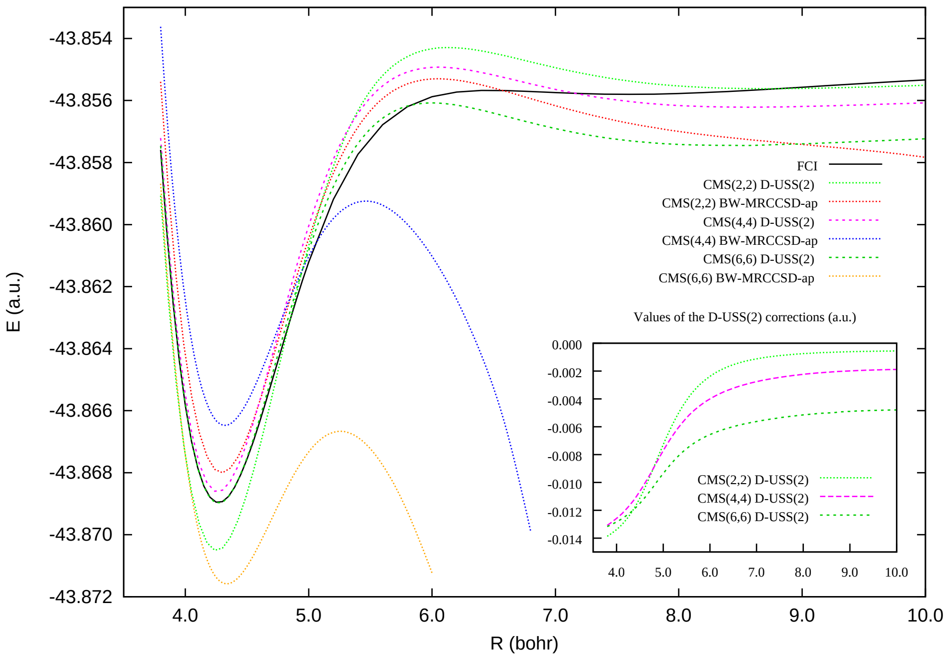

Figure 6.3: The FCI, BW-MRCCSD-ap, and D-USS(2) ground-state potential-energy curves for the linear symmetric dissociation of the linear Be_3 molecule (as described by the ANO basis set of Ref.[119]) as functions of the distance between the central and each of the lateral Be atoms in the linear Be_3 system. The complete model space (2,2), (4,4), and (6,6) were employed in the MRCC calculations. Inset: values of the D-USS(2) correction as a function of the distance.

While the FCI minimum is located at 7.60 bohr the D-USS(2) method locates it at 8.40-8.60 bohr. We believe that this may be associated with the diagonal character of the USS(2) correction and with the inclusion of up to two-body terms in the effective Hamiltonian. Of all BW-MRCCSD-ap approaches only the formulation based on a small (2,2) model space produces reasonable results, although the van der Waals minimum is not reproduced by this approach. The

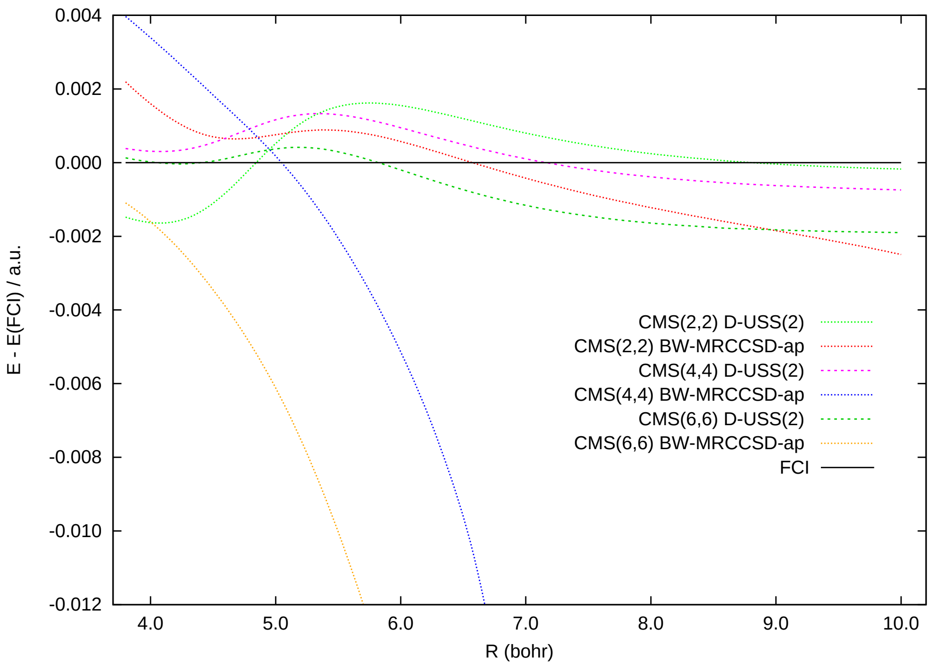

Figure 6.4: Errors of the MRCCSD methods with respect to the FCI energies for the Be_3 molecule as a function of the distance between the central and each of the lateral Be atoms in the linear Be_3 system. All energies were obtained with the ANO basis set of Ref.[119].

