

DEVELOPMENT AND APPLICATIONS OF COUPLED-CLUSTER METHODS AND
POTENTIAL ENERGY SURFACE EXTRAPOLATION SCHEMES

By

Jesse J. Lutz

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Chemistry

2011

ABSTRACT

DEVELOPMENT AND APPLICATIONS OF COUPLED-CLUSTER METHODS AND POTENTIAL ENERGY SURFACE EXTRAPOLATION SCHEMES

By

Jesse J. Lutz

The generation of highly accurate potential energy surfaces (PESs) for reactive processes represents a difficult challenge for modern electronic structure theory. Since chemical reactions often involve breaking and forming bonds or intermediate and transition state species, one must employ a methodology that provides a balanced and highly accurate description of varying levels of electronic degeneracy, but that is also practical enough to be applied to a wide range of chemical problems. Using small to medium sized systems, we examine the performance of two classes of coupled-cluster (CC) methods which are capable of accounting for the diverse electron correlation effects encountered in the majority of ground- and excited-state PES considerations. The first class of methods are the size-extensive completely renormalized (CR) CC approaches for ground-states and their equation of motion (EOM) CC extensions for excited-states, in which noniterative corrections due to higher-order excitations are added to the energies obtained with the standard CC and EOMCC approximations, such as CCSD (CC with singles and doubles) or EOMCCSD (EOMCC with singles and doubles), respectively. In particular, we focus on the left-eigenstate CR-CC(2,3) and CR-EOMCC(2,3) methods, in which a noniterative correction due to triple excitations is added to the CCSD or EOMCCSD energy, respectively, and, when necessary, a noniterative correction for quadruple excitations is also included via the CR-CC(2,3)+Q approach. A new variant of the CR-EOMCC(2,3) method, abbreviated as δ -CR-EOMCC(2,3), that can pro-

vide a size-intensive treatment of excitation energies, is discussed as well. The second class of methods considered here is the active-space variants of the electron-attached (EA) and ionized (IP) EOMCC theories, which utilize the idea of applying a linear electron-attaching or ionizing operator to the correlated, ground-state CC wave function of an N -electron closed-shell system in order to generate the ground and excited states of the related $(N \pm 1)$ -electron radical species of interest. These approaches use a physically motivated set of active orbitals to *a priori* select the dominant higher-order correlation effects to be included in the calculation, which significantly reduces the costs of the high-level EA- and IP-EOMCC approximations needed for obtaining accurate results for open-shell species without sacrificing accuracy. We have also developed a general extrapolation strategy for reducing the cost of generating PESs with correlated electronic structure methods using the concept of correlation energy scaling. Benchmark studies were performed to demonstrate typical accuracies for two types of PES extrapolation schemes, namely, the single-level PES extrapolation schemes, in which the essential quantity, the correlation energy scaling factor, is generated using only the quantum chemistry method of interest, and the dual-level PES extrapolation schemes, where lower-order approaches are used to estimate the correlation energy scaling factor corresponding to the method of interest. Unifying features of these PES extrapolation techniques are discussed, including the role of pivot geometries and base wave functions, and PES extrapolation to the complete basis set limit is examined as well. Finally, the most essential details of the new open-shell EOMCCSD and EA- and IP-EOMCC computer codes for the GAMESS software package, developed as part of this thesis research, are described.

Copyright by
JESSE J. LUTZ
2011

This dissertation is dedicated to my loving and ever supportive parents who taught me the most important things I could ever learn.

ACKNOWLEDGMENT

I owe a great deal of gratitude to my Ph.D. advisor, Professor Piotr Piecuch for countless opportunities to contribute important research which never would have been possible otherwise. He and his group were nurturing and supportive in many ways, particularly regarding research, but never limited to it. I would also like to thank the rest of my guidance committee, namely my second reader, Professor Katharine L. C. Hunt, as well as Professor Robert I. Cukier and Professor James E. Jackson for lending their time and support, and for all their comments and advice regarding research and life in general.

It is also important to mention my great appreciation for the efforts of Professor Marta Włoch and Dr. Jeffery R. Gour towards my professional development. As a result of their able guidance, I progressed much more quickly than I would have without them, and I am deeply grateful for their help and collaboration in many of the projects described in this dissertation. I would also like to thank other members of the Piecuch group, both past and present, for the help and companionship they provided, including Dr. Maricris Lodriguito, Dr. Wei Li, Dr. Jun Shen, and Mr. Jared Hansen.

Finally, I would like to acknowledge Michigan State University and the Department of Chemistry for several fellowships that supported part of my research. I would also like to acknowledge the MSU High Performance Computing Center for providing resources which were utilized for part of this research. Most of the research discussed in this thesis was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, through grants awarded to my advisor and research fellowships to me, and I am appreciative for this support as well.

TABLE OF CONTENTS

List of Tables	ix
List of Figures	xiv
1 Introduction	1
2 Project Objectives	17
3 Applications of Coupled-Cluster and Equation-of-Motion Coupled-Cluster Methods	19
3.1 Theory	19
3.1.1 Single-Reference Coupled-Cluster Theory for Ground States	20
3.1.2 Equation-of-Motion Coupled-Cluster Theory for Electronically Excited, Electron-Attached, and Ionized States	25
3.1.2.1 Excited States	26
3.1.2.2 Electron-Attached and Ionized States	29
3.1.3 The Method of Moments of Coupled-Cluster Equations	32
3.1.3.1 Completely Renormalized Coupled-Cluster and Equation-of-Motion Coupled-Cluster Approaches	35
3.1.3.2 Biorthogonal Formulation of the MMCC Equations and the CR-CC(2,3) and CR-EOMCC(2,3) Methods	43
3.1.3.3 A Size-Intensive Variant of CR-EOMCC(2,3): The δ -CR-EOMCC(2,3) Method	49
3.1.4 The Active-Space Coupled-Cluster and Equation-of-Motion Coupled-Cluster Approaches	51
3.2 Applications	56
3.2.1 The DBH24 Benchmark Database for Thermochemical Kinetics	57
3.2.2 Addition Reactions of Ethylene and Acetylene to Ozone	67
3.2.3 Mechanism of the Isomerization of Bicyclobutane to Butadiene	75
3.2.4 Excited-State Potential Energy Surfaces for the Dissociation of Water	81
3.2.5 Excitation Energies and Hydrogen-Bonding-Induced Spectral Shifts in Complexes of <i>cis</i> -7-Hydroxyquinoline	91
3.2.6 Geometries and Adiabatic Excitation Energies of CNC, C ₂ N, N ₃ , and NCO	100
4 Potential Energy Surface Extrapolation Schemes	129
4.1 Motivation	129

4.2	Theory	134
4.3	Single-Level Potential Energy Surface Extrapolation Schemes	137
4.3.1	Potential Energy Surface Extrapolation to Larger Basis Sets	137
4.3.2	Potential Energy Surface Extrapolation to the Complete Basis Set Limit	138
4.3.3	The Role of Pivot Geometries and Base Wave Functions in Single-Level PES Extrapolations	140
4.3.4	Application to the Isomerization of Bicyclobutane to Butadiene	141
4.4	Dual-Level Potential Energy Surface Extrapolation Schemes	153
4.4.1	Using Lower-Order Methods to Produce Correlation Energy Scaling Factors in High-Level Calculations	154
4.4.2	Application to Single Bond-Breaking Potential Energy Curves	156
4.4.3	Application to the Isomerization of Bicyclobutane to Butadiene	163
4.4.4	The Role of Pivot Geometries and Base Wave Functions in Dual-Level PES Extrapolations	166
5	Development of Computer Codes for the GAMESS software Package	174
5.1	Key Details of Efficient Computer Implementations	175
5.1.1	Standard Equation-of-Motion Coupled-Cluster Theory with Singles and Doubles for Open-Shell Systems	176
5.1.2	Electron-Attached and Ionized Equation-of-Motion Coupled-Cluster Theories	184
6	Summary and Concluding Remarks	187
	References	191

LIST OF TABLES

3.1	Representative barrier heights database DBH24 taken from Ref. [204]. . . .	59
3.2	Mean signed error (MSE) and mean unsigned error (MUE) of coupled cluster methods calculated with all electrons correlated compared against the DBH24 benchmark database (in kcal/mol).	61
3.3	Mean signed error (MSE) and mean unsigned error (MUE) of coupled cluster methods calculated with frozen core approximation compared against the DBH24 benchmark database (in kcal/mol).	62
3.4	Errors resulting from CR-CC(2,3) calculations of forward and reverse barrier heights, $V_f^\neq$ and $V_r^\neq$ respectively, reported as $\epsilon(V_f^\neq) / \epsilon(V_r^\neq)$ at varying basis set levels compared against DBH24 benchmark values (in kcal/mol).	66
3.5	Mean signed errors (MSE) and mean unsigned errors (MUE) resulting from CR-CC(2,3) calculations of all DBH24 forward and reverse barrier heights, $V_f^\neq$ and $V_r^\neq$ respectively, reported as $\epsilon(V_f^\neq) / \epsilon(V_r^\neq)$ at varying basis set levels (in kcal/mol).	68
3.6	Benchmark values for the C ₂ H ₂ and C ₂ H ₄ ozonolysis reaction pathways. . .	70
3.7	Energetics of stationary points relative to reactants, in kcal/mol, produced at various levels of coupled-cluster theory for the ozonolysis of ethylene and acetylene.	74
3.8	Relative enthalpies, in kcal/mol, with respect to reactant as calculated at various levels of theory for the conrotatory and disrotatory bicbut → tbut isomerization pathways.	78
3.9	Ground-state X^1A' energies for the asymmetric single-bond breaking of H ₂ O at a series of the O-H bond lengths, R , in atomic units. The last two rows give the mean unsigned errors (MUE) and non-parallelity errors (NPE) relative to the full CI PES obtained in Ref. [271].	83
3.10	Same as Table (3.9) for the three lowest-lying $^1A'$ states.	86

3.11	Same as Table (3.9) for the three lowest-lying $^3A'$ states.	87
3.12	Same as Table (3.9) for the two lowest-lying $^1A''$ states.	88
3.13	Same as Table (3.9) for the two lowest-lying $^3A''$ states.	89
3.14	The basis-set dependence of the vertical excitation energies $\omega_{\pi \rightarrow \pi^*}$ and the environment-induced shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ (in cm^{-1}) obtained with the EOMCCSD approach corresponding to the lowest $\pi \rightarrow \pi^*$ transition in the cis-7HQ chromophore and its complexes with the water and ammonia molecules.	95
3.15	The vertical excitation energies $\omega_{\pi \rightarrow \pi^*}$ and the environment-induced shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ (in cm^{-1}) obtained with the EOMCCSD/6-31+G(d), EOMCCSD/6-311+G(d), δ -CR-EOMCC(2, 3), A/6-31+G(d), and δ -CR-EOMCC(2, 3), D/6-31+G(d) approaches, and their composite EOMCC corresponding to the lowest $\pi \rightarrow \pi^*$ transition in the cis-7HQ chromophore and its various complexes.	97
3.16	Total and adiabatic excitation energies for the ground and low-lying excited states of CNC, as obtained with the different EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.	123
3.17	Total and adiabatic excitation energies for the ground and low-lying excited states of C_2N , as obtained with the different EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.	124
3.18	Comparison of the optimized equilibrium geometries for the low-lying states of CNC and C_2N , as obtained with the EA-EOMCC and SAC-CI-SDT-R/PS approaches using the DZP[4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets.	125
3.19	Total and adiabatic excitation energies for the ground and low-lying excited states of NCO, as obtained with the different IP EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.	126
3.20	Total and adiabatic excitation energies for the ground and low-lying excited states of N_3 , as obtained with the different IP EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.	127

3.21	Comparison of the optimized equilibrium geometries for the low-lying states of N ₃ and NCO, as obtained with the IP EOMCC and SAC-CI-SDT- <i>R</i> /PS approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets.	128
4.1	Computer costs of typical <i>ab initio</i> wave function calculations at the aug-cc-pVTZ basis set level, taken from Ref. [204]	133
4.2	The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut → t-but isomerization. The bicbut reactant defines the pivot geometry for the PES extrapolations.	143
4.3	The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut → t-but isomerization. The con-TS transition state defines the pivot geometry for the PES extrapolations.	144
4.4	The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut → t-but isomerization. The dis-TS transition state defines the pivot geometry for the PES extrapolations.	145
4.5	The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut → t-but isomerization. The g-but intermediate defines the pivot geometry for the PES extrapolations.	146
4.6	The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut → t-but isomerization. The gt-TS transition state defines the pivot geometry for the PES extrapolations.	147
4.7	The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut → t-but isomerization. The t-but product defines the pivot geometry for the PES extrapolations.	148

- 4.8 The calculated CR-CC(2,3)/cc-pVDZ, CR-CC(2,3)/cc-pVTZ, and CR-CC(2,3)/cc-pVQZ energies and the CBS values of the CR-CC(2,3) energies obtained using the point-wise extrapolations exploiting Eq. (3.98) and the CBS extrapolation procedure combining Eqs. (4.1)–(4.3) with Eq. (3.98) discussed in the text at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **bicbut** reactant was used to define the pivot geometry for the PES extrapolations based on Eqs. (4.1)–(4.3). 152
- 4.9 A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies of the H₂O molecule in which one of the two O-H bonds (**R**) is stretched, while keeping the other O-H bond at the equilibrium length and the H-O-H angle fixed at 104.5°. The equilibrium geometry defines the pivot geometry **R_e** and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest orbital, correlating with the 1s shell of the oxygen atom was kept frozen. 160
- 4.10 A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies for several internuclear separations $R_{\text{H-Cl}}$ of the HCl molecule. The equilibrium geometry defines the pivot geometry **R_e** and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest five orbitals, correlating with the 1s, 2s, and 2p shells of Cl, were kept frozen. 161
- 4.11 A comparison of calculated and extrapolated CR-CC(2,3)/aug-ccpVQZ energies for several internuclear separations $R_{\text{F-F}}$ of the F₂ molecule. The equilibrium geometry defines the pivot geometry **R_e** and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest two orbitals, correlating with the 1s shells of the F atoms, were kept frozen. 162
- 4.12 A comparison of calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **bicbut** reactant defines the pivot geometry **R_e** and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest four orbitals, correlating with the 1s shells of the carbon atoms, were kept frozen. 165

4.13	A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies of the H ₂ O molecule, in which one of the two O-H bonds R is stretched, while keeping the other O-H bond at the equilibrium length and the H-O-H angle fixed at 104.5°. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest orbital, correlating with the 1s orbital of the oxygen atom, was kept frozen.	169
4.14	A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies for several internuclear separations R of the HCl molecule. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest five orbitals, correlating with the 1s, 2s, and 2p shells of Cl, were kept frozen.	170
4.15	A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies for several internuclear separations R of the F ₂ molecule. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest two orbitals, correlating with the 1s orbitals on the fluorine atoms, were kept frozen. . . .	171
4.16	A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut→t-but isomerization. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest four orbitals, correlating with the 1s orbitals of the carbon atoms, were kept frozen. . . .	172
4.17	A summary of the necessary calculations and the corresponding computer resources required to utilize different tiers of the PES extrapolation scheme based on Eqs. (4.1) – (4.3) to scale the bicyclobutane isomerization pathway from the CR-CC(2,3)/cc-pVTZ level of theory to the CR-CC(2,3)/cc-pVQZ level, along with the corresponding extrapolation errors. The bicbut structure is used to provide the pivot geometry.	173
5.1	Explicit algebraic expressions for the one- and two-body matrix elements of $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$	177

LIST OF FIGURES

3.1	Stationary points along the C ₂ H ₂ (top row) and C ₂ H ₄ (bottom row) ozonolysis reaction pathways. In each row, the structures from left to right represent the van der Waals minimum (vdW), transition state (TS), and the cycloadduct, respectively. The oxygen, carbon, and hydrogen atoms are represented by the red, yellow, and grey spheres, respectively. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.	69
3.2	The conrotatory and disrotatory pathways describing the isomerization of bicyclo[1.1.0]butane to <i>trans</i> -buta-1,3-diene, along with the enthalpy values relative to the reactant at all stationary points obtained in the explicit CR-CC(2,3)/cc-pVQZ//CASSCF(10,10)/cc-pVDZ calculations (numbers in roman) and the enthalpy values obtained with the PES extrapolation procedure using the CR-CC(2,3)/cc-pVDZ, CR-CC(2,3)/cc-pVTZ, and RHF/cc-pVQZ energies at all stationary points resulting from the CASSCF(10,10)/cc-pVDZ optimizations, and the CR-CC(2,3)/cc-pVQZ correlation energy at the reactant geometry (numbers in bold italics). The available experimental enthalpy values are in parentheses. All enthalpies are in kcal/mol.	77
3.3	Cuts of the PESs for a few ground- and excited-states of a single-bond stretching model of H ₂ O as obtained with (a) the CCSD/EOMCCSD and (b) the CR-CC(2,3)/CR-EOMCC(2,3) approaches, and the TZ basis set. Lines are used to represent the CC/EOMCC data, while the corresponding full CI values are represented by points (which are identically replicated in (a) and (b)).	92
3.4	The eight hydrogen-bonded complexes of the cis-7HQ molecule.	94

Chapter 1

Introduction

Potential energy surfaces (PESs) play a central role in the theoretical description of molecular structures, properties, and reactivities, making them of great utility in many areas of chemical research, including spectroscopy and kinetics, investigations of reaction mechanisms, and the design of force fields for biological and materials science applications. Unfortunately, obtaining a molecular PES is typically very difficult due to the enormous mathematical complexity one faces when trying to solve the electronic Schrödinger equation. Development efforts in quantum chemistry have focused on the design of computationally manageable, yet reliable, approximation methods for the generation of energies and properties that can be applicable to a wide range of molecular systems. In the interest of reducing mathematical complexity some simplifying approximations are introduced from the outset, including the neglect of relativistic effects, which are small in systems with light atoms [1], and the decoupling of the nuclear and electronic wave functions, accomplished under the Born-Oppenheimer approximation [2], which allows the electronic energy of a molecule to be expressed as a function of its geometry, thus providing the conceptual basis for a PES. Despite these and

other commonly employed simplifications, the electronic Schrödinger equation remains too formidable to be solved exactly for any system with two or more electrons [3]. One of the principal challenges in the conventional determination of a PES, in which one is required to solve the electronic Schrödinger equation point-by-point for each fixed nuclear geometry of interest, is the development, implementation, and benchmarking of new approximate *ab initio* electronic structure methods that can provide suitably accurate results for the wide variety of chemical species commonly encountered when studying reactive PESs, while requiring only modest computational resources. Thus, the first goal of this dissertation is to make a significant contribution toward the development and benchmarking of leading modern methods of electronic structure theory.