larger model spaces (4,4) and (6,6) yield results of very poor quality. For the BW-MRCCSD-ap (6,6) energies the associated errors can reach even few Hartree (see Fig.6.4 and Table 6.7 and Supplementary Material). In calculating BW-MRCCSD a posteriori we eliminated all amplitudes larger than 20.0, otherwise the results would be even worse. The breakdown of the BW-MRCCSD-ap approach for the Be_3 system is clearly a consequence of the perturbative nature of the amplitudes corrections (see discussion in Section 4). It should be stressed that the D-USS(2)(6,6) errors with respect to the FCI results are less than 1 milliHartree for Rs in the $[3.80, 6.60]$ bohr interval.

7. Conclusions

Presently, parallel computing is the most promising way how to apply highly-accurate quantum chemical methods to larger systems. The development of efficient parallel implementations of MRCC methods remains one of the most important tasks to enable accurate calculations of strongly correlated systems. In this work we implemented paralell codes for the BW-MRCCSD and Mk-MRCCSD approaches and for the respective USS(2) correction. Using the two-level parallelism we demonstrated that when reference-level parallelism and task-level parallelism are combined, one can perform large-scale calculations with efficiency exceeding 90%. The preliminary tests also show a very good scalability of the Mk-MRCCSD method, which due to the presence of coupling terms introduces a significant amount of interprocessor communication compared to the BW-MRCCSD approach. Despite this fact, tests employing our processor-groups-based implementation of the Mk-MRCCSD approach have shown no deterioration of the parallel performance.

Using our code, we have also performed large-scale calculations on three selected systems - the dodecane, napthyne izomers, and polycarbenes. It was shown that both BW-BWCCSD and Mk-MRCCSD give reasonable description of the bond breaking in the dodecane. In comparison with the single-reference CCSD(T) method there is no unphysical barrier around $2.0r_e$.

The naphthyne isomers provided an interesting view of the MRCC behavior with varying distance of radical centers distance and size of basis set. The CCSD and Mk-MRCCSD methods predict that all four computed izomers have a triplet ground state, while CCSD(T) predicts a singlet ground state. The BW-MRCCSD ap method provides the S-T gap energy up to 1.5 kcal/mol, what is below the trustful accuracy threshold. The Mk-MRCCSD method compares better to results from other studies than the MR BWCCSD a posteriori.

We have confirmed by MRCC calculations that the ground state of polycarbenes $n = 1$ and $n = 2$ is the corresponding high-spin state. For the polycarbene $n = 2$ the lowest singlest state is the open-shell singlet state, what is in agreement with earlier computations. In general, for all three systems we observed that in strongly multireference cases the Mk-MRCCSD total energy is higher then the BW-MRCCSD a posteriori corrected one, but always lies between BW-MRCCSD uncorrected energy and BW-MRCCSD ap energy.

Finally, we derived a simple form of the second order USS correction to the BW-MRCCSD and Mk-MRCCSD approaches. Our studies proved that despite the fact that the BW-MRCCSD method is not size-extensive, the size-consistency errors of the D-USS(2)-BW approach remain small. On the H8 and BeH_2 examples we demonstrated that the discrepanices between USS(2)- BW and USS(2)-Mk energies are smaller than the corresponding BW-MRCCSD and Mk- MR-CCSD energy differences. In our opinion this is a strong argument for the universality of the USS(2) approaches. The algebraic form of the USS(2) corrections was scrutinized for their similarities with existing a posteriori corrections to the BW-MRCCSD approach. It was shown that the standard a posteriori correction can be related to the USS(2) correction by replacing certain types of cluster amplitudes by their reference specific perturbative estimates. This issue makes the a posteriori corrections more prone to the intruder state problem or accidental zeroing of the perturbative denominators. The USS(2) approach uses the converged values of the cluster amplitudes obtained in the BW-MRCCSD calculations rather than the pertubative estimates. Preliminary numerical tests carried out for H_2O, N_2, and Be_3 molecules with the diagonal USS(2) corrections showed that in the cases where the BW-MRCCSD-ap approach works, the D-USS(2) is capable of providing similar or better quality results (H_2O, N_2). In more challenging cases (intruder state problems, accidental degeneracies) the D-USS(2) approach can avoid the problems of the BW-MRCCSD-ap approach, as shown for the Be_3 case. We have also demonstrated that the D-USS(2) approach is less sensitive to the choice of the model space than the BWMRCCSD-ap method. Our results for H8 and BeH_2 indicate that a better balance of the correlation effects can be achieved by inclusion of the non-diagonal terms in the USS(2) formalism.

We have developed novel parallel implementations of MRCCSD methods and we shown that the code scales across 20000 cores. We believe that our algorithm paves a way for even larger MRCC calculations employing several hundreds of reference functions and large basis sets, which will make high-accurate calculations of chemically and biologically important systems tractable.