It is difficult at the outset to determine what characteristics are important when developing new *ab initio* methods of electronic structure theory. One of the ultimate goals of quantum chemistry is to provide a predictive tool which can guide experimental efforts. In order to be considered a quantitatively predictive model, an *ab initio* method must be able to reliably produce numerical values for reaction energies which are accurate to within $\simeq 1$ kcal/mol of widely acknowledged benchmark data, a threshold often called “chemical accuracy”. It should also be clear what cases a method is and is not appropriate for, and, when necessary, it should be clear what to do to improve a poor result. A shortcoming of many *ab initio* methods is that they do not offer a balanced treatment of the short-range or “dynamical” and long-range or “non-dynamical” electronic interactions and therefore are unreliable for applications involving chemically reactive processes, since the relative importance of these competing effects can vary rapidly moving from one region to another on the corresponding PES. The goal is to develop a method which can give a balanced treatment

of both types of electronic interactions, while also having a straightforward and logical way of improving the result in the case that the predictions are shown to be inadequate.

The starting point for single-reference (SR) electronic structure methods is the independent particle model (IPM), usually Hartree-Fock (HF), approximation [4–7], and the more practical discretized algebraic form of the resulting equations based on the linear combination of atomic orbitals (LCAO) self-consistent field (SCF) formalism, which is the origin of the basis set approximation defining *ab initio* models [8]. The HF method produces the best single determinant description of the electronic wave function of interest and although it is by now well established that the HF approximation yields over 99% of the total energy, its ability to provide a reliable description of chemical phenomena is very limited due to the need to describe relatively small energy differences in characterizing chemical processes. The inadequacies of HF have been shown time and again, but as an outstanding early example, Wahl’s study of the F_2 molecule demonstrated that an SCF description predicts that F_2 is unbound [9]. Despite the completely unphysical relative energetics often produced by the HF approximation, the LCAO SCF formalism is still a valuable tool as it provides the HF model, a critical element to our conceptual understanding of electronic structure and chemical reactivity. To retain the HF picture and be able to make predictive calculations, methods which improve upon the basic molecular orbital approximation must be developed. The focus of these so-called ‘post-HF methods’ is to accurately recover the small amount of energy neglected at the HF level, a quantity which is known as the correlation energy. The most popular *ab initio* post-HF approaches can be grouped into the variational and perturbative classes. Variational approaches provide an upper bound to the exact energy, as in the configuration interaction (CI) theory [10–14], while perturbative approaches, such as the

many-body perturbation theory (MBPT) [15–22] and coupled cluster (CC) methods [23–27], have other advantages which will be discussed shortly. To understand how these approaches may be utilized to produce powerful new *ab initio* electronic structure approaches, their individual strengths and weaknesses must first be understood.

The CI method is the most straightforward way of accounting for electronic correlation. The CI wave function is a linear combination of the reference (usually HF) determinant and Slater determinants obtained by exciting electrons from occupied to unoccupied orbitals. If this determinantal expansion is complete (full CI), the correlation energy and total energy of the molecular system will be exact within whatever basis set approximation was used in the calculation. Due to the fact that the number of Slater determinants grows factorially with the size of the system, full CI calculations are impractical for systems of more than a few electrons. Approximations are commonly made in order to reduce this expense, for example, by truncating the CI expansion to include only the simplest classes of excitations, singles and doubles (SD), which is called the CISD method. A particularly appealing feature of the CI methods is that, by systematically adding classes of excitations, e.g., CI with singles, doubles, and triples, CI with singles, doubles, triples, and quadruples, and so on, one may obtain a series of increasingly accurate energies approaching the full CI value. Without this systematically improvable nature, it is difficult to claim a particular level of convergence in electronic structure calculations unless there is benchmark data available for comparison. Unfortunately, truncated CI methods lack size-extensivity, i.e., the correct linear scaling with the number of electrons, and the hierarchy converges slowly with the rank of excitation included. While the lack of size-extensivity can be approximately accounted for at the CISD level by adding Davidson corrections [28, 29], the slow convergence with level of excitation

to the full CI limit makes the CI method comparatively inefficient in the context of other modern electronic structure methods.

Perturbation theory offers another systematically improvable hierarchy of methods for determining the correlation energy. Although these methods are not variational, they are strictly size-extensive at every level. In the SR-MBPT approach, the electronic Hamiltonian is divided into an unperturbed part, corresponding to a single-determinantal reference description, usually the HF determinant, and a perturbation part, which usually describes the electronic correlation. In this form of MBPT, introduced by Møller and Plesset (MP) [15], the perturbation corrections to the reference wave function and energy are then calculated using the Rayleigh-Schrödinger perturbation theory. Systematic improvements in the electronic energy may be made by considering perturbative corrections of increasing orders, i.e., second-order MP (MP2), third-order MP (MP3), and so on. Additionally, comparisons of each order of perturbation with various CI ranks have made possible further division within levels of perturbation, allowing for the selection of certain classes of excitations. As an example, one may choose to calculate only certain components within the fourth-order of MP theory, e.g., MP4 with only contribution from doubles (MP4D), doubles and quadruples (MP4DQ), singles, doubles, and quadruples (MP4SDQ), or singles, doubles, triples, and quadruples (full MP4). This allows for certain computationally inexpensive higher-order corrections to be included without significantly increasing time requirements. In general, if the unperturbed HF determinant describing the molecular system is close to the exact wave function for that system, the convergence of the MBPT series is usually very rapid; however, when chemical bonds are stretched, the MBPT series becomes divergent. Thus, MBPT produces very good energies near the equilibrium geometries of molecules; however,

unlike CI, in regions of PESs where bonds are broken or formed MBPT energies become unphysical, making them inadequate for describing reactive PESs.

The conventional SR coupled-cluster (CC) theory is currently considered the preeminent *ab initio* method for ground-state calculations of modest-sized molecular systems. Contrary to the CI approach, which is characterized by a linear expansion of configuration state functions, the CC approach uses an exponential ansatz for the wave function, inherently assures that truncated forms of CC theory remain size-extensive, and produces a much faster convergence to the full CI wave function. It can be shown using a perturbation theory analysis that the improved convergence of CC as compared with the same level of truncation in CI theory is due to higher-order excitations being folded in as products of lower-order excitations by the exponential form of the CC ansatz. At the same time, thanks to the use of diagram factorization techniques commonly employed in efficient computer implementations of CC methods, the computer costs of CC calculations are similar to those characterizing the CI approaches truncated at the same excitation levels. This is why CC methods can offer higher accuracy at relatively lower costs as compared with CI or MBPT methods and even though the energies produced are not variational, they are typically considerably more accurate than those produced by CI at the same level of truncation. Excited states may also be accessed in CC theory through the equation-of-motion (EOM) CC formalism [30–34] or its symmetry-adapted-cluster configuration-interaction (SAC-CI) [35–39] or linear-response CC analogs [40–44]. The most appealing electronic structure methods come as a result of combining the different abovementioned approaches. EOMCC, for example, is just a CI-like expansion starting from an approximate ground-state wave function obtained using CC theory. Approaches based on the CC and EOMCC formalisms which utilize perturbation

theory to obtain inexpensive corrections due to higher-order correlation effects also exist and will be discussed in more detail shortly.

All of these appealing features of CC theory make it the most promising electronic structure approach of those considered so far, but there are still open problems yet to be completely addressed in CC theory. For instance, the most widely used and computationally practical CC approximation, the CC with singles and doubles (CCSD) method [45–48], fails completely when applied to describe a PES involving bond breaking. The CCSD method, which is based on an iterative procedure with central processing unit (CPU) steps scaling as $n_o^2 n_u^4$, where n_o and n_u are the numbers of occupied and unoccupied orbitals in the reference model, respectively, or as $\mathcal{N}^6$, where $\mathcal{N}$ is the size of the system expressed as the sum of the exponents of the basis functions, neglects the important triply excited, quadruply excited, and other higher-order clusters needed to describe bond breaking. Unfortunately, the CC method with singles, doubles, and triples (CCSDT) [49, 50] and the CC approach with singles, doubles, triples, and quadruples (CCSDTQ) [51–54] which include these clusters are prohibitively expensive for anything but small molecules, as they require iterative steps that scale as $n_o^3 n_u^5$ ($\mathcal{N}^8$), and $n_o^4 n_u^6$ ($\mathcal{N}^{10}$), respectively. A parallel problem exists in the case of the EOMCC or response CC methods where the basic singles and doubles approximation (EOMCCSD) [31–33], which is characterized by iterative $n_o^2 n_u^4$ ($\mathcal{N}^6$) scaling steps, fails to describe excited states dominated by two-electron and other many-electron transitions out of the ground state. Again, the EOMCC theory with singles, doubles, and triples (EOMCCSDT) [55–58] and the EOMCC approach with singles, doubles, triples, and quadruples (EOMCCSDTQ) [59], which can describe such states, do not offer a suitable alternative in the majority of applications because of their prohibitively expensive iterative $\mathcal{N}^8$

and $\mathcal{N}^{10}$ scaling steps, respectively. The massive success of the SRCC theories in molecular applications [60–66], particularly for nondegenerate ground-states of molecules near their equilibrium geometry and electronically excited states dominated by one-electron transitions, in addition to substantial progress in recent years in code parallelization [67–77], and in the development of various local correlation CC techniques (see, e.g., Refs. [78–83]), including, for example the cluster-in-molecule CC method developed by the Piecuch group [84–86], continues to stimulate parallel effort toward the extension of CC theory to handle quasidegenerate states characterizing bond breaking and many-electron excitations with the same or close to the same level of computational effort as that required by CCSD.

The most natural way to handle quasi-degenerate electronic states is to either turn to one of the variants of multi-reference (MR) perturbation theory, such as the popular MC-QDPT2 [87,88] and CASPT2 [89] methods, which are designed to handle large nondynamical correlation effects and low-order dynamical correlation effects, the MRCI approaches, such as the popular MRCI(Q) approximation [90,91], or the genuine MRCC methods, which, in analogy to the previously discussed SR analogs, offer a better treatment of the dynamical correlation effects as compared with the MRMBPT or MRCI methods. The genuine MRCC methods can be categorized into two types. The first is the hierarchy of the Fock-space or valence-universal methods [92,93], in which a single valence-universal wave operator operates on the system of interest and its ions which are obtained by removing one, two, etc. active electrons from active orbitals. This is a convenient formalism when one is interested in ionization potentials and electron affinities, however, these methods are unfavorable when a wide range of geometries must be considered, as is the case in the generation of reactive PESs. The other category of genuine MRCC methods, the Hilbert-space or state-universal

(SU) approaches [94], employ the Jeziorski-Monkhorst wave function ansatz in conjunction with the multi-root Bloch wave-operator formalism. The SU-MRCC approaches have been shown to produce very accurate results for the ground and excited states of systems undergoing severe geometrical transformations (see, e.g., Refs. [95–100]), making them a very attractive choice within the context of PES generation. However, their routine use for such applications is complicated by many factors. First, a model space of reference states must be pre-defined by the user, which often contains many states irrelevant to a given problem, and then a truncation scheme must be chosen which is compatible with the model space. This requires expert-level decisions on a case-by-case basis for each molecular problem. Then, even when appropriate choices are made, intruder states may appear [95–97], which can severely complicate interpretation of results. The aforementioned ambiguous parameters become even more undesirable when considered in the context of practical implementation. The need to accomodate such general choices in combination with the massive number of cluster amplitudes which must be computed make writing general SU-MRCC computer programs excessively difficult. While widespread routine use of genuine MRCC methods seems unlikely in the near future, these methods have provided great insight into ways to improve existing SRCC formalisms and inspired activity toward MRCC approaches dealing with a single quantum state. We refer the reader to Ref. [101], and references therein for further detailed discussion.

Due to the difficulties with implementing practical, user-friendly MRCC methods and the prohibitive computational expenses characterizing the SR CCSDT and CCSDTQ approaches and their EOM analogs, substantial research effort has been directed toward developing approximate approaches for including higher-order correlation effects within a SRCC formal-

ism. This led to both approximate iterative approaches, such as CCSDT- n [102–104], and CCSDTQ-1 [105], as well as the more popular non-iterative approaches, such as CCSD[T] [104, 106] and CCSD(T) [107], where triply excited T_3 clusters are approximated using perturbative arguments which take the form of a relatively inexpensive non-iterative $n_o^3 n_u^4$ ($\mathcal{N}^7$) scaling step in addition to the cost of the underlying CCSD calculation. Methods were also derived for cases when quadruple excitations should not be disregarded, where perturbative noniterative corrections for both triples and quadruples are added, e.g., the CCSD(TQ $_f$) method [108], characterized by a $n_o^2 n_u^5$ ($\mathcal{N}^7$) scaling step. These and other noniterative perturbative CC approaches have become extremely popular because they efficiently account for most of the important connected triple or triple and quadruple excitations in a user-friendly (“black-box”) fashion, while avoiding the steep iterative $\mathcal{N}^8$ or $\mathcal{N}^{10}$ scaling steps required by full CCSDT or CCSDTQ. Unfortunately, their applicability is limited to molecules near the equilibrium geometries, since the perturbative arguments used to derive the noniterative corrections of CCSD(T) and similar approaches fail when bonds are stretched or broken. A parallel challenge is found within the development of the standard response CC and EOMCC approximations for excited states, in which the effects of triply or triply and quadruply excited configurations are estimated using arguments originating from MBPT (see, e.g., Refs. [109–114]). The perturbatively corrected EOMCC methods, such as, for example, EOMCCSD($\tilde{T}$) [110], where the basic EOMCCSD approximation is corrected for triples via a noniterative $n_o^3 n_u^4$ or $\mathcal{N}^7$ scaling correction, break down for excited states having larger contributions due to doubly excited configurations, which is particularly common for excited-state PESs along bond breaking coordinates (see, e.g., Ref. [115]).

Quite a few approaches have been suggested in recent years which attempt to overcome

the failures of the conventional perturbative SR CC and EOMCC methods at larger internuclear separations and for excited states dominated by many-electron transitions, while avoiding the complexity of the genuine MRCC approaches. Examples which will be considered here include the externally corrected CC methods, such as the reduced MRCC (RMR-CC) approaches [116–122], the completely renormalized (CR) CC methods [65, 115, 123–132], and the active-space CC theories [54–57, 133–149] (see Ref. [101] for a recent review). All of these methods are related in that they are concerned with improving the description of bond-breaking processes and other cases involving electronic quasidegeneracies, while relying on a SR-like formulation. Each of these methods will now be discussed, paying particular attention to the strengths of each approach and making mention of selected recent applications in the literature.

The externally corrected SRCC methods represent an alternative to the perturbatively-derived corrections in which the CCSD equations corrected for terms containing the triply (T_3) and quadruply (T_4) excited clusters are solved after replacing T_3 and T_4 amplitudes by their values obtained in the cluster analysis of some non-CC wave function which exhibits good behavior at large internuclear separations, such as the projected unrestricted HF [150, 151], valence bond [152–154], multiconfigurational SCF or complete active-space SCF (CASSCF) [155–157], or MRCI wave functions [116–121]. Although all of these methods help in bond breaking situations, the latter, MRCI-corrected CCSD approach, referred to as RMR-CCSD, and its RMR-CCSD(T) extension [121, 122], have shown the most substantial improvements in the CCSD results. Unfortunately, the generation of the MRCISD wave function is very expensive when compared to the CCSD approach and MRCI is not size-extensive, which limits the applicability of the RMR-CC approaches to smaller systems. One

can partly address these deficiencies by turning to the so-called (N,M) -CCSD methods [158]. This notation implies that an M -reference general-model-space (GMS) SU-MRCCSD calculation [99,100] is corrected for higher-order clusters by drawing the relevant information from M pertinent wave functions [159] produced by an N -reference MRCISD calculation. Extensive benchmarking and testing have shown the RMR-CCSD method and its aforementioned extension with a perturbative correction for triples, RMR-CCSD(T) [121,122,160–172], as well as the (N,M) -CCSD methods [158,173–181] to be quite accurate in practice, although all of these approaches are very complex and require a lot of expertise.