Bibliography

[1] J. Paldus, J. Pittner, and P. Čársky, *Multireference coupled-cluster methods: Recent developments*, pp. 455–490. Berlin: Springer, 2010.

[2] B. Jeziorski and H. J. Monkhorst, "Coupled-cluster method for multideterminantal reference states," *Phys. Rev. A*, vol. 24, pp. 1668–1681, 1981.

[3] B. Jeziorski and J. Paldus, "Spin-adapted multireference coupled-cluster approach: Linear approximation for two closed-shell-type reference configurations," *The Journal of Chemical Physics*, vol. 88, no. 9, pp. 5673–5687, 1988.

[4] K. Jankowski, L. Meissner, and J. Wasilewski, "Davidson-type corrections for quasidegenerate states," *International Journal of Quantum Chemistry*, vol. 28, no. 6, pp. 931–942, 1985.

[5] P. Piecuch and J. Paldus, "Application of Hilbert-space coupled-cluster theory to simple $(H_2)_2$ model systems. II. Nonplanar models," *Phys. Rev. A*, vol. 49, pp. 3479–3514, May 1994.

[6] P. Piecuch and J. Paldus, "Orthogonally spin-adapted state-universal coupled-cluster formalism: Implementation of the complete two-reference theory including cubic and quartic coupling terms," *J. Chem. Phys.*, vol. 101, pp. 5875–5890, 1994.

[7] S. A. Kucharski and R. J. Bartlett, "Hilbert space multireference coupled-cluster methods. I. the single and double excitation model," *J. Chem. Phys.*, vol. 95, pp. 8227–8238, 1991.

[8] A. Balková, S. A. Kucharski, L. Meissner, and R. J. Bartlett *Theor. Chim. Acta*, vol. 80, pp. 335–348, 1991.

[9] A. Balková, S. A. Kucharski, L. Meissner, and R. J. Bartlett, "The multireference coupled-cluster method in Hilbert space: An incomplete model space application to the LiH molecule," *J. Chem. Phys.*, vol. 95, pp. 4311–4316, 1991.

[10] K. Kowalski and P. Piecuch, "New classes of noniterative energy corrections to multi-reference coupled-cluster energies," *Mol. Phys.*, vol. 102, pp. 2425–2449, 2004.

[11] J. Pittner and P. Piecuch, "Method of moments for the continuous transition between the Brillouin-Wigner-type and Rayleigh-Schrödinger-type multireference coupled cluster theories," *Mol. Phys.*, vol. 107, pp. 1362–3028, 2009.

[12] I. Hubač in *New Methods in Quantum Theory* (A. Tsipis, V. S. Popov, D. R. Herschbach, and J. S. Avery, eds.), vol. 8 of *NATO ASI Series 3: High Technology*, pp. 183–202, Dordrecht: Kluwer, 1996.

[13] J. Mášik and I. Hubač, "Multireference Brillouin-Wigner coupled-cluster theory. Single-root approach.," *Adv. Quant. Chem.*, vol. 31, pp. 75–104, 1998.

[14] J. Pittner, P. Nachtigall, P. Čársky, J. Mášik, and I. Hubač, "Assessment of the single-root multireference Brillouin-Wigner coupled cluster method. test calculations on CH_2, SiH_2, and twisted ethylene," *J. Chem. Phys.*, vol. 110, pp. 10275–10282, 1999.

[15] J. Pittner, "Continuous transition between Brillouin–Wigner and Rayleigh–Schrödinger perturbation theory, generalized Bloch equation, and Hilbert space multireference coupled cluster," *The Journal of Chemical Physics*, vol. 118, no. 24, pp. 10876–10889, 2003.

[16] U. S. Mahapatra, B. Datta, and D. Mukherjee, "A size-consistent state-specific multireference coupled cluster theory: Formal developments and molecular applications," *J. Chem. Phys.*, vol. 110, no. 13, pp. 6171–6188, 1999.

[17] U. S. Mahapatra, B. Datta, and D. Mukherjee, "Development of a size-consistent state-specific multireference perturbation theory with relaxed model-space coefficients," *Chem. Phys. Lett.*, vol. 299, pp. 42–50, 1999.

[18] F. A. Evangelista, W. D. Allen, H. F. Schaefer, and III, "High-order excitations in state-universal and state-specific multireference coupled cluster theories: Model systems," *The Journal of Chemical Physics*, vol. 125, no. 15, p. 154113, 2006.

[19] F. A. Evangelista, W. D. Allen, and H. F. Schaefer III, "Coupling term derivation and general implementation of state-specific multireference coupled cluster theories," *J. Chem. Phys.*, vol. 127, p. 024102, 2007.

[20] K. Bhaskaran-Nair, O. Demel, and J. Pittner, "Multireference state-specific Mukherjee's coupled cluster method with non-iterative triexcitations," *J. Chem. Phys.*, vol. 129, p. 184105, 2008.

[21] K. Bhaskaran-Nair, O. Demel, and J. Pittner, "Multireference mukherjee's coupled cluster method with triexcitations in the linked formulation: efficient implementation and applications," *J. Chem. Phys.*, vol. 132, p. 154105, 2010.

[22] O. Demel, K. Bhaskaran-Nair, and J. Pittner, "Uncoupled multireference state-specific Mukherjee's coupled cluster method with triexcitations," *J. Chem. Phys.*, vol. 133, p. 134106, 2010.

[23] K. Bhaskaran-Nair, O. Demel, J. Šmydke, and J. Pittner, "Multireference state-specific Mukherjee's coupled cluster method with noniterative triexcitations using uncoupled approximation," *The Journal of Chemical Physics*, vol. 134, no. 15, p. 154106, 2011.

[24] S. Das, D. Mukherjee, and M. Kallay, "Full implementation and benchmark studies of Mukherjee's state-specific multireference coupled-cluster ansatz," *The Journal of Chemical Physics*, vol. 132, no. 7, p. 074103, 2010.

[25] U. S. Mahapatra and S. Chattopadhyay, "Potential energy surface studies via a single root multireference coupled cluster theory," *J. Chem. Phys.*, vol. 133, no. 7, p. 074102, 2010.

[26] C. Bloch, "Sur la theorie des perturbations des etats lies," *Nucl. Phys.*, vol. 6, no. 3, pp. 329–347, 1958.

[27] F. Jorgensen, "Effective Hamiltonians," *Mol. Phys.*, vol. 29, no. 4, pp. 1137–1164, 1975.

[28] C. E. Soliverez, "General-theory of effective-Hamiltonians," *Phys. Rev. A*, vol. 24, no. 1, pp. 4–9, 1981.

[29] P. Durand, "Direct determination of effective-Hamiltonians by wave-operator methods .1. general formalism," *Phys. Rev. A*, vol. 28, no. 6, pp. 3184–3192, 1983.

[30] T. H. Schucan and H. A. Weidenmüller, "Effective interaction in nuclei and its perturbation expansion - algebraic approach," *Ann. Phys.*, vol. 73, no. 1, pp. 108–135, 1972.

[31] T. H. Schucan and H. A. Weidenmüller, "Perturbation-theory for effective interaction in nuclei," *Ann. Phys.*, vol. 76, no. 2, pp. 483–509, 1973.

[32] K. Kowalski and P. Piecuch, "Complete set of solutions of multireference coupled-cluster equations: The state-universal formalism," *Phys. Rev. A*, vol. 61, p. 052506, 2000.

[33] X. Li and J. Paldus, "General-model-space state-universal coupled-cluster theory: Connectivity conditions and explicit equations," *J. Chem. Phys.*, vol. 119, pp. 5320–5333, 2003.

[34] X. Li and J. Paldus, "n-reference, m-state coupled-cluster method: Merging the state-universal and reduced multireference coupled-cluster theories," *J. Chem. Phys.*, vol. 119, pp. 5334–5345, 2003.

[35] X. Li and J. Paldus, "The general-model-space state-universal coupled-cluster method exemplified by the LiH molecule," *J. Chem. Phys.*, vol. 119, pp. 5346–5357, 2003.

[36] P. Piecuch and J. I. Landman, "Parallelization of multi-reference coupled-cluster method," *Parallel Computing*, vol. 26, no. 7-8, pp. 913–943, 2000.

[37] E. Prochnow, M. E. Harding, and J. Gauss, "Parallel calculation of CCSDT and Mk-MRCCSDT energies," *J. Chem. Theory Comput.*, vol. 6, no. 8, pp. 2339–2347, 2010.

[38] I. Lindgren and J. Morrison, *Atomic Many-Body Theory.* Berlin: Springer, 1982.

[39] J. Pittner, "Continuous transition between brillouin-wigner and rayleigh-schrödinger perturbation theory, generalized bloch equation, and hilbert space multireference coupled cluster," *J. Chem. Phys.*, vol. 118, pp. 10876–10889, 2003.

[40] I. Hubač, J. Mášik, P. Mach, J. Urban, and P. Babinec, "Multireference Brillouin-Wigner coupled-cluster theory," in *Computational Chemistry. Reviews of Current Trends.* (J. Leszczynski, ed.), vol. 3, pp. 1–48, Singapore: World Scientific, 1999.

[41] J. Mášik and I. Hubač, "Multireference Brillouin-Wigner coupled-cluster theory: Hilbert space approach," in *Quantum Systems in Chemistry and Physics: Trends in Methods and Applications* (R. McWeeny, J. Maruani,

Y. G. Smeyers, and S. Wilson, eds.), pp. 283–308, Dordrecht: Kluwer Academic Publishers, 1997.

[42] P. Čársky, V. Hrouda, V. Sychrovský, I. Hubač, P. Babinec, P. Mach, J. Urban, and J. Mášik *Collect. Czech. Chem. Commun.*, vol. 60, p. 1419, 1995.

[43] I. Hubač and S. Wilson, *Brillouin-Wigner methods for many-body systems*, vol. 21. Berlin: Springer, 2010.