The ground-state CR-CC [65,123–131] and excited-state CR-EOMCC [65,115,123,126,129,132] methods represent another class of approaches which were developed with the intention of removing the pervasive failings of the conventional CC/EOMCC perturbative methods. These approaches are based on the more general formalism of the method of moments of CC equations (MMCC) [65,123–125,127–129,131,182,183]. In analogy to the conventional perturbative methods, the CR-CC and CR-EOMCC methods allow one to calculate noniterative state-specific energy corrections corresponding to selected higher-order excitations which are added to the energies obtained from conventional CC/EOMCC calculations, such as CCSD or EOMCCSD. These CR-CC and CR-EOMCC corrections are based on the asymmetric energy expressions and resulting moment expansions which form the underlying framework for all MMCC methods. The original CR-CC approaches, such as CR-CCSD(T) [124,125] suffered from small errors due to a lack of strict size-extensivity [65], but these issues were addressed in the more recent left-eigenstate CR-CC approaches, including CR-CC(2,3), which are based on yet another form of the moment expansion of the full CI energy that defines the biorthogonal MMCC formalism [127–132]. The main

advantage of the CR-CC, CR-EOMCC, and other MMCC approaches is that the resulting ground- and excited-state energies are not dependent on any choice of active orbitals or other subjective parameters that one has to choose in MR calculations. Another advantage is the computational expense of the CR-CC and CR-EOMCC methods, which is on the order of the conventional perturbative SRCC methods. For example, the noniterative triples correction in the CR-CC(2,3) approach scales as $n_o^3 n_u^4$ or $\mathcal{N}^7$, in analogy to the scaling of CCSD(T). The CR-CC(2,3) method has already been proven to be very accurate and robust, particularly in applications involving single bond breaking [127–130, 184–188], mechanistic studies involving biradicals [127, 128, 130, 131, 189–193], and singlet-triplet gaps in biradical/magnetic systems [130, 131, 192]. Similar successes have been reported for the CR-EOMCC(2,3) method [129, 132] and its CR-EOMCCSD(T) predecessor [65, 115, 123, 126], where noniterative, $n_o^3 n_u^4$ or $\mathcal{N}^7$ scaling corrections for triples are added to the underlying EOMCCSD energies. For example, the CR-EOMCC(2,3) method has recently been shown to accurately reproduce adiabatic excitation energies for various closed- and open-shell molecules which are believed to be dominated by two-electron transitions out of the ground-state [132].

The third and final effort toward constructing a SR formalism capable of handling stronger non-dynamical correlations which will be considered here are the active-space CC and EOMCC methods [54, 133–149] (see Ref. [101] for a recent review). By specifically targeting the higher-order cluster and excitation amplitudes which become large in such situations by assigning a small subset of active orbitals defining these excitations, highly accurate results may be obtained while avoiding much of the expense of the higher-order parent methods. For example, the CCSDt and CCSDtq methods are based on the idea of

selecting T_3 and T_4 clusters within the CCSDT and CCSDTQ systems of equations using active orbitals [54, 133–146]. When properly implemented, the most expensive CPU steps of the full CCSDt and CCSDtq approaches scale as $N_o N_u n_o^2 n_u^4$ and $N_o^2 N_u^2 n_o^2 n_u^4$, respectively, where N_o and N_u are the numbers of active occupied and unoccupied orbitals, respectively. All active-space approaches have a few distinct advantages over competing methods. As an example, take again the CCSDt and CCSDtq approaches. These methods recover the exact results of their parent CCSDT and CCSDTQ approaches, respectively, in the limit that all orbitals are assigned as active. They are also systematically improvable, approaching the CCSDT and CCSDTQ limits as the number of active orbitals is increased. Given an appropriate selection of a usually small number of active orbitals, the active-space CC results are typically virtually perfect when compared to the values produced by the parent methods. Another advantage is the fact that all active-space CC methods are characterized by relatively low computational scalings, which are small prefactors times the $n_o^2 n_u^4$ steps of the CCSD type. Naturally, following the development of the ground-state CCSDt and CCSDtq active-space approaches, the EOMCCSDt and EOMCCSDtq methods were developed, with the first implementation of EOMCCSDt and the proposal for all such EOMCC-based methods occurring in the Piecuch group [55–58]. In addition to the EOMCCSDt approach, another class of the active-space EOMCC methods was also developed by the Piecuch group, namely the active-space electron attached (EA) and ionized (IP) EOMCC approaches [147–149], which may be used to generate ground and excited states of valence open-shell systems out of a related closed-shell ground-state, as in radical species. Generally, the EA- and IP-EOMCC methods (see Refs. [66, 147–149] and references therein for information) have the distinct advantage over the traditional open-shell EOMCC approaches based

on restricted open-shell HF (ROHF) or unrestricted HF (UHF) references in the fact that they automatically generate orthogonally spin-adapted wave functions (a difficult thing to accomplish in CC theory). For comparison, traditional ROHF- or UHF-based open-shell CC/EOMCC methods introduce spin-contamination to the resulting wave functions. Spin-contamination can have non-negligible effects on the energy of the states generated, but, more importantly, it prohibits the designation of the spin-symmetry of a given state, causing the identification of a particular state to become most inconvenient. This is an important issue, particularly when the excited states of open-shell systems are examined, which will be returned to periodically throughout this dissertation.

While promising electronic structure approaches have been and continue to be developed, many of which provide a highly accurate and balanced description of chemical species typically encountered while scanning molecular PESs, reliable methods are still prohibitively expensive when one faces the calculation of the hundreds or thousands of points which are typically needed to sample these PESs. Electronic structure methods could certainly benefit from an auxiliary approach for predicting points on the PES based on inexpensive calculations with either less expensive quantum chemistry approaches or smaller basis sets, which would ameliorate the staggering computational expense of generating hundreds or thousands of points using high levels of electronic structure theory. To address this problem, an *ab initio* extrapolation scheme has been proposed that predicts the PES corresponding to expensive high-level calculations from the results of a series of comparatively inexpensive lower-level calculations using the concept of correlation energy scaling. The PES extrapolation scheme of this type was originally suggested in Ref. [184] and was designed such that it can be used in conjunction with any typical post-HF or post-CASSCF electronic structure method.

Thus, the second principal goal of this dissertation is to develop and perform benchmark studies on the PES extrapolation schemes based on correlation energy scaling ideas, which allow one to generate chemically reactive molecular PESs with very little or almost no information from high-level calculations while retaining the desired accuracy that high-level *ab initio* electronic structure methods provide.

Finally, after new methods are developed and shown to be useful in benchmark studies, general-purpose computer programs should be written and distributed based on the successful theories to give scientists and engineers around the world a tool for interpreting, or even predicting, experimental results. The third and final goal in this dissertation is to outline the key details for the computer implementations of the open-shell EOMCCSD and EA- and IP-EOMCC methods which were written for the GAMESS electronic structure software package [194], a freely available suite of computer codes with tens of thousands of registered users. The three goals of this dissertation, summarized in Chapter 2, are addressed in this thesis, with a chapter being devoted to each. Thus, Chapter 3 is devoted to the development and application of new ground and excited-state CC/EOMCC methods for chemically reactive systems including the CR-CC/EOMCC methods and the active-space EA- and IP-EOMCC approaches. Chapter 4 begins with a discussion motivating the need for auxiliary methods to aid in the generation of molecular PESs and moves toward a presentation of the theory and various applications which help demonstrate various ways the PES extrapolation schemes based on correlation energy scaling may be used. Chapter 5 covers the development of computer codes for the GAMESS software package in detail, presenting programmable equations for a few of the theories implemented in this work as well as a brief discussion of how they are solved in practice.

Chapter 2

Project Objectives

The main objectives of this work are:

- A. Performing benchmark calculations using the new generations of CR-CC approaches, including barrier heights of hydrogen transfer, heavy-atom transfer, nucleophilic substitution, and unimolecular and association reactions, and PESs for addition and isomerization reactions involving species with varying degrees of electronic degeneracy in order to demonstrate what levels of theory are appropriate in different situations.
- B. Developing and performing benchmark applications for the new generations of CR-EOMCC approaches, including the calculations of vertical excitation energies and environment-induced spectral shifts of organic chromophores and the calculation of excited-state PESs along bond-breaking channels.
- C. Performing benchmark applications for the EA- and IP-EOMCC methods including geometry optimizations and the calculation of adiabatic excitation energies of small open-shell molecules.

- D. Developing the PES extrapolation schemes based on correlation energy scaling and performing benchmark applications to demonstrate the full range of capabilities offered and typical accuracies which should be expected in practice.
- E. Outlining the key details of computer implementations of the ROHF-based EOM-CCSD and RHF- or ROHF-based EA- and IP-EOMCC programs recently developed for GAMESS.

Chapter 3

Applications of Coupled-Cluster and Equation-of-Motion Coupled-Cluster Methods

3.1 Theory

As explained in the Introduction, the SR CC and EOMCC methods are the preeminent methods for the determination of electronic energies and properties in chemistry. The majority of this dissertation is concerned with the new generations of the CC and EOMCC methods which are useful in situations where the conventional CC and EOMCC approximations fail. We begin by reviewing the conventional CC and EOMCC theories in Sects. (3.1.1) and (3.1.2), respectively. The CR-CC and CR-EOMCC approaches and the active-space CC and EOMCC methods, with special emphasis on their EA- and IP-EOMCC extensions are discussed afterwards, in Sects. (3.1.3) and (3.1.4), respectively.

3.1.1 Single-Reference Coupled-Cluster Theory for Ground States

In the SRCC theory, the ground-state wave function $|\Psi_0\rangle$ of an N -electron system is expressed using the exponential ansatz,

$$|\Psi_0\rangle = e^T |\Phi\rangle, \quad (3.1)$$

where T is the cluster operator and $|\Phi\rangle$ is an IPM reference configuration, e.g., the HF determinant (throughout this thesis, the RHF or ROHF determinant). Typically, we truncate the many-body expansion of T at a conveniently chosen excitation level m_T , to obtain an approximate T , i.e., $T \simeq T^{(A)}$, hoping that one can reach the desired accuracies with $m_T \ll N$. The truncated cluster operator $T^{(A)}$ defining the approximate CC method A is given by

$$T^{(A)} = \sum_{n=1}^{m_T} T_n, \quad (3.2)$$

with

$$T_n = \sum_{i_1 < \dots < i_n, a_1 < \dots < a_n} t_{a_1 \dots a_n}^{i_1 \dots i_n} a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1}, \quad (3.3)$$

where T_n is the n -body component of $T^{(A)}$, $t_{a_1 \dots a_n}^{i_1 \dots i_n}$ are the cluster amplitudes, $i_1, i_2, i_3, \dots$ or $i, j, k, \dots$ ($a_1, a_2, a_3, \dots$ or $a, b, c, \dots$) are the spin-orbitals occupied (unoccupied) in the reference determinant $|\Phi\rangle$, and a^p (a_p) are the creation (annihilation) operators associated with the orthonormal spin-orbital basis set $\{|p\rangle\}$. In Eq. (3.2), m_T defines the maximum many-body component included in the truncated cluster operator $T^{(A)}$, returning the exact, full CI ground-state wave function when $m_T = N$ (recall that N is the number of electrons in the system of interest). When the cluster operator is truncated such that $m_T < N$, Eq.

(3.1) leads to the well-known hierarchy of standard CC approximations: CCSD when T is truncated at doubly excited clusters ($T \simeq T^{(\text{CCSD})} = T_1 + T_2, m_T = 2$); CCSDT when T is truncated at triply excited clusters ($T \simeq T^{(\text{CCSDT})} = T_1 + T_2 + T_3, m_T = 3$), CCSDTQ when T is truncated at quadruply excited clusters ($T \simeq T^{(\text{CCSDTQ})} = T_1 + T_2 + T_3 + T_4, m_T = 4$), etc.

The SRCC equations are formally obtained by inserting the CC wave function $|\Psi_0\rangle$, Eq. (3.1), into the electronic Schrödinger equation,

$$H|\Psi_0\rangle = E_0|\Psi_0\rangle, \quad (3.4)$$

premultiplying both sides of Eq. (3.4) on the left by $e^{-T^{(A)}}$ to obtain the connected cluster form of the Schrödinger equation [25–27],

$$\bar{H}^{(A)}|\Phi\rangle = E_0|\Phi\rangle, \quad (3.5)$$

with

$$\bar{H}^{(A)} = e^{-T^{(A)}} H e^{T^{(A)}} = (H e^{T^{(A)}})_C, \quad (3.6)$$

and projecting Eq. (3.5) onto the excited determinants $|\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle = a^{a_1} \dots a^{a_n} a_{i_1} \dots a_{i_n} |\Phi\rangle$ corresponding to the many-body components T_n included in $T^{(A)}$. Here, the subscript C indicates the connected part of a given operator expression. Eq. (3.6) is known as the similarity-transformed Hamiltonian of the CC theory. The resulting system of equations for cluster amplitudes $t_{a_1 \dots a_n}^{i_1 \dots i_n}$ defining $T^{(A)}$ has the following general form:

$$\langle \Phi_{i_1 \dots i_n}^{a_1 \dots a_n} | \bar{H}^{(A)} | \Phi \rangle = 0, \quad i_1 < \dots < i_n, \quad a_1 < \dots < a_n, \quad (3.7)$$

where $n = 1, \dots, m_T$ and $|\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle$ are the n -tuply excited determinants relative to the reference determinant $|\Phi\rangle$. As an example, the CCSD amplitude equations are obtained by projecting Eq. (3.5), where $T^{(A)} = T^{(\text{CCSD})} = T_1 + T_2$, onto all singly and doubly excited determinants, $|\Phi_i^a\rangle$ and $|\Phi_{ij}^{ab}\rangle$, respectively, such that

$$\begin{aligned}\langle \Phi_i^a | \bar{H}^{(\text{CCSD})} | \Phi \rangle &= 0, \\ \langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle &= 0,\end{aligned}\tag{3.8}$$

where

$$\bar{H}^{(\text{CCSD})} = e^{-(T_1+T_2)} H e^{(T_1+T_2)} = (H_N e^{T_1+T_2})_C \tag{3.9}$$

is the similarity-transformed Hamiltonian of the CCSD approach. These equations are solved for the one- and two-body cluster amplitudes, t_a^i and t_{ab}^{ij} , respectively, which appear in the definitions of the one- and two-body cluster operators,

$$T_1 = \sum_{i,a} t_a^i a^a a_i \tag{3.10}$$

and

$$T_2 = \sum_{i < j, a < b} t_{ab}^{ij} a^a a^b a_j a_i. \tag{3.11}$$

As explained in the Introduction, the most expensive CPU steps of the CCSD calculations, based on Eq. (3.8) scale as $n_o^2 n_u^4$ or $\mathcal{N}^6$, where $\mathcal{N}$ is a measure of the system size.

Once the general system of nonlinear, energy-independent equations for cluster amplitudes, Eq. (3.7), is solved for $T^{(A)}$, the energy $E_0^{(A)}$ corresponding to the standard SRCC

method A is calculated by projecting Eq. (3.5) onto the reference determinant $|\Phi\rangle$ such that

$$E_0^{(A)} = \langle \Phi | (H e^{T^{(A)}})_C | \Phi \rangle. \quad (3.12)$$

By introducing the normal product form of the Hamiltonian, $H_N = H - \langle \Phi | H | \Phi \rangle$, this equation can be rewritten as

$$\Delta E_0^{(A)} = \langle \Phi | (H_N e^{T^{(A)}})_C | \Phi \rangle, \quad (3.13)$$

where $\Delta E_0^{(A)}$ is the total energy of the system relative to reference energy, i.e., $\Delta E_0^{(A)} = E_0^{(A)} - \langle \Phi | H | \Phi \rangle$, which is equivalent to the correlation energy when $|\Phi\rangle$ is the HF state. For CCSD and all higher-order CC methods it is interesting to note that, at least for the quantum-chemistry Hamiltonians of interest in this dissertation, which include only two-body electron-electron interactions, the SRCC energy expression is

$$E_0^{(A)} = \langle \Phi | H | \Phi \rangle + \langle \Phi | [H_N (T_1 + T_2 + \frac{1}{2} T_1^2)]_C | \Phi \rangle \quad (3.14)$$

or

$$\Delta E_0^{(A)} = \langle \Phi | [H_N (T_1 + T_2 + \frac{1}{2} T_1^2)]_C | \Phi \rangle. \quad (3.15)$$

From Eq. (3.15) it can be seen that $\Delta E_0^{(A)}$ depends only on T_1 and T_2 clusters, independent of the excitation level m_T defining SRCC method A as long as $m_T \geq 2$. It follows then that the T_1 and T_2 clusters obtained at the level of the basic CCSD approximation are already sufficient to calculate the CC energy in a complete manner. This does not make the CCSD approach the exact theory. Adding higher-than-doubly excited clusters in the

cluster expansion can significantly improve the quality of the CC energy through coupling of higher-order clusters with the T_1 and T_2 amplitudes in the SRCC equations, Eq. (3.7). This suggests, however, an approximate treatment of higher-than-doubly excited cluster amplitudes may be sufficient to improve the quality of the CC energy in a satisfactory way.

As a final note about the ground-state CC theory, it should be mentioned that spin-contamination can become an issue depending on the choice of method used to generate the reference determinant. For example, if one is interested in describing a closed-shell molecule, the RHF reference $|\Phi\rangle$ may be employed, which is itself a spin eigenfunction, $\hat{S}^2|\Phi\rangle = S(S+1)\hbar^2|\Phi\rangle$, where $\hat{S}$ is the total spin operator, as is any approximate CC wave function produced from it through Eq. (3.1). In this case, spin-contamination is not a problem in the ground-state wave function or for any wave functions derived from it, since T and e^T commute with $\hat{S}^2$. However, if instead one wishes to describe an open-shell system, i.e., a radical or a system with even number of electrons of non-singlet multiplicity, the restricted open-shell HF (ROHF) or unrestricted HF (UHF) references may be employed, but one must then be careful when doing so. While the ROHF wave function is spin-adapted (UHF wave functions are not), in contrast to the RHF closed-shell case, a CC wavefunction constructed from an ROHF reference determinant $|\Phi\rangle$ will not be automatically spin-adapted due to the nonlinear nature of CC theory. The spin-contamination can introduce small errors in energies and present other difficulties when UHF-based and ROHF-based CC wave functions are employed. This issue will be further discussed in the next section of this dissertation which addresses the generation of electronically excited, electron-attached, and ionized states out of ground-state CC wave functions.