[44] I. Hubač, J. Pittner, and P. Čársky, "Size-extensivity correction for the state-specific multireference Brillouin-Wigner coupled-cluster theory," *J. Chem. Phys.*, vol. 112, pp. 8779–8784, 2000.

[45] J. Pittner and O. Demel, "Towards the multireference Brillouin-Wigner coupled clusters method with iterative connected triples: MR BWCCSDT-α approximation," *J. Chem. Phys.*, vol. 122, p. 181101, 2005.

[46] O. Demel and J. Pittner, "Multireference Brillouin-Wigner coupled clusters method with noniterative perturbative connected triples," *J. Chem. Phys.*, vol. 124, p. 144112, 2006.

[47] J. Pittner and J. Šmydke, "Analytic gradient for the multireference Brillouin–Wigner coupled–cluster method and for the state-universal multireference coupled–cluster method," *J. Chem. Phys.*, vol. 127, p. 114103, 2007.

[48] O. Demel and J. Pittner, "Multireference Brillouin-Wigner coupled clusters method with singles, doubles, and triples: Efficient implementation and comparison with approximate approaches," *J. Chem. Phys.*, vol. 128, p. 104108, 2008.

[49] J. Pittner, O. Demel, P. Čársky, and I. Hubač, "Four-reference state-specific Brillouin-Wigner coupled-cluster method: Study of the IBr molecule.," *Int. J. Mol. Sci.*, vol. 2, pp. 281–290, 2002.

[50] J. Pittner, P. Čársky, and I. Hubač, "Four- and 8-reference state-specific Brillouin-Wigner coupled-cluster method: Study of the singlet oxygen," *Int. J. Quant. Chem.*, vol. 90, pp. 1031–1037, 2002.

[51] J. Pittner, J. Šmydke, P. Čársky, and I. Hubač, "State-specific Brillouin-Wigner multireference coupled cluster study of the F_2 molecule: Assessment of the a posteriori size-extensivity correction," *J. Molec. Struc. (Theochem)*, vol. 547, pp. 239–244, 2001.

[52] I. S. K. Kerkines, J. Pittner, P. Čársky, A. Mavridis, and I. Hubač, "On the ground states of CaC and ZnC: a multireference Brillouin-Wigner coupled cluster study.," *J. Chem. Phys.*, vol. 117, pp. 9733–9739, 2002.

[53] V. I. Teberekidis, I. S. K. Kerkines, C. A. Tsipis, P. Čársky, and A. Mavridis, "Ground states of BeC and MgC: Comparative multireference Brillouin-Wigner coupled cluster and configuration interaction study," *Int. J. Quant. Chem.*, vol. 102, pp. 762–774, 2005.

[54] S. Kardahakis, J. Pittner, P. Čársky, and A. Mavridis, "Multireference configuration interaction and coupled cluster calculations on the $x^3\sigma^-$, $a^1\delta$, and $b^1\sigma^+$ states of the NF molecule," *Int. J. Quant. Chem.*, vol. 104, pp. 458–467, 2005.

[55] M. Tobita, S. A. Perera, M. Musial, R. J. Bartlett, M. Nooijen, and J. S. Lee, "Critical comparison of single-reference and multireference coupled-cluster methods: Geometry, harmonic frequencies, and excitation energies of N_2O_2," *J. Chem. Phys.*, vol. 119, pp. 10713–10723, 2003.

[56] J. C. Sancho-García, J. Pittner, P. Čársky, and I. Hubač, "Multireference coupled-cluster calculations on the energy of activation in the automerization of cyclobutadiene: assessment of the state-specific multireference Brillouin-Wigner theory," *J. Chem. Phys.*, vol. 112, pp. 8785–8788, 2000.

[57] J. Pittner, P. Nachtigall, P. Čársky, and I. Hubač, "A multireference coupled cluster study of the singlet-triplet gap in tetramethyleneethane molecule," *J. Phys. Chem. A*, vol. 105, pp. 1354–1356, 2001.

[58] O. Rey-Puiggros, J. Pittner, P. Čársky, P. Stampfuß, and W. Wenzel, "Multireference Brillouin-Wigner coupled cluster singles and doubles (MR BWCCSD) and multireference doubles configuration interaction (MRDCI) calculations for the Bergman cyclization reaction," *Collect. Czech. Chem. Commun.*, vol. 68, pp. 2309–2321, 2003.

[59] O. Demel, J. Pittner, P. Čársky, and I. Hubač, "MR BWCCSD study of the singlet–triplet separation in alkylcarbenes," *J. Phys. Chem. A*, vol. 108, pp. 3125–3128, 2004.

[60] I. S. K. Kerkines, P. Čársky, and A. Mavridis *J. Phys. Chem. A*, vol. 109, p. 10148, 2005.

[61] J. Brabec and J. Pittner, "The singlet-triplet gap in trimethylenmethane and the ring-opening of methylenecyclopropane: a multireference Brillouin-Wigner coupled cluster study," *J. Phys. Chem. A*, vol. 110, pp. 11765–11769, 2006.

[62] L. Veis, P. Čársky, J. Pittner, and J. Michl, "Coupled cluster study of polycyclopentanes: Structure and properties of C_5H_{2n}, $n = 0 - 4$," *Collect. Czech. Chem. Commun.*, vol. 73, pp. 1525–1551, 2008.

[63] J. Paldus, X. Li, and N. D. Petraco, "General-model-space state-universal coupled-cluster method: Diagrammatic approach," *J. Math. Chem.*, vol. 35, pp. 215–251, 2004.

[64] S. Hirata, "Tensor contraction engine: Abstraction and automated parallel implementation of configuration-interaction, coupled-cluster, and many-body perturbation theories," *J. Phys. Chem. A*, vol. 107, no. 46, pp. 9887–9897, 2003.

[65] S. Hirata, "Tensor Contraction Engine manual," 2004.

[66] X. Z. Li and J. Paldus, "Energetics of naphthynes - performance of reduced multi-reference coupled-cluster methods for diradicals(1)," *Canadian Journal of Chemistry-Revue Canadienne De Chimie*, vol. 87, no. 7, pp. 917–926, 2009.

[67] J. Thom H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen," *The Journal of Chemical Physics*, vol. 90, no. 2, pp. 1007–1023, 1989.

[68] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, "Self-consistent molecular orbital methods. XX. a basis set for correlated wave functions," *J. Chem. Phys.*, vol. 72, pp. 650–654, 1980.

[69] P.-O. Widmark, P.-Å. Malmqvist, and B. Roos, "Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. I. first row atoms," *Theor. Chim. Acta*, vol. 77, pp. 291–306, 1990.

[70] D. G. Fedorov, R. M. Olson, K. Kitaura, M. S. Gordon, and S. Koseki, "A new hierarchical parallelization scheme: Generalized distributed data interface (GDDI), and an application to the fragment molecular orbital method (FMO)," *J. Chem. Phys.*, vol. 25, no. 6, pp. 872–880, 2004.

[71] J. Hutter and A. Curioni, "Car-Parrinello molecular dynamics on massively parallel computers," *ChemPhysChem*, vol. 6, no. 9, pp. 1788–1793, 2005.

[72] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, "Quantum espresso: a modular and open-source software project for quantum simulations of materials," *Journal of Physics: Condensed Matter*, vol. 21, no. 39, p. 395502, 2009.

[73] J. Dongarra, "Preface: Basic linear algebra subprograms technical (BLAST) forum standard II," *International Journal of High Performance Computing Applications*, vol. 16, pp. 115–199, Summer 2002.

[74] J. Dongarra, "Preface: Basic linear algebra subprograms technical (BLAST) forum standard I," *International Journal of High Performance Computing Applications*, vol. 16, pp. 1–111, Spring 2002.

[75] L. S. Blackford, J. Demmel, J. Dongarra, I. Duff, S. Hammarling, G. Henry, M. Heroux, L. Kaufman, A. Lumsdaine, A. Petitet, R. Pozo, K. Remington, and R. C. Whaley, "An updated set of basic linear algebra subprograms (BLAS)," *ACM Trans. Math. Softw.*, vol. 28, pp. 135–151, June 2002.

[76] P. Pulay, "Convergence acceleration of iterative sequences - the case of SCF iteration," *Chem. Phys. Lett.*, vol. 73, no. 2, pp. 393–398, 1980.

[77] P. Pulay, "Improved SCF convergence acceleration," *J. Comput. Chem.*, vol. 3, no. 4, pp. 556–560, 1982.

[78] W. J. Hehre, R. Ditchfield, and J. A. Pople, "Self—Consistent Molecular Orbital Methods. XII. Further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules," *The Journal of Chemical Physics*, vol. 56, no. 5, pp. 2257–2261, 1972.

[79] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, "Ab-initio calculation of vibrational absorption and circular-dichroism spectra using density-functional force-fields," *J. Phys. Chem.*, vol. 98, no. 45, pp. 11623–11627, 1994.

[80] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. de Jong, "NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations," *Comp. Phys. Commun.*, vol. 181, no. 9, pp. 1477–1489, 2010.

[81] V. Tipparaju, E. Aprá, W. Yu, and J. S. Vetter, *Enabling a highly-scalable global address space model for petascale computing.* CF '10, New York, NY, USA: ACM, 2010.

[82] S. Alam, R. Barrett, J. Kuehn, and S. Poole, *Performance analysis and projections for Petascale applications on Cray XT series systems.* 2009.

[83] M. Kleinschmidt, C. M. Marian, M. Waletzke, and S. Grimme, "Parallel multireference configuration interaction calculations on mini-beta-carotenes and beta-carotene," *J. Chem. Phys.*, vol. 130, p. 044708, 2009.

[84] J. P. Ceron-Carrasco, A. Requena, and C. M. Marian, "Theoretical study of the low-lying excited states of beta-carotene isomers by a multireference configuration interaction method," *Chem. Phys.*, vol. 373, no. 1-2, pp. 98–103, 2010.