3.1.2 Equation-of-Motion Coupled-Cluster Theory for Electronically Excited, Electron-Attached, and Ionized States

The ground-state CC theory, described in Sect. (3.1.1), can be extended to excited, electron-attached, and ionized states by application of a linear excitation operator R_μ to the CC ground state, $|\Psi_0\rangle$. This leads to the EOMCC formalisms. Many different EOMCC methods may be formulated by modifying the basic definition of R_μ . In this dissertation, the particle-conserving excitation energy (EE) EOMCC as well as the particle-nonconserving electron-attached (EA) EOMCC and ionized (IP) EOMCC theories are considered. While discussions of the EA- and IP-EOMCC models are constrained to schemes where only one electron is added or removed, other approaches in these categories can easily be imagined with more than one electron added or removed. In the remainder of this section, an overview is presented introducing important concepts common to all EOMCC theories, while the specific details defining the EE-EOMCC and the EA- and IP-EOMCC theories are outlined in Sections (3.1.2.1) and (3.1.2.2), respectively.

In general, an ansatz may be written expressing the exact excited-state, electron-attached, or ionized wave function $|\Psi_\mu\rangle$ corresponding to state μ of interest as a linear excitation operator R_μ applied to the ground-state SRCC wave function, i.e.,

$$|\Psi_\mu\rangle = R_\mu|\Psi_0\rangle, \quad (3.16)$$

where $|\Psi_0\rangle$ is defined by Eq. (3.1). In the exact EOMCC theory, the cluster operator T and the exciting, electron-attaching, or ionizing operator R_μ are sums of all relevant many-body components needed to generate the system of interest. To obtain R_μ one must solve

the eigenvalue problem resulting from the substitution of Eq. (3.16) into the Schrödinger equation, $H|\Psi_\mu\rangle = E_\mu|\Psi_\mu\rangle$, to obtain

$$(\bar{H}_{N,\text{open}} R_{\mu,\text{open}})_C |\Phi\rangle = \omega_\mu R_\mu |\Phi\rangle, \quad (3.17)$$

in the subspace spanned by all determinants corresponding to the many-body components included in R_μ . Here, $\bar{H}_{N,\text{open}} = (H_N e^T)_{C,\text{open}} = e^{-T} H_N e^T - (H_N e^T)_{C,\text{closed}}$ is the similarity-transformed Hamiltonian of the CC theory in the normal-ordered form relative to the Fermi vacuum $|\Phi\rangle$, where the subscripts “open” and “closed” refer to the open (i.e., having external lines) and closed (i.e., having no external lines) parts of a given operator expression, obtained by solving the corresponding ground-state CC equations for T as described in the previous section, and $\omega_\mu = E_\mu - E_0$ is the vertical excitation (or electron-attachment or ionization) energy.

3.1.2.1 Excited States

In the particle conserving EE-EOMCC theory, excited-state energies and wave functions are obtained for an N -electron system by the application of a linear excitation operator R_μ of the form

$$R_\mu^{(A)} = R_{\mu,0}^{(A)} + R_{\mu,\text{open}}^{(A)} \equiv r_{\mu,0} \mathbf{1} + \sum_{n=1}^{m_R} R_{\mu,n}, \quad (3.18)$$

where $m_R \leq N$ and

$$R_{\mu,n} = \sum_{i_1 < \dots < i_n, a_1 < \dots < a_n} r_{\mu,a_1 \dots a_n}^{i_1 \dots i_n} a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1}, \quad (3.19)$$

onto the N -electron ground-state SRCC wave function $|\Psi_0^{(A)}\rangle$. The EE-EOMCC method is characterized by the linear excitation operator, $R_\mu^{(A)}$, in Eq. (3.18) having exactly the same number of creation and annihilation operators in each many-body component $R_{\mu,n}$. Just as in the cluster expansion of the ground-state wave function, in practice the many-body expansion of $R_\mu^{(A)}$ in Eq. (3.18) is truncated at some excitation level $m_R < N$ (usually the same level of excitation as in the cluster operator used in the preceding ground-state CC calculation, i.e., $m_R = m_T$). To determine the amplitudes $r_{\mu,a_1\dots a_n}^{i_1\dots i_n}$ with $n > 1$, we solve the eigenvalue problem given by Eq. (3.18) by diagonalizing the similarity-transformed Hamiltonian $\bar{H}^{(A)}$, Eq. (3.6), in a space spanned by the excited determinants $|\Phi_{i_1\dots i_n}^{a_1\dots a_n}\rangle$, with $n = 1, \dots, m_R$, corresponding to the many-body excitation operators included in $R_\mu^{(A)}$. In general, in order for Eq. (3.18) to hold and to obtain a size-intensive description [44, 196] of vertical excitation energies, m_R should not exceed m_T [34], but, as already mentioned, one typically chooses $m_R = m_T$.

For example, in the EOMCCSD method (note that the EE-EOMCC methods are often abbreviated as simply EOMCC), where R_μ is approximated as

$$R_\mu^{(\text{CCSD})} = R_{\mu,0} + R_{\mu,1} + R_{\mu,2}, \quad (3.20)$$

with

$$R_{\mu,1} = \sum_{i,a} r_{\mu,a}^i a^a a_i \quad (3.21)$$

and

$$R_{\mu,2} = \sum_{i < j, a < b} r_{\mu,ab}^{ij} a^a a^b a_j a_i, \quad (3.22)$$

we obtain the singly and doubly excited $r_{\mu,a}^i$ and $r_{\mu,ab}^{ij}$ amplitudes and EOMCSD vertical

excitation energies

$$\omega_{\mu}^{(\text{CCSD})} = E_{\mu}^{(\text{CCSD})} - E_0^{(\text{CCSD})} \quad (3.23)$$

by solving the system

$$\langle \Phi_i^a | (\bar{H}_{N,\text{open}}^{(\text{CCSD})} R_{\mu,\text{open}}^{(\text{CCSD})})_C | \Phi \rangle = \omega_{\mu}^{(\text{CCSD})} r_{\mu,a}^i \quad (3.24)$$

$$\langle \Phi_{ij}^{ab} | (\bar{H}_{N,\text{open}}^{(\text{CCSD})} R_{\mu,\text{open}}^{(\text{CCSD})})_C | \Phi \rangle = \omega_{\mu}^{(\text{CCSD})} r_{\mu,ab}^{ij}. \quad (3.25)$$

In other words, the EOMCCSD amplitudes and energies are determined by diagonalizing the matrix representing the similarity-transformed Hamiltonian of CCSD, Eq. (3.9), in a space of all singly and doubly excited determinants,

$$\bar{H}^{\text{CCSD}} = \begin{pmatrix} \bar{H}_{\text{SS}} & \bar{H}_{\text{SD}} \\ \bar{H}_{\text{DS}} & \bar{H}_{\text{DD}} \end{pmatrix}. \quad (3.26)$$

In analogy to the ground-state CCSD calculations, the most expensive CPU steps needed to do this scale as $n_o^2 n_u^4$ or $\mathcal{N}^6$.

While the EE-EOMCC methods can generate very accurate excited-state energetics and properties, spin-contamination of the ground-state when open-shell systems are treated with ROHF and UHF references can be sometimes problematic, introducing small errors into the calculations and preventing the identification of multiplicities of excited states, which is a useful guide when attempting to sort out states, particularly in situations involving near-degeneracies. One way to address this issue is to employ particle-nonconserving EOMCC approaches such as the EA- and IP-EOMCC methods, which build open-shell ground and excited states out of a related closed-shell CC wave function generated with a RHF reference,

automatically assuring that the resulting ground- and excited-state wave functions are spin-adapted. A brief description of the EA- and IP-EOMCC methods is given in the following section.

3.1.2.2 Electron-Attached and Ionized States

In the particle-nonconserving EA- and IP-EOMCC approaches, ground- and excited-state wave functions are generated by solving the eigenvalue problem given by

$$(\bar{H}_{N,\text{open}} R_\mu^{(N\pm 1)})_C |\Phi\rangle = \omega_\mu^{(N\pm 1)} R_\mu^{(N\pm 1)} |\Phi\rangle, \quad (3.27)$$

where $R_\mu^{(N\pm 1)}$ are particle-nonconserving operators, generating electronic states of $(N \pm 1)$ -electron systems, given by

$$R_\mu^{(N+1)} = \sum_{n=0}^{m_R} R_{\mu,(n+1)p-nh} \quad (3.28)$$

and

$$R_\mu^{(N-1)} = \sum_{n=0}^{m_R} R_{\mu,(n+1)h-np}, \quad (3.29)$$

where

$$R_{\mu,(n+1)p-nh} = \sum_{i_1 < \dots < i_n, a < a_1 < \dots < a_n} r_{aa_1 \dots a_n}^{i_1 \dots i_n} a^a a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1} \quad (3.30)$$

and

$$R_{\mu,(n+1)h-np} = \sum_{i_1 < \dots < i_n < i, a_1 < \dots < a_n} r_{a_1 \dots a_n}^{ii_1 \dots i_n} a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1} a_i, \quad (3.31)$$

with $m_R = N$ in the exact case and $m_R < N$ in the approximate schemes. In the above equations, $|\Phi\rangle$ is an N -electron reference determinant (e.g., the RHF reference) and T , used to define $\bar{H}_N$ in Eq. (3.27) and $|\Psi_0\rangle$ in Eq. (3.16), is the cluster operator of the SRCC theory, as

applied to the N -electron reference system. If a closed-shell reference system is used to define the wave functions $|\Psi_\mu\rangle$ through Eq. (3.16), with R_μ being one of the $R_\mu^{(N\pm 1)}$ operators, the similarity transformed Hamiltonian defining the EA-EOMCC and IP-EOMCC methods commutes with the S^2 and S_z operators. The result is that these methods produce open-shell eigenstates which are orthogonally spin-adapted, which means that spin-contamination issues which plague the standard spin-orbital-based open-shell EOMCC implementations that utilize ROHF or UHF references are avoided entirely.

In general, and in analogy to the EE-EOMCC case, when constructing approximate EA-EOMCC and IP-EOMCC schemes, the connected form of the eigenvalue problem displayed in Eq. (3.27) and the size intensivity of the resulting electron-attachment or ionization energies

$$\omega_\mu^{(N\pm 1)} = E_\mu^{(N\pm 1)} - E_0^{(N)} \quad (3.32)$$

are retained when $m_R \leq m_T$ [34]. The common approaches to designing the EA- and IP-EOMCC approximations are built upon approximate N -electron CCSD reference wave functions with $R_\mu^{(N\pm 1)}$ truncated such that $m_R = m_T - 1$ or $m_R = m_T$. In the basic EA-EOMCCSD($2p-1h$) and IP-EOMCCSD($2h-1p$) methods, we use $m_R = 1$ and $m_T = 2$. Thus, after solving the CCSD equations for an N -electron reference system, we diagonalize the similarity transformed Hamiltonian $\bar{H}^{(\text{CCSD})}$ in the $(N+1)$ -electron subspace of the $|\Phi^a\rangle = a^a|\Phi\rangle$ and $|\Phi_j^{ab}\rangle = a^a a^b a_j|\Phi\rangle$ determinants in the EA-EOMCCSD($2p-1h$) case and the $(N-1)$ -electron subspace of the $|\Phi_i\rangle = a_i|\Phi\rangle$ and $|\Phi_{ij}^b\rangle = a^b a_j a_i|\Phi\rangle$ determinants in the IP-EOMCCSD($2h-1p$) case, obtaining the $1p$ amplitudes r_a and the $2p-1h$ amplitudes r_{ab}^j , along with the corresponding energies $\omega_\mu^{(N+1)}$, in the former case, or the $1h$ amplitudes r_i , the $2h-1p$ amplitudes r_{ij}^b , and the corresponding energies $\omega_\mu^{(N-1)}$ in the latter

case. While these methods provide an initial approximation, the EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) approaches, in which $m_R = m_T = 2$, provide a much better description by incorporating the $3p-2h$ and $3h-2p$ components in the electron-attaching and ionizing operators, $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$, respectively. Unfortunately, the full inclusion of the $R_{\mu,3p-2h}$ and $R_{\mu,3h-2p}$ terms in the EA- and IP-EOMCC calculations comes at a rather high price, increasing the $\mathcal{N}^5$ -like $n_o n_u^4$ and $n_o^2 n_u^3$ operations defining the iterative diagonalization steps of EA-EOMCCSD($2p-1h$) and IP-EOMCCSD($2h-1p$), respectively, to the $\mathcal{N}^7$ -like $n_o^2 n_u^5$ and $n_o^3 n_u^4$ steps. In the EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) schemes, the similarity-transformed Hamiltonian $\bar{H}^{(\text{CCSD})}$, obtained in the CCSD calculations for the N -electron reference system, is diagonalized in the $N + 1$ -electron subspace spanned by the $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, and $|\Phi_{jk}^{abc}\rangle = a^a a^b a^c a_k a_j |\Phi\rangle$ determinants in the former case or the $N - 1$ -electron subspace spanned by the $|\Phi_i\rangle$, $|\Phi_{ij}^a\rangle$, and $|\Phi_{ijk}^{ab}\rangle = a^b a^c a_k a_j a_i |\Phi\rangle$ determinants in the latter case. From these diagonalizations, the r_a , r_{ab}^j , and $3p-2h$ amplitudes r_{abc}^{jk} , along with the corresponding energies ω_μ^{N+1} , or the r^i , r_b^{ij} , and $3h-2p$ amplitudes r_{bc}^{ijk} , along with the corresponding energies ω_μ^{N-1} are produced, which define the results of the EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) calculations, respectively. In many cases, it appears that the $3p-2h$ and $3h-2p$ effects brought through the $R_\mu^{(N\pm 1)}$ operators play a much more significant role than the triply excited components of the cluster operator T . In fact, it is usually not necessary to include the T_3 clusters in T until the $4p-3h$ and $4h-3p$ effects become important. In Sect. (3.1.4) the active-space variants of the EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) methods are presented, which is one way to retain the accuracy of these methods while avoiding their steep computer cost increase with n_o and n_u described above.

3.1.3 The Method of Moments of Coupled-Cluster Equations

The conventional CC and EOMCC methods summarized in Sects. (3.1.1) and (3.1.2) are useful, but one usually faces a challenge of having to correct the results of the low-level CC/EOMCC calculations, such as CCSD or EOMCCSD, for the higher-order correlation effects neglected by lower levels of CC/EOMCC theory without making the calculations prohibitively expensive. One would also like to make sure that the corrections to the CCSD, EOMCCSD, or other CC/EOMCC energies are robust in situations, such as bond breaking or excited states dominated by two-electron transitions, where the traditional perturbative corrections of the CCSD(T) type fail. The MMCC theory summarized below provides such robust and computationally attractive corrections to the CCSD, EOMCCSD, and other conventional CC/EOMCC energies.

The central focus of the MMCC theory is obtaining the non-iterative, state-specific, energy corrections

$$\delta_{\mu}^{(A)} \equiv E_{\mu} - E_{\mu}^{(A)}, \quad (3.33)$$

which recover the exact, full CI energies E_{μ} when added to the corresponding ground- and excited-state energies, $E_{\mu}^{(A)}$, obtained from the conventional CC/EOMCC approximation A . The goal of any method based on MMCC theory is to estimate these corrections using the underlying moment energy expansions, such that the resulting MMCC energies, defined as

$$E_{\mu}^{(\text{MMCC})} = E_{\mu}^{(A)} + \delta_{\mu}^{(\text{MMCC})}, \quad (3.34)$$

are good approximations to the corresponding exact energies E_{μ} .

All MMCC corrections $\delta_{\mu}^{(A)}$ are obtained via expansions of E_{μ} in terms of the general-

ized moments of CC/EOMCC equations defining a given CC/EOMCC approximation. The ground-state moments are simply projections of the connected cluster form of the Schrödinger equation, Eq. (3.5), on the excited determinants, $|\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle$ with $n > m_A$ disregarded in the conventional CC calculations

$$\mathfrak{M}_{0,a_1 \dots a_n}^{i_1 \dots i_n}(m_A) = \langle \Phi_{i_1 \dots i_n}^{a_1 \dots a_n} | \bar{H}^{(A)} | \Phi \rangle, \quad (3.35)$$

where $m_A = m_T$ is the maximum level of excitation included in the CC calculation being corrected. The excited-state moments needed to correct the EOMCC energies resulting from truncating T and R_μ at the m_A -body components (so that $m_T = m_R = m_A$) are projections of the EOMCC equations on the excited determinants $|\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle$,

$$\mathfrak{M}_{\mu,a_1 \dots a_n}^{i_1 \dots i_n}(m_A) = \langle \Phi_{i_1 \dots i_n}^{a_1 \dots a_n} | \bar{H}_{\text{open}}^{(A)} R_{\mu,\text{open}}^{(A)} | \Phi \rangle. \quad (3.36)$$

These moments are central to the ground- and excited-state MMCC theory and will be shown to be crucial quantities for evaluating the desired CR-CC and CR-EOMCC energy corrections in the following sections.

Several ways of expressing the $\delta_\mu^{(A)}$ corrections in terms of moments $\mathfrak{M}_{\mu,a_1 \dots a_n}^{i_1 \dots i_n}(m_A)$ have been proposed to date [65, 123–125, 127–129, 131, 182, 183]. The original and historically oldest formula, obtained in [124] for the ground states and [182] for excited states has the

following form:

$$\begin{aligned}
\delta_\mu^{(A)} &\equiv E_\mu - E_\mu^{(A)} \\
&= \sum_{n=m_A+1}^N \sum_{k=m_A+1}^n \langle \Psi_\mu | C_{n-k}(m_A) M_{\mu,k}(m_A) | \Phi \rangle / \\
&\quad \langle \Psi_\mu | R_\mu^{(A)} e^{T^{(A)}} | \Phi \rangle.
\end{aligned} \tag{3.37}$$

Here,

$$C_{n-k}(m_A) = (e^{T^{(A)}})_{n-k} \tag{3.38}$$

are the $(n - k)$ -body components of the exponential wave operator $e^{T^{(A)}}$, defining the CC method A , $|\Psi_\mu\rangle$ is the full CI ground- ($\mu = 0$) or excited- ($\mu > 0$) state, and

$$M_{\mu,k}(m_A) = \sum_{i_1 < \dots < i_k, a_1 < \dots < a_k} \mathfrak{M}_{\mu,a_1\dots a_k}^{i_1\dots i_k}(m_A) a^{a_1} \dots a^{a_n} a_{i_n} \dots a_{i_1} \quad (n \geq 1) \tag{3.39}$$

where moments $\mathfrak{M}_{\mu,a_1\dots a_k}^{i_1\dots i_k}(m_A)$ are defined by Eqs. (3.35) and (3.36). Thus, Eq. (3.37) states that one has to calculate quantities $C_{n-k}(m_A)$, Eq. (3.38), and moments $\mathfrak{M}_{\mu,a_1\dots a_k}^{i_1\dots i_k}(m_A)$, with $k > m_A$, to determine the noniterative energy correction $\delta_\mu^{(A)}$, Eq. (3.33). The $C_{n-k}(m_A)$ terms are very easy to calculate. The zero-body term, $C_0(m_A)$, equals 1; the one-body term, $C_1(m_A)$, equals T_1 ; the two-body term, $C_2(m_A)$, equals $T_2 + \frac{1}{2}T_1^2$ if $m_A \geq 2$; the three-body term $C_3(m_A)$ equals $T_1T_2 + \frac{1}{6}T_1^3$ if $m_A = 2$ and $T_3 + T_1T_2 + \frac{1}{6}T_1^3$ if $m_A \geq 3$, etc. The computation of moments $\mathfrak{M}_{\mu,a_1\dots a_n}^{i_1\dots i_n}(m_A)$ for the most interesting cases of correcting the CCSD or EOMCCSD energies ($m_A = 2$) is straightforward too, particularly if we limit ourselves to the corrections due to triples ($k = 3$) or quadruples ($k = 4$).

As an example, if one is interested in recovering the exact ground-state energy E_0 through

the addition of the full correction $\delta_0^{(\text{CCSD})}$ to the CCSD energy $E_\mu^{(\text{CCSD})}$ (where $m_A = 2$), one has to consider the generalized moments of the CCSD equations $\mathfrak{M}_{0,a_1\dots a_k}^{i_1\dots i_k}(2)$ with $k > 2$. After a quick diagrammatic analysis, we can show that this seemingly long expansion contains moments $\mathfrak{M}_{0,a_1\dots a_k}^{i_1\dots i_k}(2)$ with $k \leq 6$ only, since the electronic Hamiltonian contains only up to two-body interactions. Thus, all terms which make up the correction $\delta_0^{(\text{CCSD})}$ approaching the exact energy contain relatively few moments, namely,

$$\mathfrak{M}_{0,a_1\dots a_k}^{i_1\dots i_k}(2) = \langle \Phi_{i_1\dots i_k}^{a_1\dots a_k} | \bar{H}^{(\text{CCSD})} | \Phi \rangle, \quad k = 3 - 6. \quad (3.40)$$

The projections of the CCSD equations on higher-than-hextuply excited configurations do not have to be calculated, since for Hamiltonians containing up to two-body interactions the generalized moments $\mathfrak{M}_{0,a_1\dots a_k}^{i_1\dots i_k}(2)$ with $k > 6$ vanish. Similar simplifications occur in the case of correcting the excited-state EOMCCSD energies, where the only moments $\mathfrak{M}_{0,a_1\dots a_k}^{i_1\dots i_k}(2)$ that matter are those with $k = 3 - 8$. Although this is a considerable reduction of the computer effort, it is usually not computationally feasible to calculate up to six-body or higher moments to obtain a given MMCC correction. The CR-CC and CR-EOMCC methods discussed in Sects. (3.1.3.1)-(3.1.3.3) address this issue by focusing on the approximate corrections due to triples and quadruples which use moments $\mathfrak{M}_{\mu,a_1\dots a_k}^{i_1\dots i_k}(2)$ with $k = 3$ and 4 only and simplify Eq. (3.37) accordingly.

3.1.3.1 Completely Renormalized Coupled-Cluster and Equation-of-Motion Coupled-Cluster Approaches

In general, the CR-CC and CR-EOMCC approaches are obtained by approximating $|\Psi_\mu\rangle$ in Eq. (3.37) by a quasi-perturbative form that brings information about the desired cor-

relation effects we want $E_\mu^{(A)}$ to be corrected for. In particular, the CR-CCSD(T) [124], CR-CCSD(TQ) [124, 125] and CR-EOMCCSD(T) [115] methods, which are of interest in this dissertation, are obtained when the wave functions $|\Psi_\mu\rangle$ in Eq. (3.37) are approximated by low-order MBPT-like expressions. In the CR-CCSD(T) method which corrects the ground-state CCSD energy $E_0^{(\text{CCSD})}$ for triply excited clusters, $|\Psi_0\rangle$ in Eq. (3.37), where $m_A = 2$, is replaced by the following second-order-type, MBPT(2)[SDT]-like expression

$$|\Psi_0^{\text{CCSD(T)}}\rangle = (1 + T_1 + T_2 + T_3^{[2]} + Z_3)|\Phi\rangle, \quad (3.41)$$

where T_1 and T_2 are the singly and doubly excited clusters obtained in the CCSD calculations, the

$$T_3^{[2]}|\Phi\rangle = R_0^{(3)}(V_N T_2)_C|\Phi\rangle \quad (3.42)$$

term is an approximation of the connected triples (T_3) contribution, which is correct through second order, and

$$Z_3|\Phi\rangle = R_0^{(3)}V_N T_1|\Phi\rangle \quad (3.43)$$

is the disconnected triples correction, which distinguishes the CCSD(T) approach from its CCSD[T] predecessor. Conventional MBPT notation is used here, in which $R_0^{(3)}$ designates the three-body component of the MBPT reduced resolvent and V_N is the two-body part of H_N . In the CR-CCSD(TQ) method (note that here and elsewhere the so-called CR-CCSD(TQ),b variant is implied and that a discussion of the other variants, which can be found, for example, in [65, 123], will be omitted for the sake of brevity), which corrects the ground-state CCSD energy for the combined effect of triples and quadruples, $|\Psi_0\rangle$ in Eq. (3.37), where $m_A = 2$, is replaced by the following second-order-type, MBPT(2)[SDTQ]-like

expression

$$|\Psi_0^{\text{CCSD(TQ)}}\rangle = |\Psi_0^{\text{CCSD(T)}}\rangle + \frac{1}{2}T_2^2|\Phi\rangle, \quad (3.44)$$

where $|\Psi_0^{\text{CCSD(T)}}\rangle$ is given by Eq. (3.41).

Once these approximations have been established, the following compact formulas for the CR-CCSD(T) and CR-CCSD(TQ) energies can be written:

$$E_0^{(\text{CR-CCSD(T)})} = E_0^{(\text{CCSD})} + N^{\text{CR(T)}}/D^{(\text{T})} \quad (3.45)$$

and

$$E_0^{(\text{CR-CCSD(TQ)})} = E_0^{(\text{CCSD})} + N^{\text{CR(TQ)}}/D^{(\text{TQ})}, \quad (3.46)$$

where the $N^{\text{CR(T)}}$ and $N^{\text{CR(TQ)}}$ numerators are defined as

$$N^{\text{CR(T)}} = \langle \Phi | (T_3^{[2]})^\dagger M_{0,3}(2) | \Phi \rangle + \langle \Phi | (Z_3)^\dagger M_{0,3}(2) | \Phi \rangle \quad (3.47)$$

and

$$N^{\text{CR(TQ)}} = N^{\text{CR(T)}} + \frac{1}{2} \langle \Phi | (T_2^\dagger)^2 [T_1 M_{0,3}(2) + M_{0,4}(2)] | \Phi \rangle, \quad (3.48)$$

and the $D^{(\text{T})}$ and $D^{(\text{TQ})}$ denominators, representing the overlaps between the $|\Psi_0^{\text{CCSD(T)}}\rangle$ and $|\Psi_0^{\text{CCSD(TQ)}}\rangle$ wave functions, Eqs. (3.41), and (3.44), respectively, with the CCSD

ground state, as in Eq. (3.37), are calculated as

$$\begin{aligned}
D^{(T)} &\equiv \langle \Psi_0^{\text{CCSD}(T)} | e^{T_1+T_2} | \Phi \rangle \\
&= 1 + \langle \Phi | T_1^\dagger T_1 | \Phi \rangle + \langle \Phi | T_2^\dagger (T_2 + \frac{1}{2} T_1^2) | \Phi \rangle \\
&\quad + \langle \Phi | (T_3^{[2]})^\dagger (T_1 T_2 + \frac{1}{6} T_1^3) | \Phi \rangle \\
&\quad + \langle \Phi | Z_3^\dagger (T_1 T_2 + \frac{1}{6} T_1^3) | \Phi \rangle
\end{aligned} \tag{3.49}$$

and

$$\begin{aligned}
D^{(TQ)} &\equiv \langle \Psi_0^{\text{CCSD}(TQ)} | e^{T_1+T_2} | \Phi \rangle \\
&= D^{(T)} + \frac{1}{2} \langle \Phi | (T_2^\dagger)^2 (\frac{1}{2} T_2^2 + \frac{1}{2} T_1^2 T_2 + \frac{1}{24} T_1^4) | \Phi \rangle.
\end{aligned} \tag{3.50}$$

The quantities $M_3(2)$ and $M_4(2)$ in Eqs. (3.47) and (3.48) are expressed in terms of the triply and quadruply excited moments of the CCSD equations, easily calculated as in Eqs. (3.39) and (3.40), with k set at 3 and 4, respectively. Specifically,

$$M_{0,3}(2) | \Phi \rangle = \sum_{i < k < j, a < b < c} \mathfrak{M}_{0,abc}^{ijk}(2) | \Phi_{ijk}^{abc} \rangle, \tag{3.51}$$

where

$$\begin{aligned}
\mathfrak{M}_{0,abc}^{ijk}(2) &= \langle \Phi_{ijk}^{abc} | [H_N (T_2 + T_1 T_2 + \frac{1}{2} T_2^2 + \frac{1}{2} T_1^2 T_2 \\
&\quad + \frac{1}{2} T_1 T_2^2 + \frac{1}{6} T_1^3 T_2)]_C | \Phi \rangle,
\end{aligned} \tag{3.52}$$

and

$$M_{0,4}(2)|\Phi\rangle = \sum_{i<k<j<l, a<b<c<d} \mathfrak{M}_{0,abcd}^{ijkl}(2) |\Phi_{ijkl}^{abcd}\rangle, \quad (3.53)$$

where

$$\begin{aligned} \mathfrak{M}_{0,abcd}^{ijkl}(2) = & \langle \Phi_{ijkl}^{abcd} | [H_N(\frac{1}{2}T_2^2 + \frac{1}{2}T_1T_2^2 + \frac{1}{6}T_2^3 \\ & + \frac{1}{4}T_1^2T_2^2)]_C | \Phi \rangle. \end{aligned} \quad (3.54)$$

It is interesting to note that if we replace moments $\mathfrak{M}_{0,abc}^{ijk}(2)$ and $\mathfrak{M}_{0,abcd}^{ijkl}(2)$, entering the above equations through the $M_3(2)|\Phi\rangle$ and $M_4(2)|\Phi\rangle$ quantities as shown above, by their lowest-order estimates and the overlap denominators $D^{(T)}$ and $D^{(TQ)}$ by 1, the CR-CCSD(T) and CR-CCSD(TQ) approaches reduce to the standard CCSD(T) and CCSD(TQ) methods. For example, if we replace $\mathfrak{M}_{0,abc}^{ijk}(2)$ by the lowest-order estimate $\langle \Phi_{ijk}^{abc} | (V_N T_2)_C | \Phi \rangle$ and $D^{(T)}$ by 1, the CR-CCSD(T) energy, Eq. (3.45), simplifies to

$$E_0^{(\text{CCSD(T)})} = E_0^{(\text{CCSD})} + E_{\text{T}}^{[4]} + E_{\text{ST}}^{[5]}, \quad (3.55)$$

where

$$E_{\text{T}}^{[4]} = \langle \Phi | (T_3^{[2]})^\dagger (V_N T_2)_C | \Phi \rangle \quad (3.56)$$

and

$$E_{\text{ST}}^{[5]} = \langle \Phi | (Z_3)^\dagger (V_N T_2)_C | \Phi \rangle, \quad (3.57)$$

which is the well-known formula for the CCSD(T) energy. In other words, the CCSD(T) and CCSD(TQ) methods can be derived through the MMCC theory and are shown to be natural simplifications of the CR-CCSD(T) and CR-CCSD(TQ) methods, respectively. It is even

more important to note that the assumption that $D^{(T)}$ and $D^{(TQ)}$ can be approximated by 1 is correct only for molecules near the equilibrium geometries. As bonds are stretched or broken, $D^{(T)}$ and $D^{(TQ)}$ increase, damping the unphysical values of the conventional (T) and (TQ) correction. This is the key idea of “renormalization” of non-iterative energy corrections behind the CR-CC and CR-EOMCC approaches, which assume excessive values in conventional perturbative approaches of the CCSD(T) type when bonds are broken.

The CR-EOMCC methodologies for excited-states are obtained in a similar manner to their ground-state analogs, such as CR-CCSD(T) and CR-CCSD(TQ). For example, the CR-EOMCCSD(T) method is obtained by replacing the wave function $|\Psi_\mu\rangle$ in Eq. (3.37) with perturbative expressions resulting from an analysis of the EOMCCSDT equations, such as

$$\begin{aligned}
|\Psi_\mu\rangle = & \{R_{\mu,0} + (R_{\mu,1} + R_{\mu,0}T_1) + [R_{\mu,2} + R_{\mu,1}T_1 \\
& + R_{\mu,0}(T_2 + \frac{1}{2}T_1^2)] + [\tilde{R}_{\mu,3} + R_{\mu,2}T_1 \\
& + R_{\mu,1}(T_2 + \frac{1}{2}T_1^2) \\
& + R_{\mu,0}(T_1T_2 + \frac{1}{6}T_1^3)]\}|\Phi\rangle,
\end{aligned} \tag{3.58}$$

where T_1 and T_2 are the singly and doubly excited clusters obtained in the CCSD calculations and $R_{\mu,0}$, $R_{\mu,1}$, and $R_{\mu,2}$ are the reference, singly excited, and doubly excited components of the EOMCCSD excitation operator $R_\mu^{(\text{CCSD})}$. The approximate triply excited components of the EOMCC excitation operator, which enter Eq. (3.58) and which can be determined

through perturbative analysis of the EOMCCSDT equations, are defined as follows:

$$\tilde{R}_{\mu,3} = \sum_{i < j < k, a < b < c} \tilde{r}_{\mu,abc}^{ijk} a^a a^b a^c a_k a_j a_i, \quad (3.59)$$

where

$$\tilde{r}_{\mu,abc}^{ijk} = \mathfrak{M}_{\mu,abc}^{ijk}(2) / D_{\mu,abc}^{ijk}, \quad (3.60)$$

with moments $\mathfrak{M}_{\mu,abc}^{ijk}(2)$ given by

$$\begin{aligned} \mathfrak{M}_{\mu,ijk}^{abc}(2) &= \langle \Phi_{ijk}^{abc} | (\bar{H}_2^{(\text{CCSD})} R_{\mu,2})_C | \Phi \rangle + \langle \Phi_{ijk}^{abc} | [\bar{H}_3^{(\text{CCSD})} (R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle \\ &+ \langle \Phi_{ijk}^{abc} | (\bar{H}_4^{(\text{CCSD})} R_{\mu,1})_C | \Phi \rangle + r_{\mu,0}^{(\text{CCSD})} \langle \Phi_{ijk}^{abc} | \bar{H}_3^{(\text{CCSD})} | \Phi \rangle, \end{aligned} \quad (3.61)$$

and the perturbative denominator $D_{\mu,abc}^{ijk}$ given by

$$\begin{aligned} D_{\mu,abc}^{ijk} &= E_{\mu}^{(\text{CCSD})} - \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \\ &= \omega_{\mu}^{(\text{CCSD})} - \langle \Phi_{ijk}^{abc} | \bar{H}_{\text{open}}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \\ &= \omega_{\mu}^{(\text{CCSD})} - \langle \Phi_{ijk}^{abc} | \bar{H}_1^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \\ &\quad - \langle \Phi_{ijk}^{abc} | \bar{H}_2^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \\ &\quad - \langle \Phi_{ijk}^{abc} | \bar{H}_3^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle. \end{aligned} \quad (3.62)$$

Here, $\omega_{\mu}^{(\text{EOMCCSD})}$ represents the EOMCCSD vertical excitation energy defined by Eq.