[85] W. Li, P. Piecuch, J. R. Gour, and S. Li, "Local correlation calculations using standard and renormalized coupled-cluster approaches," *J. Chem. Phys.*, vol. 131, no. 11, p. 114109, 2009.

[86] P. Piecuch and M. Wloch, "Renormalized coupled-cluster methods exploiting left eigenstates of the similarity-transformed Hamiltonian," *The Journal of Chemical Physics*, vol. 123, no. 22, p. 224105, 2005.

[87] P. J. Stephens, F. J. Devlin, C. S. Ashvar, C. F. Chabalowski, and M. J. Frisch, "Theoretical calculation of vibrational circular dichroism spectra," *Faraday Discuss.*, vol. 99, pp. 103–119, 1994.

[88] A. D. Becke, "Density-functional thermochemistry. III. the role of exact exchange," *J. Chem. Phys.*, vol. 98, no. 7, pp. 5648–5652, 1993.

[89] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, pp. 785–789, Jan 1988.

[90] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Can. J. Phys.*, vol. 58, no. 8, pp. 1200–1211, 1980.

[91] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, "Basis set exchange: A community database for computational sciences," *J. Chem. Inf. Model.*, vol. 47, no. 3, pp. 1045–1052, 2007.

[92] D. Feller, "The role of databases in support of computational chemistry calculations," *J. Comp. Chem.*, vol. 17, no. 13, pp. 1571–1586, 1996.

[93] P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, "Single-reference, size-extensive, non-iterative coupled-cluster approaches to bond breaking and biradicals," *Chem. Phys. Lett.*, vol. 418, no. 4-6, pp. 467 – 474, 2006.

[94] Y. Ge, M. S. Gordon, and P. Piecuch, "Breaking bonds with the left eigenstate completely renormalized coupled-cluster method," *J. Chem. Phys.*, vol. 127, p. 174106, 2007.

[95] D. B. Borders and T. W. Doyle, eds., *Enediyne antibiotics as antitumor agents*. 270 Madison Avenue, New York, NY 10016: Marcel Dekker, 1995.

[96] N. Koga and H. Iwamura, "From high spin polycarbenes to carbene-based magnets," in *Carbene chemistry: from fleeting intermediates to powerful reagents* (G. Bertrand, ed.), ch. 9, pp. 271–296, 270 Madison avenue, New York, NY 10016: FontisMedia & Marcel Dekker, 2002.

[97] W. Mizukami, Y. Kurashige, and T. Yanai, "Communication: Novel quantum states of electron spins in polycarbenes from ab initio density matrix renormalization group calculations," *J. Chem. Phys.*, vol. 133, no. 9, p. 091101, 2010.

[98] D. Zgid and M. Nooijen, "The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space," *J. Chem. Phys.*, vol. 128, no. 14, p. 144116, 2008.

[99] K. Kowalski, "A universal state-selective approach to multireference coupled-cluster non-iterative corrections," *The Journal of Chemical Physics*, vol. 134, no. 19, p. 194107, 2011.

[100] J. F. Stanton, "Why CCSD(T) works: a different perspective," *Chemical Physics Letters*, vol. 281, no. 1–3, pp. 130 – 134, 1997.

[101] S. A. Kucharski and R. J. Bartlett, "Noniterative energy corrections through fifth-order to the coupled cluster singles and doubles method," *The Journal of Chemical Physics*, vol. 108, no. 13, pp. 5243–5254, 1998.

[102] S. A. Kucharski and R. J. Bartlett, "Sixth-order energy corrections with converged coupled cluster singles and doubles amplitudes," *The Journal of Chemical Physics*, vol. 108, no. 13, pp. 5255–5264, 1998.

[103] T. D. Crawford and J. F. Stanton, "Investigation of an asymmetric triple-excitation correction for coupled-cluster energies," *International Journal of Quantum Chemistry*, vol. 70, no. 4-5, pp. 601–611, 1998.

[104] S. R. Gwaltney and M. Head-Gordon, "A second-order correction to singles and doubles coupled-cluster methods based on a perturbative expansion of a similarity-transformed Hamiltonian," *Chemical Physics Letters*, vol. 323, no. 1–2, pp. 21 – 28, 2000.

[105] S. R. Gwaltney, C. D. Sherrill, M. Head-Gordon, and A. I. Krylov, "Second-order perturbation corrections to singles and doubles coupled-cluster methods: General theory and application to the valence optimized doubles model," *The Journal of Chemical Physics*, vol. 113, no. 9, pp. 3548–3560, 2000.

[106] S. Hirata, M. Nooijen, I. Grabowski, and R. J. Bartlett, "Perturbative corrections to coupled-cluster and equation-of-motion coupled-cluster energies: A determinantal analysis," *The Journal of Chemical Physics*, vol. 114, no. 9, pp. 3919–3928, 2001.

[107] Y. J. Bomble, J. F. Stanton, M. Kallay, and J. Gauss, "Coupled-cluster methods including noniterative corrections for quadruple excitations," *The Journal of Chemical Physics*, vol. 123, no. 5, p. 054101, 2005.