(3.23),

$$\omega_{\mu}^{(\text{EOMCCSD})} = E_{\mu}^{(\text{EOMCCSD})} - E_0^{(\text{CCSD})}, \quad (3.63)$$

$\bar{H}_{\text{open}}^{(\text{CCSD})}$ is the open part of $\bar{H}^{(\text{CCSD})}$ (all diagrams of $\bar{H}^{(\text{CCSD})}$ that have external lines),

and $\bar{H}_k^{(\text{CCSD})}$ is a k -body component of $\bar{H}^{(\text{CCSD})}$. The variant of CR-EOMCCSD(T) described here is variant CR-EOMCCSD(T),ID. While other variants exist [115], we do not discuss them here for the sake of brevity. Relationships exist between the CR-EOMCC methods and their conventional perturbative EOMCC counterparts, such as EOMCCSD(T) or EOMCCSD($\tilde{T}$), and further details can be found elsewhere [115]. It should be emphasized that the CR-CCSD(T), CR-CCSD(TQ), and CR-EOMCCSD(T) approaches are not only related to the conventional CCSD(T), CCSD(TQ), and EOMCCSD(T)-type methods in a straightforward manner, as described above, but they also have similar computer costs. For example, in analogy to CCSD(T), the CR-CCSD(T) and CR-EOMCCSD(T) approaches have CPU steps that scale as $n_o^2 n_u^4$ ($\mathcal{N}^6$) in the CCSD/EOMCCSD part and $n_o^3 n_u^4$ ($\mathcal{N}^7$) in the triples correction parts. Similarly, in analogy to the factorized formulation of CCSD(TQ) [108], the costs of CR-CCSD(TQ) scale as $n_o^2 n_u^4$ ($\mathcal{N}^6$) in the CCSD part and $n_o^3 n_u^4 + n_o^2 n_u^5$ ($\mathcal{N}^7$) in the (TQ) parts.

In the next section, an alternative, biorthogonal formulation of the MMCC equations is presented and the left-eigenstate CR-CC and CR-EOMCC approaches, specifically the CR-CC(2,3) and CR-EOMCC(2,3) methods that result from it, are discussed. These methods have been shown to be even more accurate than the CR-CCSD(T), CR-EOMCCSD(T), and other CR-CC/CR-EOMCC approaches derived out of Eq. (3.37), so that much of our benchmarking effort of the CR-CC and CR-EOMCC methods presented in this thesis focuses on the CR-CC(2,3) and CR-EOMCC(2,3) theories discussed in the next section.

3.1.3.2 Biorthogonal Formulation of the MMCC Equations and the CR-CC(2,3) and CR-EOMCC(2,3) Methods

An alternative formulation of CR-CC methods that satisfies the property of size-extensivity in the ground state, which the CR-CCSD(T) and CR-CCSD(TQ) approaches violate somewhat, was developed in Refs. [127–132]. The resulting approaches, such as CR-CC(2,3) and CR-EOMCC(2,3) originate from the so-called biorthogonal MMCC formalism presented in Refs. [127, 128]. The biorthogonal formulation of the MMCC formalism redefines the correction $\delta_\mu^{(A)}$ by introducing the following ansatz for the exact bra state, $\langle \Psi_\mu |$, entering Eq. (3.37):

$$\langle \Psi_\mu | = \langle \Phi | \mathcal{L}_\mu e^{-T^{(A)}}, \quad (3.64)$$

where $\mathcal{L}_\mu$ is a deexcitation operator defined as

$$\mathcal{L}_\mu = \sum_{n=0}^N \mathcal{L}_{\mu,n}, \quad (3.65)$$

with

$$\mathcal{L}_{\mu,n} = \sum_{i_1 < \dots < i_n, a_1 < \dots < a_n} \ell_{\mu, i_1 \dots i_n}^{a_1 \dots a_n} a^{i_1} \dots a^{i_n} a_{a_n} \dots a_{a_1}. \quad (3.66)$$

By substituting Eq. (3.64) into Eq. (3.37) [131], or by considering the appropriate asymmetric energy expressions [127, 128], one can rewrite the $\delta_\mu^{(A)}$ correction in the following

way:

$$\begin{aligned}
\delta_\mu^{(A)} &= \sum_{n=m_A+1}^{N_{\mu,A}} \langle \Phi | \mathcal{L}_{\mu,n} M_{\mu,n}(m_A) | \Phi \rangle \\
&= \sum_{n=m_A+1}^{N_{\mu,A}} \sum_{i_1 < \dots < i_n, a_1 < \dots < a_n} \ell_{\mu, i_1 \dots i_n}^{a_1 \dots a_n} \mathfrak{M}_{\mu, a_1 \dots a_n}^{i_1 \dots i_n}(m_A), \tag{3.67}
\end{aligned}$$

where $M_{\mu,n}(m_A)$ has already been defined by Eqs. (3.39). Thus, for example, when one wishes to obtain the CR-CC corrections to CCSD/EOMCCSD energies due to triple excitations, then $A = \text{CCSD}$ and $m_A = 2$, as before, and, the ground-state ($\mu = 0$) CR-CC(2,3) or excited-state ($\mu > 0$) CR-EOMCC(2,3) energies, $E_\mu(2, 3)$, can be calculated in the following manner:

$$E_\mu(2, 3) = E_\mu^{(\text{CCSD})} + \delta_\mu(2, 3), \tag{3.68}$$

where

$$\delta_\mu(2, 3) = \langle \Phi | \mathcal{L}_{\mu,3} M_{\mu,3}(2) | \Phi \rangle = \sum_{i < j < k, a < b < c} \ell_{\mu, ijk}^{abc} \mathfrak{M}_{\mu, abc}^{ijk}(2), \tag{3.69}$$

with $\mathfrak{M}_{\mu, abc}^{ijk}(2)$ representing the triply excited moments of the CCSD ($\mu = 0$) or EOMCCSD ($\mu > 0$) equations defined by Eqs. (3.52) and (3.61), respectively.

Since the triply excited moments of the CCSD/EOMCCSD equations are already well-defined, the determination of the three-body amplitudes $\ell_{\mu, ijk}^{abc}$ entering Eq. (3.69) becomes the primary focus. Since the exact values of these amplitudes may not be obtained without solving the full CI problem for the exact bra state $\langle \Psi_\mu |$, a method for the approximate determination of the $\ell_{\mu, ijk}^{abc}$ amplitudes has to be proposed. Following Refs. [127, 128], the derivation of the $\ell_{\mu, ijk}^{abc}$ amplitudes used in practical CR-CC(2,3) and CR-EOMCC(2,3) calculations begins by defining the approximate form of the deexcitation operator $\mathcal{L}_\mu$ which

parameterizes the full CI state $\langle \Psi_\mu |$ as

$$\mathcal{L}_\mu \approx L_\mu^{(\text{CCSD})} + \mathcal{L}_{\mu,3}, \quad (3.70)$$

where $\mathcal{L}_{\mu,3}$ is the three-body component of $\mathcal{L}_\mu$ of interest, which is expressed in the usual way as

$$\mathcal{L}_{\mu,3} = \sum_{i < j < k, a < b < c} \ell_{\mu,ijk}^{abc} a^i a^j a^k a_c a_b a_a, \quad (3.71)$$

with $\ell_{\mu,ijk}^{abc}$ representing the desired triply deexcited amplitudes to be determined, and $L_\mu^{(\text{CCSD})}$ is the deexcitation operator that defines the left or bra CCSD/EOMCCSD states via the equation [33, 34]

$$\langle \Psi_\mu | = \langle \Phi | L_\mu^{(\text{CCSD})} e^{-T_1 - T_2}, \quad (3.72)$$

where

$$L_\mu^{(\text{CCSD})} = \delta_{\mu,0} \mathbf{1} + L_{\mu,1} + L_{\mu,2}, \quad (3.73)$$

is obtained by solving the so-called left CCSD/EOMCCSD equations [33, 34] for the corresponding $\ell_{\mu,i}^a$ and $\ell_{\mu,ij}^{ab}$ amplitudes (the $\mu = 0$ variant of the system of the left CCSD/EOMCCSD equations is equivalent to the system of the so-called “lambda” equations of the analytic gradient CCSD theory [197]). The explicit, computationally tractable form of the approximate $\ell_{\mu,ijk}^{abc}$ amplitudes, which enter the CR-CC(2,3) and CR-EOMCC(2,3) corrections, are then obtained by substituting the approximate expression for $\mathcal{L}_\mu$ given by Eq. (3.70) into the exact form of the similarity-transformed bra eigenvalue problem for $\mathcal{L}_\mu$,

$$\langle \Phi | \mathcal{L}_\mu \bar{H}^{(\text{A})} = E_\mu \langle \Phi | \mathcal{L}_\mu, \quad (3.74)$$

where A is set at CCSD (and m_A at 2), and the resulting equation is right projected onto the triply excited determinants $|\Phi_{ijk}^{abc}\rangle$ to obtain

$$\begin{aligned} \langle \Phi | L_\mu^{(\text{CCSD})} \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle &+ \sum_{l < m < n, d < e < f} \langle \Phi_{lmn}^{def} | \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \ell_{\mu,lmn}^{def} \\ &= E_\mu \ell_{\mu,ijk}^{abc}. \end{aligned} \quad (3.75)$$

The exact energy E_μ in Eq. (3.75) is then replaced by the corresponding CCSD/ EOMCCSD energy $E_\mu^{(\text{CCSD})}$, and the triples-triples block of the matrix representing $\bar{H}^{(\text{CCSD})}$ in the second term on the left-hand side of Eq. (3.75) is replaced by its diagonal, as in the Epstein-Nesbet partitioning [198,199]. The result of all of these operations is the following formula for the approximate $\ell_{\mu,ijk}^{abc}$ amplitudes in terms of the many-body components of the similarity-transformed Hamiltonian of CCSD, one- and two-body components of $L_\mu^{(\text{CCSD})}$, and in the $\mu > 0$ case, the EOMCCSD excitation energies $\omega_\mu^{(\text{CCSD})}$:

$$\ell_{\mu,ijk}^{abc} = N_{\mu,ijk}^{abc} / D_{\mu,abc}^{ijk}, \quad (3.76)$$

where the numerator $N_{\mu,ijk}^{abc}$ and denominator $D_{\mu,abc}^{ijk}$ are defined as follows:

$$\begin{aligned} N_{\mu,ijk}^{abc} &= \langle \Phi | L_\mu^{(\text{CCSD})} \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \\ &= \langle \Phi | [(L_{\mu,1} \bar{H}_2^{(\text{CCSD})})_{DC} + (L_{\mu,2} \bar{H}_1^{(\text{CCSD})})_{DC} \\ &\quad + (L_{\mu,2} \bar{H}_2^{(\text{CCSD})})_C] | \Phi_{ijk}^{abc} \rangle, \end{aligned} \quad (3.77)$$

and

$$\begin{aligned}
D_{\mu,abc}^{ijk} &= E_{\mu}^{(\text{CCSD})} - \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle \\
&= \omega_{\mu}^{(\text{CCSD})} - \sum_{n=1}^3 \langle \Phi_{ijk}^{abc} | \bar{H}_n^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle.
\end{aligned} \tag{3.78}$$

Note that the denominator $D_{\mu,abc}^{ijk}$ used here is the same as that of the CR-EOM-CCSD(T) approach (see Eq. 3.62). The CR-CC(2,3) ($\mu = 0$) and CR-EOMCC(2,3) ($\mu > 0$) approaches are obtained by substituting Eqs. (3.52) and (3.61) for $\mathfrak{M}_{\mu,abc}^{ijk}(2)$ and Eq. (3.76) for $\ell_{\mu,ijk}^{abc}$, where $N_{\mu,ijk}^{abc}$ and $D_{\mu,abc}^{ijk}$ are given by Eqs. (3.77) and (3.78), respectively, into the triples correction formula, Eq. (3.69), which is subsequently added to the CCSD/EOMCCSD energy $E_{\mu}^{(\text{CCSD})}$ to obtain the total energy $E_{\mu}(2,3)$, as in Eq. (3.68).

In both the CR-CC(2,3) and CR-EOMCC(2,3) theories, the exact treatment of the Epstein-Nesbet-like denominator $D_{\mu,abc}^{ijk}$, as in Eq. (3.78), where no terms in $D_{\mu,abc}^{ijk}$ are neglected, characterizes the most complete variant of the CR-CC(2,3) and CR-EOMCC(2,3) approaches designated as CR-CC(2,3),D or CR-EOMCC(2,3),D, respectively. By neglecting selected terms in Eq. (3.78) for $D_{\mu,abc}^{ijk}$, we obtain approximate CR-CC(2,3) and CR-EOMCC(2,3) schemes. Let us focus on CR-EOMCC(2,3) for this discussion, which contains CR-CC(2,3) as a special case corresponding to $\mu = 0$. Variant C of the CR-EOMCC(2,3) theory, designated as the CR-EOMCC(2,3),C approach, is obtained by ignoring the three-body component of $\bar{H}^{(\text{CCSD})}$ in Eq. (3.78), i.e., the $\langle \Phi_{ijk}^{abc} | \bar{H}_3^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle$ term, while keeping the contributions to $D_{\mu,abc}^{ijk}$ from the one- and two-body components of $\bar{H}^{(\text{CCSD})}$ intact. The CR-EOMCC(2,3),B approach is obtained by ignoring the two- and three-body components of $\bar{H}^{(\text{CCSD})}$ in Eq. (3.78), leaving only the one-body contribution $\langle \Phi_{ijk}^{abc} | \bar{H}_1^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle$ in

$D_{\mu,abc}^{ijk}$. Finally, variant A of the CR-EOMCC(2,3) approach is obtained by replacing the Epstein-Nesbet-like denominator $D_{\mu,abc}^{ijk}$, Eq. (3.78), by the Møller-Plesset-like denominator for triple excitations, $\omega_{\mu}^{(\text{CCSD})} - (\epsilon_a + \epsilon_b + \epsilon_c - \epsilon_i - \epsilon_j - \epsilon_k)$, where ϵ_p 's are the single-particle energies associated with spin-orbitals p (diagonal elements of the Fock matrix). An analogous discussion may be made for the ground-state case resulting in the A, B, C, and D variants of CR-CC(2,3).

In analogy to the CR-CCSD(T) approach discussed in the previous subsection, one can extend the CR-CC(2,3) method to higher-than-triple excitations as in, for example, the CR-CC(2,4) scheme [127, 128, 166]. The CR-CC(2,4) approach, when implemented fully, combines the $\mathcal{N}^6$ -type steps of CCSD with the $\mathcal{N}^7$ ($n_o^3 n_u^4$) steps of CR-CC(2,3) needed to determine the triples correction, and the $\mathcal{N}^9$ ($n_o^4 n_u^5$) steps needed to calculate the analogous correction due to quadruples. In order to address this CPU-time increase, Piecuch et al. proposed the so-called CR-CC(2,3)+Q method, also tested in this thesis, where one calculates the ground-state energy as follows [186]

$$E^{\text{CR-CC(2,3)+Q}} = E^{\text{CR-CC(2,3)}} + E^{\text{CR-CCSD(TQ)}} - E^{\text{CR-CCSD(T)}}, \quad (3.79)$$

i.e., one adds the quadruples correction extracted from the CR-CCSD(TQ) calculations to the CR-CC(2,3) energy. This has an advantage over CR-CC(2,4) in the fact that the CPU-time costs of the CR-CCSD(TQ) calculations in the quadruples correction part scale as $\mathcal{N}^7$ ($n_o^2 n_u^5$) with the system size $\mathcal{N}$, as opposed to the $\mathcal{N}^9$ steps of CR-CC(2,4). As a result, the CR-CC(2,3)+Q approach is almost as affordable as the CR-CC(2,3) method itself, while bringing information about connected quadruply excited clusters that become important in multiple bond breaking situations [186, 187].

The CR-CC(2,3) approach is capable of breaking bonds and, unlike its CR-CCSD(T) predecessor, is size-extensive. Unfortunately, the $\delta_\mu(2,3)$ corrections to the EOMCCSD energies, defining CR-EOMCCSD(2,3), violate the property of size-intensivity of the EOMCC excitation energies [44, 115, 196]. Although this violation is often unimportant, it is useful to consider the possibility of restoring size intensivity in CR-EOMCC(2,3). This aspect is discussed next.