[108] A. G. Taube and R. J. Bartlett, "Improving upon CCSD(T): Lambda CCSD(T). I. Potential energy surfaces," *The Journal of Chemical Physics*, vol. 128, no. 4, p. 044110, 2008.

[109] G. D. P. III and R. J. Bartlett, "A full coupled-cluster singles and doubles model: The inclusion of disconnected triples," *The Journal of Chemical Physics*, vol. 76, no. 4, pp. 1910–1918, 1982.

[110] I. Hubač and S. Wilson, "On the use of Brillouin-Wigner perturbation theory for many-body systems," *Journal of Physics B: Atomic, Molecular and Optical Physics*, vol. 33, no. 3, p. 365, 2000.

[111] K. Kowalski and P. Piecuch, "The method of moments of coupled-cluster equations and the renormalized CCSD[T], CCSD(T), CCSD(TQ), and CCSDT(Q) approaches," *The Journal of Chemical Physics*, vol. 113, no. 1, pp. 18–35, 2000.

[112] F. A. Evangelista, E. Prochnow, J. Gauss, H. F. Schaefer, and III, "Perturbative triples corrections in state-specific multireference coupled cluster theory," *The Journal of Chemical Physics*, vol. 132, no. 7, p. 074107, 2010.

[113] J. Pittner, "TINYMRCC: Suite of MRCC codes developed at Dr. Pittner's group," 2006.

[114] S. Hirata, "Tensor Contraction Engine: Abstraction and automated parallel implementation of configuration-interaction, coupled-cluster, and many-body perturbation theories," *The Journal of Physical Chemistry A*, vol. 107, no. 46, pp. 9887–9897, 2003.

[115] S. Hirata, "Symbolic algebra in quantum chemistry," *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, vol. 116, pp. 2–17, 2006. 10.1007/s00214-005-0029-5.

[116] J. Brabec, J. Pittner, H. J. J. van Dam, E. Aprà, and K. Kowalski, "Parallel implementation of multireference coupled-cluster theories based on the reference-level parallelism," *Journal of Chemical Theory and Computation*, vol. 8, no. 2, pp. 487–497, 2012.

[117] J. Thom H. Dunning, "Gaussian basis functions for use in molecular calculations. I. contraction of (9s5p) atomic basis sets for the first-row atoms," *The Journal of Chemical Physics*, vol. 53, no. 7, pp. 2823–2833, 1970.

[118] J. M. Junquera-Hernandez, J. Sanchez-Marin, G. L. Bendazzoli, and S. Evangelisti, "Full configuration interaction calculation of Be_3," *The Journal of Chemical Physics*, vol. 120, no. 18, pp. 8405–8411, 2004.

[119] J. M. Junquera-Hernandez, J. Sanchez-Marin, G. L. Bendazzoli, and S. Evangelisti, "Full configuration interaction calculation of singlet excited states of Be_3," *The Journal of Chemical Physics*, vol. 121, no. 15, pp. 7103–7109, 2004.

[120] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, "Application of Hilbert-space coupled-cluster theory to simple $(H_2)_2$ model systems: Planar models," *Phys. Rev. A*, vol. 47, pp. 2738–2782, Apr 1993.

[121] P. Piecuch and J. Paldus, "Property evaluation using the two-reference state-universal coupled-cluster method," *The Journal of Physical Chemistry*, vol. 99, no. 42, pp. 15354–15368, 1995.

[122] G. D. Purvis, R. Shepard, F. B. Brown, and R. J. Bartlett, "C_{2v} insertion pathway for BeH_2: A test problem for the coupled-cluster single and double excitation model," *International Journal of Quantum Chemistry*, vol. 23, no. 3, pp. 835–845, 1983.

[123] J. Olsen, P. Jorgensen, H. Koch, A. Balkova, and R. J. Bartlett, "Full configuration–interaction and state of the art correlation calculations on water in a valence double-zeta basis with polarization functions," *The Journal of Chemical Physics*, vol. 104, no. 20, pp. 8007–8015, 1996.

[124] K. Kowalski and P. Piecuch, "Renormalized CCSD(T) and CCSD(TQ) approaches: Dissociation of the N_2 triple bond," *The Journal of Chemical Physics*, vol. 113, no. 14, pp. 5644–5652, 2000.

[125] X. Li and J. Paldus, "Reduced multireference coupled cluster method: Rovibrational spectra of N_2," *The Journal of Chemical Physics*, vol. 113, no. 22, pp. 9966–9977, 2000.

[126] X. Li and J. Paldus, "Partially linearized, fully size-extensive, and reduced multireference coupled-cluster methods. II. Applications and performance," *The Journal of Chemical Physics*, vol. 128, no. 14, p. 144119, 2008.

[127] X. Li and J. Paldus, "Full potential energy curve for N_2 by the reduced multireference coupled-cluster method," *The Journal of Chemical Physics*, vol. 129, no. 5, p. 054104, 2008.

[128] "See supplementary material at http://dx.doi.org/10.1063/1.3692969 for more Be_3 geometries,"

List of Figures

List of Tables