3.1.3.3 A Size-Intensive Variant of CR-EOMCC(2,3): The δ -CR- EOMCC(2,3) Method

As shown in Refs. [132, 200], although the ground-state variants of CR-CC(2,3) are size-extensive, their excited-state CR-EOMCC(2,3) analogs do not satisfy the property of size-intensivity [44, 126, 196], i.e., the vertical excitation energy of a non-interacting system $A+B$, in which fragment A is excited, resulting from the CR-EOMCC(2,3) calculations, is not the same as that obtained for the isolated system A . The lack of size intensivity of the CR-EOMCC(2,3) excitation energies can be traced back to the presence of the size-extensive contribution [200, 201]

$$\beta_\mu = \sum_{i < j < k, a < b < c} (r_{\mu,0} \ell_{\mu,ijk}^{abc} - \ell_{0,ijk}^{abc}) \mathfrak{M}_{0,abc}^{ijk}(2) \quad (3.80)$$

in the CR-EOMCC(2,3) vertical excitation energy

$$\omega_\mu^{(\text{CR-EOMCC}(2,3))} = E_\mu^{(\text{CR-EOMCC}(2,3))} - E_0^{(\text{CR-CC}(2,3))}. \quad (3.81)$$

Using the above equations for the CR-EOMCC(2,3) energies, particularly Eq. (3.80), we can decompose the CR-EOMCC(2,3) excitation energy as follows [200, 201]:

$$\omega_{\mu}^{(\text{CR-EOMCC}(2,3))} = \omega_{\mu}^{(\text{CCSD})} + \alpha_{\mu} + \beta_{\mu}. \quad (3.82)$$

Here, $\omega_{\mu}^{(\text{CCSD})}$ is the vertical excitation energy of EOMCCSD, Eq. (3.23),

$$\alpha_{\mu} = \sum_{i < j < k, a < b < c} \ell_{\mu,ijk}^{abc} \tilde{\mathfrak{M}}_{\mu,abc}^{ijk}(2), \quad (3.83)$$

where $\tilde{\mathfrak{M}}_{\mu,abc}^{ijk}(2) = \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} (R_{\mu,1} + R_{\mu,2}) | \Phi \rangle$ is the contribution to the triply excited moment $\mathfrak{M}_{\mu,abc}^{ijk}(2)$ of EOMCCSD due to the one- and two-body components of the EOM-CCSD excitation operator $R_{\mu}^{(\text{CCSD})}$, Eq. (3.20), and β_{μ} is the quantity defined by Eq. (3.80). Since the EOMCCSD approach is rigorously size intensive and, as shown in Refs. [132, 200], the α_{μ} term is size intensive as well, the $[\omega_{\mu}^{(\text{CCSD})} + \alpha_{\mu}(2, 3)]$ part of the CR-EOMCC(2,3) excitation energy $\omega_{\mu}^{(\text{CR-EOMCC}(2,3))}$ is a size-intensive quantity. Unfortunately, the β_{μ} term defined by Eq. (3.80), being a size-extensive contribution that does not cancel out, grows with the size of the system [132, 200], destroying the size intensity of $\omega_{\mu}^{(\text{CR-EOMCC}(2,3))}$. In order to implement the rigorously size-intensive variant of CR-EOMCC(2,3), designated as δ -CR-EOMCC(2, 3) [201], the problematic β_{μ} term in Eq. (3.82) is simply neglected such that the excitation energy is redefined as follows:

$$\omega_{\mu}^{(\delta\text{-CR-EOMCC}(2,3))} = \omega_{\mu}^{(\text{CCSD})} + \alpha_{\mu}, \quad (3.84)$$

with α_μ given by Eq. (3.83). The resulting δ -CR-EOMCC(2,3) approach [201] provides a size-intensive description of the excitation energies and, by defining the total energy of a given electronic state μ , i.e., E_μ , as a sum of the size-extensive ground-state CR-CC(2,3) energy and size-intensive excitation energy $\omega_\mu^{(\delta\text{-CR-EOMCC}(2,3))}$, Eq. (3.84), so that

$$\begin{aligned}
E_\mu &= E_0^{(\text{CR-CC}(2,3))} + \omega_\mu^{(\delta\text{-CR-EOMCC}(2,3))} \\
&= E_\mu^{(\text{CCSD})} + \sum_{i < j < k, a < b < c} \ell_{0,ijk}^{abc} \mathfrak{M}_{0,abc}^{ijk}(2) \\
&\quad + \sum_{i < j < k, a < b < c} \ell_{\mu,ijk}^{abc} \tilde{\mathfrak{M}}_{\mu,abc}^{ijk}(2).
\end{aligned} \tag{3.85}$$

While the addition of noniterative corrections is one way to correct for higher-order excitations not included in lower-order CC/EOMCC approximations, in the next section yet another inexpensive approach to account for higher-order effects in CC/EOMCC calculations is considered.

3.1.4 The Active-Space Coupled-Cluster and Equation-of-Motion Coupled-Cluster Approaches

In Sect. (3.1.3) methods for obtaining a partial account of the T_3 and T_4 clusters, and their excited-state analogs, all based on noniterative MMCC corrections were considered. Another practical way to account for higher-than-doubly excited clusters in the SRCC considerations is by exploiting the ideas originally presented in Refs. [54, 133–136, 144], where the CCSDt and CCSDtq active-space CC equations were explored (see, also, Refs. [137–140] and Ref. [101] for a review). The basic language of the active-space CC and EOMCC methods follows from MRCC theory, where one subpartitions the one-electron basis of occupied

and unoccupied spin-orbitals used in the conventional SRCC-style considerations into (i) core or inactive occupied spin-orbitals, designated as $\mathbf{i}, \mathbf{j}, \dots$, (ii) active occupied spin-orbitals, designated as $\mathbf{I}, \mathbf{J}, \dots$, (iii) active unoccupied spin-orbitals, designated as $\mathbf{A}, \mathbf{B}, \dots$, and (iv) virtual or inactive unoccupied spin-orbitals, designated as $\mathbf{a}, \mathbf{b}, \dots$. The active-space methods CCSDt and CCSDtq, for example, then restrict the higher-order cluster components, T_3 or T_3 and T_4 , respectively, which must be calculated, to a small subset of all triples and quadruples defined via active orbitals. This restriction greatly reduces the computer costs, when compared to parent CCSDT and CCSDTQ approaches, as discussed in the Introduction. The CPU-time determining steps of CCSDt and CCSDtq are $N_o N_u n_o^2 n_u^4$ and $N_o^2 N_u^2 n_o^2 n_u^4$, respectively, where N_o and N_u are the numbers of active occupied and active unoccupied orbitals, i.e., the costs of CCSDt and CCSDtq calculations scale as relatively small prefactors times the costs of the corresponding CCSD calculations.

As an example of how the equations for the active-space methods are obtained, let us take the CCSDtq method as an example. The conventional CCSDTQ equations, which must be solved for the t_a^i , t_{ab}^{ij} , t_{abc}^{ijk} , and t_{abcd}^{ijkl} amplitudes by projecting the connected cluster form of the Schrödinger equation, Eq. (3.5), in which $T^{(A)} = T_1 + T_2 + T_3 + T_4$, on the singly, doubly, triply, and quadruply excited determinants, can be written as follows:

$$\langle \Phi_i^a | \mathbf{CCSD} + (H_N T_3)_C | \Phi \rangle = 0, \quad (3.86)$$

$$\langle \Phi_{ij}^{ab} | \mathbf{CCSD} + [H_N (T_3 + T_1 T_3 + T_4)]_C | \Phi \rangle = 0, \quad (3.87)$$

$$\langle \Phi_{ijk}^{abc} | \mathbf{CCSD} + [H_N (T_3 + T_1 T_3 + T_4 + T_1 T_4 + T_2 T_3 + \frac{1}{2} T_1^2 T_3)]_C | \Phi \rangle = 0, \quad (3.88)$$

$$\begin{aligned} \langle \Phi_{ijkl}^{abcd} | \mathbf{CCSD} + [H_N(T_3 + T_1 T_3 + T_4 + T_1 T_4 + T_2 T_3 + \frac{1}{2} T_1^2 T_3 \\ + T_2 T_4 + \frac{1}{2} T_3^2 + \frac{1}{2} T_1^2 T_4 + T_1 T_2 T_3 + \frac{1}{6} T_1^3 T_3)]_C | \Phi \rangle = 0 \quad , \end{aligned} \quad (3.89)$$

where $\mathbf{CCSD}$ designates all terms that contain T_1 and T_2 clusters only and terms that do not contain cluster operators at all. Once the system of equations, Eqs. (3.86) – (3.89) is solved, the energy is obtained using Eq. (3.14). The CCSDT method could be obtained from the CCSDTQ approach described here by simply setting $T_4 = 0$ and solving the system of equations formed by Eqs. (3.86) – (3.88) only.

In order to introduce the active-space CCSDtq formalism, the T_3 and T_4 clusters must be restricted to the internal and semi-internal excitations of the following types:

$$t_3 = \sum_{\mathbf{I} > \mathbf{J} > \mathbf{C}, a > b > \mathbf{C}} t_{ab\mathbf{C}}^{\mathbf{I}jk} a^a a^b a^{\mathbf{C}} a^k a^j a^{\mathbf{I}}, \quad (3.90)$$

$$t_4 = \sum_{\mathbf{I} > \mathbf{J} > \mathbf{C} > \mathbf{D}, a > b > \mathbf{C} > \mathbf{D}} t_{ab\mathbf{CD}}^{\mathbf{I}Jkl} a^a a^b a^{\mathbf{C}} a^l a^k a^{\mathbf{J}} a^{\mathbf{I}}, \quad (3.91)$$

where $\mathbf{I}$ and $\mathbf{J}$ are summed over only active occupied orbitals and $\mathbf{C}$ and $\mathbf{D}$ are summed over only active unoccupied orbitals. The CCSDtq system of equations for the relevant t_a^i , t_{ab}^{ij} , $t_{ab\mathbf{C}}^{\mathbf{I}jk}$, and $t_{ab\mathbf{CD}}^{\mathbf{I}Jkl}$ amplitudes has the form of CCSDTQ equations in which T_3 and T_4 are replaced by t_3 and t_4 , respectively. We obtain

$$\langle \Phi_i^a | \mathbf{CCSD} + (H_N t_3)_C | \Phi \rangle = 0, \quad (3.92)$$

$$\langle \Phi_{ij}^{ab} | \mathbf{CCSD} + [H_N(t_3 + T_1 t_3 + t_4)]_C | \Phi \rangle = 0, \quad (3.93)$$

$$\langle \Phi_{ijk}^{ab\mathbf{C}} | \mathbf{CCSD} + [H_N(t_3 + T_1 t_3 + t_4 + T_1 t_4 + T_2 t_3 + \frac{1}{2} T_1^2 t_3)]_C | \Phi \rangle = 0, \quad (3.94)$$

$$\begin{aligned}
\langle \Phi_{\mathbf{I}\mathbf{J}kl}^{ab\mathbf{CD}} | & \text{CCSD} + [H_N(t_3 + T_1 t_3 + t_4 + T_1 t_4 + T_2 t_3 + \frac{1}{2} T_1^2 t_3 \\
& + T_2 t_4 + \frac{1}{2} t_3^2 + \frac{1}{2} T_1^2 t_4 + T_1 T_2 t_3 + \frac{1}{6} T_1^3 t_3)]_C |\Phi\rangle = 0 \quad , \quad (3.95)
\end{aligned}$$

where t_3 and t_4 are defined by Eqs. (3.90) and (3.91), respectively. The CCSDt method could be obtained from the CCSDtq equations described here by setting $t_4 = 0$ and solving the system of equations given by Eqs. (3.92) – (3.94) only.

The active-space methods of the CCSDt and CCSDtq types were shown to be effective for excited-state theories as well, for example, in Refs. [55–57], where the EOMCCSDt approach, an active-space variant of EOMCCSDT was reported for the first time. While it is rather straightforward to generalize the active-space methods to particle conserving EOMCC theories, such as EOMCCSDT and EOMCCSDTQ, in this dissertation the focus is on the active-space variants of particle non-conserving EOMCC theories, in particular, the EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) methods. The full inclusion of the $R_{\mu,3p-2h}$ and $R_{\mu,3h-2p}$ components of the $R_{\mu}^{(N+1)}$ and $R_{\mu}^{(N-1)}$ operators in the EA- and IP-EOMCC calculations needed to obtain an accurate description of electronic excitations in radicals comes at a high price, increasing the $\mathcal{N}^5$ -like $n_o n_u^4$ and $n_o^2 n_u^3$ operations defining the iterative diagonalization steps of the base EA-EOMCCSD($2p-1h$) and IP-EOMCCSD($2h-1p$) schemes to the $\mathcal{N}^7$ -like $n_o^2 n_u^5$ and $n_o^3 n_u^4$ steps, respectively. One way to retain the accuracy of the higher-order EA- and IP-EOMCC schemes with the $3p-2h$ and $3h-2p$ excitations, while avoiding this steep computer cost increase, is to use the active-space variants of the EA/IP-EOMCC methods with higher-than $2p-1h/2h-1p$ excitations described in Refs. [147–149].

In analogy to the ground-state active-space CC approaches described above, in the active-space EA- and IP-EOMCC methods one divides the available orbitals of the N -electron

reference system into core, active occupied, active unoccupied, and virtual categories, and uses active orbitals to define the electron attaching and ionizing operators $R_\mu^{(N+1)}$ and $R_\mu^{(N-1)}$, respectively. In particular, the active-space EA-EOMCCSD($3p-2h$){ N_u } approach using N_u active unoccupied orbitals is obtained by replacing the $3p-2h$ component $R_{\mu,3p-2h}$ of the electron attaching operator $R_\mu^{(N+1)}$, Eq. (3.28), by

$$r_{\mu,3p-2h} = \sum_{j>k, \mathbf{A}<b<c} r_{\mathbf{A}bc}^{jk} a^{\mathbf{A}b} a^c a_k a_j. \quad (3.96)$$

The relatively small set of the unknown amplitudes $r_{\mathbf{A}bc}^{jk}$ defining $r_{\mu,3p-2h}$, Eq. (3.96), in which at least one of the three unoccupied spin-orbital indices is active, and the remaining $1p$ and $2p-1h$ amplitudes r_a and r_{ab}^j that enter the $(N+1)$ -electron wave functions of the active-space EA-EOMCCSD($3p-2h$){ N_u } approach are obtained by diagonalizing the similarity-transformed Hamiltonian of CCSD, Eq. (3.9), obtained in the ground-state CCSD calculations for the N -electron reference system, in the subspace of the $(N+1)$ -electron Hilbert space spanned by the $|\Phi^a\rangle$, $|\Phi_j^{ab}\rangle$, and $|\Phi_{jk}^{\mathbf{A}bc}\rangle$ determinants. Similarly, the active-space IP-EOMCCSD($3h-2p$){ N_o } approach using N_o active occupied orbitals is obtained by replacing the $3h-2p$ component $R_{\mu,3h-2p}$ of the ionizing operator $R_\mu^{(N-1)}$, Eq. (3.29), by

$$r_{\mu,3h-2p} = \sum_{\mathbf{I}>j>k, b<c} r_{bc}^{\mathbf{I}jk} a^b a^c a_k a_j a_{\mathbf{I}}, \quad (3.97)$$

where the relatively small set of the unknown amplitudes $r_{bc}^{\mathbf{I}jk}$ defining $r_{\mu,3h-2p}$, Eq. (3.97), in which at least one of the three occupied spin-orbital indices is active, and the remaining $1h$ and $2h-1p$ amplitudes r^i and r_b^{ij} that define the $(N-1)$ -electron wave functions of the active-space IP-EOMCCSD($3h-2p$){ N_o } approach are obtained by diagonalizing the similarity-

transformed Hamiltonian obtained in the N -electron CCSD calculations, Eq. (3.9), in the subspace of the $(N - 1)$ -electron Hilbert space spanned by the $|\Phi_i\rangle$, $|\Phi_{ij}^b\rangle$, and $|\Phi_{\mathbf{I}jk}^{bc}\rangle$ determinants. In Sect. (3.2.5) the full and active-space variants of the EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) methods are used to optimize the geometries of the ground and low-lying excited states of four open-shell molecules, CNC, C₂N, NCO and N₃, and determine the corresponding adiabatic excitation energies. The results provided in that section will demonstrate typical accuracies of the active-space EA- and IP-EOMCC methods as compared with their parent approaches.

3.2 Applications

In this section, typical accuracies of the methods described in Sect. (3.1) are demonstrated using various chemically relevant benchmarks and applications. For all of the studied systems, the relevant benchmark data are provided for comparison, originating either from experimentally measured quantities or from quantum-chemical calculations performed at or near the full CI level. In Sect. (3.2.1), the CR-CC(2,3) method will be used to calculate barrier heights for a large variety of simple chemical reactions, for which the activation barriers are well established and which are characterized by largely SR transition states, to see if it can match the performance of CCSD(T) in cases where CCSD(T) is accurate. A more MR case is presented in Sect. (3.2.2) and then an extremely biradical case is presented in Sect. (3.2.3) to demonstrate the relative performance of the conventional CCSD(T) and CCSD(TQ) methods, as compared with the performance of the CR-CC(2,3) and CR-CC(2,3)+Q theories including corrections for the connected triply and quadruply excited clusters, which are specifically designed to handle such situations. In Sects. (3.2.4)

and (3.2.5), excited-state systems are considered in order to demonstrate accuracies of the CR-EOMCC(2,3) method for calculating vertical excitations and the δ -CR-EOMCC(2,3) approach for calculating adiabatic excitation energies. Finally, in Sect. (3.2.6) the advantages and accuracies of various levels of EA- and IP-EOMCC methods are discussed in more challenging cases of many-electron excitations in open-shell systems. Although the results presented in this section are chosen to demonstrate the strengths of the CC and EOMCC methods, with particular attention paid to our CR-CC, CR-EOMCC, and active-space EA- and IP-EOMCC approaches, the weaknesses of each method will be pointed out as well.

3.2.1 The DBH24 Benchmark Database for Thermochemical Kinetics

This section provides a systematic comparison of the performance of CCSD, CCSD(T), and variants A-D of the CR-CC(2,3) theory for a diverse collection of reaction barrier heights. Since it is rather inconvenient to test such methods on the typical large benchmark databases such as NHTBH38/04 [202] or Database/3 [203], as calculating the necessary chemical species can quickly become excessively time consuming when many methods and basis sets are examined, a representative benchmark database, DBH24 [204], was developed as a more feasible alternative to these large databases, designed specifically with computationally more intensive *ab initio* methods in mind. DBH24 is composed of 24 barrier heights which were determined to be the most statistically representative subset of all 38 of the forward and reverse barrier heights of NHTBH38/04 and the 44 hydrogen-transfer barrier heights of Database/3. DBH24 consists of four types of reactions, namely, hydrogen transfer (HT), heavy-atom transfer (HAT), nucleophilic substitution (NS), and unimolecular and associ-

ation (UA) reactions. There are three reactions (six barrier heights because forward and reverse reactions are considered) for each type of reaction in the DBH24 database. The six barrier heights of each reaction are denoted as HATBH6, NSBH6, UABH6, and HTBH6, respectively.

When the DBH24 benchmark database was introduced in Ref. [204], the performance of many *ab initio* wave function and density functional theory methods was examined alongside some semi-empirical and composite approaches with the goal of determining the best approaches (measured in terms of accuracy, consistency, and computational efficiency) for the calculation of barrier heights for routine use in thermochemical kinetics. The benchmark values in DBH24 were obtained using high-level theoretical methods, such as Weizmann-1 (W1) [205] or MRCI calculations, or, in a few cases, benchmark values were derived from experimental data. The full list of reactions and forward and reverse barrier height benchmark values are given in Table (3.1). All calculations reported in Ref. [204] were based on reactant, product, and transition structures optimized at the QCISD/MG3 level with the spin-restricted formalism for closed-shell systems and the fully spin-unrestricted formalism for open-shell systems. The effect of spin-orbit coupling was added to the energies of the Cl and OH radicals, which lower their energies by 0.84 and 0.20 kcal/mol, respectively. Among all of the single-level (i.e. not composite) wave function methods tested in Ref. [204], those based on CC theory, especially CCSD(T), proved to be the most accurate.

Following this initial study, we extended the work reported in Ref. [204] in Ref. [206] to see if we could objectively determine whether the CR-CC(2,3) method could statistically outperform the CCSD(T) approach on the DBH24 benchmark set due to the improved form of the triples correction in CR-CC(2,3). We also used the opportunity to test the depen-

Table 3.1: Representative barrier heights database DBH24 taken from Ref. [204].

database	reaction	$V_f^{\neq a}$	$V_r^{\neq a}$
HATBH6	$\text{H} + \text{N}_2\text{O} \rightarrow \text{OH} + \text{N}_2$	18.14	83.22
	$\text{H} + \text{ClH} \rightarrow \text{HCl} + \text{H}$	18.00	18.00
	$\text{CH}_3 + \text{FCl} \rightarrow \text{CH}_3\text{F} + \text{Cl}$	7.43	61.01
NSBH6	$\text{Cl}^- \cdots \text{CH}_3\text{Cl} \rightarrow \text{ClCH}_3 \cdots \text{Cl}^-$	13.61	13.61
	$\text{F}^- \cdots \text{CH}_3\text{Cl} \rightarrow \text{FCH}_3 \cdots \text{Cl}^-$	2.89	29.62
	$\text{OH}^- + \text{CH}_3\text{F} \rightarrow \text{HOCH}_3 + \text{F}^-$	-2.78	17.33
UABH6	$\text{H} + \text{N}_2 \rightarrow \text{HN}_2$	14.69	10.72
	$\text{H} + \text{C}_2\text{H}_4 \rightarrow \text{CH}_3\text{CH}_2$	1.72	41.75
	$\text{HCN} \rightarrow \text{HNC}$	48.16	33.11
HTBH6	$\text{OH} + \text{CH}_4 \rightarrow \text{CH}_3 + \text{H}_2\text{O}$	6.7	19.6
	$\text{H} + \text{OH} \rightarrow \text{O} + \text{H}_2$	10.7	13.1
	$\text{H} + \text{H}_2\text{S} \rightarrow \text{H}_2 + \text{HS}$	3.6	17.3

^a $V_f^{\neq}$ denotes forward barrier height and $V_r^{\neq}$ denotes reverse barrier height (in kcal/mol).

dence of the results on the quality of the basis set and the effect of freezing core orbitals. Calculations were performed using the CCSD, CCSD(T), and CR-CC(2,3) approaches to calculate the forward and reverse barrier heights for all of the reactions included in the DBH24 database using five different basis sets of triple-zeta quality with and without applying the frozen core approximation. The five basis sets used in our study [206] were MG3S [207], and four correlation consistent basis sets, namely, aug-cc-pVTZ [208–210], aug-cc-pV(T+d) [211], aug-cc-pCVTZ [208, 209, 212, 213], and aug-cc-pCV(T+d)Z. Note that MG3S is identical to 6-311+G(3d2f,2df,2p) for H-Si and is similar to 6-311+G(3d2f), but improved [214] for P-Ar. The aug-cc-pV(T+d)Z basis set is the same as aug-cc-pVTZ except that it has a single extra d function for the second row atoms from Al through Ar, and the other d functions of aug-cc-pVTZ are also optimized for these atoms. The aug-cc-pCV(T+d)Z basis set is same as

aug-cc-pCVTZ basis set except that all valence d functions are taken from aug-cc-pV(T+d)Z plus two d functions describing inner shells are taken from aug-cc-pCVTZ. As in the original study by Zheng et al. [204], the geometries used in our calculations [206] were optimized using the QCISD/MG3 level and the energies of the Cl and OH radicals were corrected for spin-orbit effects.

The entire set of reaction barrier heights for the DBH24 database, as calculated with the CCSD, CCSD(T), and CR-CC(2,3),A-D approaches combined with the five triple-zeta basis sets mentioned above is supplied in the Supporting Information to Ref. [206]. The calculated mean signed errors (MSEs) and mean unsigned errors (MUEs) obtained from these calculations, taken from Ref. [206], are reported in Tables (3.2) and (3.3), with Table (3.2) collecting results from all-electron calculations with all orbitals are correlated and Table (3.3) collecting results from calculations with the core orbitals frozen. It can be seen from these two tables that the CCSD results are relatively poor, typically producing MUE values around or above 2.0 kcal/mol, but the CCSD(T) and CR-CC(2,3),A-D methods significantly improve the CCSD activation energies, especially in conjunction with the augmented correlation consistent basis sets. The mean unsigned errors of CR-CC(2,3) and CCSD(T) with the MG3S basis set are about 0.9-1.0 kcal/mol both when correlating all electrons and when correlating only valence electrons, whereas the mean unsigned errors characterizing the CR-CC(2,3) and CCSD(T) results for the augmented correlation consistent basis sets vary between 0.4 and 0.75 kcal/mol. We conclude from this that the MG3S basis set offers an economical triple-zeta basis set alternative for systems which are too large for the correlation consistent basis sets to be affordable.

Table 3.2: Mean signed error (MSE) and mean unsigned error (MUE) of coupled cluster methods calculated with all electrons correlated compared against the DBH24 benchmark database (in kcal/mol).

Method	HATBH6		NSBH6		UABH6		HTBH6		DBH24
	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MUE
MG3S									
CCSD(full)	4.36	4.36	2.29	2.29	1.76	1.76	2.55	2.55	2.74
CCSD(T)(full)	0.92	1.24	-0.01	0.74	0.70	0.70	0.93	1.04	0.93
CR-CC(2,3),A(full)	1.50	1.61	0.72	0.54	0.93	0.93	1.13	1.14	1.06
CR-CC(2,3),B(full)	1.72	1.77	0.47	0.57	0.96	0.96	1.19	1.19	1.12
CR-CC(2,3),C(full)	1.16	1.35	0.10	0.59	0.82	0.82	0.98	1.04	0.95
CR-CC(2,3),D(full)	1.17	1.35	0.10	0.60	0.82	0.82	0.98	1.04	0.95
aug-cc-pVTZ									
CCSD(full)	2.85	2.85	1.83	1.83	1.28	1.28	1.06	1.06	1.76
CCSD(T)(full)	-0.72	0.84	-0.52	0.64	0.17	0.34	-0.72	0.72	0.64
aug-cc-pCVTZ									
CCSD(full)	3.61	3.61	2.12	2.12	1.13	1.13	1.75	1.75	2.15
CCSD(T)(full)	-0.03	0.61	-0.26	0.46	0.01	0.28	-0.05	0.45	0.45
CR-CC(2,3),A(full)	0.55	0.84	0.04	0.34	0.23	0.38	0.14	0.54	0.52
CR-CC(2,3),B(full)	0.77	0.97	0.23	0.34	0.26	0.41	0.22	0.56	0.57
CR-CC(2,3),C(full)	0.42	0.76	-0.22	0.63	0.22	0.43	0.11	0.55	0.59
CR-CC(2,3),D(full)	0.42	0.76	-0.22	0.63	0.22	0.43	0.11	0.55	0.59
aug-cc-pCV(T+d)Z									
CCSD(full)	3.60	3.60	2.15	2.15	1.13	1.13	1.76	1.76	2.16
CCSD(T)(full)	-0.05	0.58	-0.24	0.44	0.01	0.28	-0.05	0.45	0.44

Table 3.3: Mean signed error (MSE) and mean unsigned error (MUE) of coupled cluster methods calculated with frozen core approximation compared against the DBH24 benchmark database (in kcal/mol).

Method	HATBH6		NSBH6		UABH6		HTBH6		DBH24
	MSE	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MUE
MG3S									
CCSD	4.43	4.43	2.03	2.03	1.58	1.58	2.62	2.62	2.67
CCSD(T)	1.06	1.37	-0.25	0.94	0.53	0.53	1.04	1.10	0.98
CR-CC(2,3),A	1.63	1.76	0.03	0.75	0.76	0.76	1.23	1.23	1.12
CR-CC(2,3),B	1.85	1.91	0.22	0.63	0.80	0.80	1.29	1.29	1.16
CR-CC(2,3),C	1.28	1.49	-0.17	0.83	0.66	0.66	1.08	1.10	1.02
CR-CC(2,3),D	1.29	1.49	-0.18	0.83	0.65	0.65	1.08	1.10	1.02
aug-cc-pVTZ									
CCSD	3.54	3.54	1.66	1.66	1.03	1.11	1.72	1.72	2.01
CCSD(T)	0.01	0.91	-0.67	0.68	-0.06	0.40	-0.04	0.57	0.64
CR-CC(2,3),A	0.58	1.20	-0.39	0.44	0.17	0.47	0.15	0.62	0.68
CR-CC(2,3),B	0.80	1.32	-0.20	0.35	0.19	0.49	0.23	0.65	0.70
CR-CC(2,3),C	0.46	1.13	-0.70	0.77	0.16	0.48	0.11	0.64	0.75
CR-CC(2,3),D	0.46	1.13	-0.71	0.77	0.15	0.48	0.11	0.64	0.75
aug-cc-pV(T+d)Z									
CCSD	3.41	3.41	1.82	1.82	1.03	1.11	1.69	1.69	2.01
CCSD(T)	-0.13	0.67	-0.53	0.62	-0.06	0.40	-0.06	0.54	0.56
CR-CC(2,3),A	0.45	0.88	-0.24	0.39	0.17	0.47	0.13	0.60	0.58
CR-CC(2,3),B	0.67	1.00	-0.05	0.30	0.19	0.49	0.20	0.63	0.61
CR-CC(2,3),C	0.30	0.80	-0.53	0.60	0.16	0.48	0.09	0.62	0.62
CR-CC(2,3),D	0.31	0.80	-0.54	0.60	0.15	0.48	0.09	0.62	0.62

Although the CCSD(T)(full)/aug-cc-pCV(T+d)Z method gives the best results among all the tested methods, with a mean unsigned error of only 0.44 kcal/mol, it is our opinion that this is not the best combination of method and basis set for routine applications of the type of the benchmark study examined here. There are several reasons in support of this conclusion. First, it is clear by a comparison of Tables (3.2) and (3.3) that all-electron CCSD(T) and CR-CC(2,3) calculations generally give only slightly better results than those produced by frozen-core calculations using the MG3S, aug-cc-pVTZ, or aug-cc-pV(T+d)Z basis sets. As an example, the CR-CC(2,3),D/MG3S calculations with all electrons correlated produced MUEs ranging from 0.60 to 1.35 kcal/mol, while the same combination of method and basis set produced MUEs of 0.65 to 1.49 kcal/mol under the frozen-core approximation. In general, it is not recommended to use valence-optimized basis sets when including both core and core-valence correlations, since this is not only more expensive, but also a potential source of problems [215]. By examining MUEs in Table (3.2) corresponding to CCSD(T) calculations using the aug-cc-pVTZ and aug-cc-pCVTZ basis sets, which have ranges from 0.34 to 0.84 kcal/mol and 0.28 to 0.61 kcal/mol, respectively, it is clear that the accuracy systematically increases when the core-optimized basis sets are used in all-electron calculations. However, since the results only improve slightly it is our opinion that employing the core-optimized basis sets is not practical due to the increase in the number of basis functions composing these basis sets and the additional correlated orbitals required to perform the corresponding all-electron calculations, which, taken together, make the calculations significantly more expensive than their frozen-core analogs. It is also clear from Tables (3.2) and (3.3) that the overall accuracy of the CR-CC(2,3),A-D approaches is practically the same as that of CCSD(T). In particular, since CR-CC(2,3),D is shown to reproduce the high accuracy

of CCSD(T) in the barrier height calculations for the reactions from the DBH24 database, which are largely of the SR type, while also offering a significantly better performance in more MR cases, as demonstrated in Sects. (3.2.2) and (3.2.3), we recommend it for applications where an accurate treatment of triples is required, particularly when paired with the aug-cc-pV(T+d)Z basis set in the frozen-core approximation.

Additional unpublished work was performed addressing the question of whether the MSEs and MUEs characterizing the CR-CC(2,3) calculations could be further reduced by saturating the basis set. The increase in accuracy observed when switching from the aug-cc-pVTZ to aug-cc-pCVTZ basis set can be interpreted in one of two ways. Either correlating the core electrons has a significant effect on the accuracy of these calculations, or, since there is a large disparity between the number of basis functions in the two sets, a large portion of the errors reported in Ref. [206] were due to basis set incompleteness. To investigate this issue, CR-CC(2,3) energies were calculated with the aug-cc-pV(Q+d)Z basis set for all relevant species in the DBH24 benchmark database. This basis set was chosen because it follows aug-cc-pV(T+d)Z in the hierarchy of basis sets having the general form aug-cc-pV(X +d)Z, where X is known as the cardinal number of the basis set, and, having obtained energies at the aug-cc-pV(T+d)Z and aug-cc-pV(Q+d)Z basis set levels, it is then possible to extrapolate the electronic correlation energy to the complete basis set (CBS) limit, ΔE_∞ , using one of the existing empirical laws defining the dependence of the electronic correlation energy ΔE on X , such as

$$\Delta E(X) = \Delta E_\infty + AX^{-3}, \quad (3.98)$$

where $\Delta E(X)$ is the correlation energy obtained with the aug-cc-pV(X +d)Z basis set and ΔE_∞ and A are the parameters determined from fitting $\Delta E(X)$ to the calculated correla-

tion energies. Preference is always given to the aug-cc-pV(T+d)Z and aug-cc-pV(Q+d)Z basis set data in these fits, rather than the easier to obtain aug-cc-pV(D+d)Z data, because the aug-cc-pV(D+d)Z data are well known to produce rather poor fits to the function given by Eq. (3.98). While the CBS-limit correlation energy can be obtained by the above extrapolation technique, the reference energy, which was provided by the RHF method for the closed-shell DBH24 species and by the ROHF method for open-shell species, was explicitly calculated using a very large, nearly complete, basis set, i.e., the aug-cc-pV(6+d)Z basis set. This has been shown to be a more accurate approach than extrapolation for obtaining approximate CBS-limit reference energies [216] and the calculations do not become prohibitively expensive even for the largest molecules of interest in this study. Both the fast (exponential) convergence behavior of the reference energy and the unreliable nature of correlation energies resulting from correlation consistent DZ-type basis sets with cardinal number 2 are thoroughly discussed and well illustrated for the H₂O system in Ref. [217]. Thus, the approximate CBS-limit total energy reported in this section is constructed as a sum of the energy produced by an aug-cc-pV(6+d)Z reference calculation and a CBS-limit correlation energy, obtained by entering the aug-cc-pV(T+d)Z and aug-cc-pV(Q+d)Z basis set level correlation energies into the extrapolation formula given by Eq. (3.98).

Table 3.4: Errors resulting from CR-CC(2,3) calculations of forward and reverse barrier heights, $V_f^\ddagger$ and $V_r^\ddagger$ respectively, reported as $\epsilon(V_f^\ddagger) / \epsilon(V_r^\ddagger)$ at varying basis set levels compared against DBH24 benchmark values (in kcal/mol).

	HATBH6		
	$\text{H} + \text{N}_2\text{O} \rightarrow \text{OH} + \text{N}_2$	$\text{H} + \text{HCl} \rightarrow \text{HCl} + \text{H}$	$\text{CH}_3 + \text{FCl} \rightarrow \text{CH}_3\text{F} + \text{Cl}$
aug-cc-pV(D+d)Z	-1.17/3.84	1.92/1.92	-2.96/-0.98
aug-cc-pV(T+d)Z	-0.99/2.16	0.24/0.24	-0.48/0.69
aug-cc-pV(Q+d)Z	-0.37/1.85	-0.10/-0.10	0.12/0.66
CBS-limit	0.07/1.69	-0.41/-0.41	0.48/0.85
	NSBH6		
	$\text{Cl}^- \dots \text{CH}_3\text{Cl} \rightarrow \text{ClCH}_3 \dots \text{Cl}^-$	$\text{F}^- \dots \text{CH}_3\text{Cl} \rightarrow \text{FCH}_3 \dots \text{Cl}^-$	$\text{OH}^- + \text{CH}_3\text{F} \rightarrow \text{HOCH}_3 + \text{F}^-$
aug-cc-pV(D+d)Z	-1.31/-1.31	-0.79/-2.16	-2.81/-1.58
aug-cc-pV(T+d)Z	-1.08/-1.08	-0.14/-1.08	0.19/-0.04
aug-cc-pV(Q+d)Z	-0.93/-0.93	-0.01/-0.34	0.52/0.59
CBS-limit	-0.93/-0.93	0.05/0.12	0.73/0.99
	UABH6		
	$\text{H} + \text{N}_2 \rightarrow \text{HN}_2$	$\text{H} + \text{C}_2\text{H}_4 \rightarrow \text{C}_2\text{H}_5$	$\text{HCN} \rightarrow \text{HNC}$
aug-cc-pV(D+d)Z	0.13/ 0.32	0.63/ 0.19	-1.92/-1.24
aug-cc-pV(T+d)Z	0.00/ 0.56	0.34/ 1.00	-0.67/-0.30
aug-cc-pV(Q+d)Z	0.08/ 0.53	0.38/ 0.86	-0.50/-0.08
CBS-limit	0.08/ 0.58	0.39/ 0.79	-0.36/ 0.04
	HTBH6		
	$\text{OH} + \text{CH}_4 \rightarrow \text{CH}_3 + \text{H}_2\text{O}$	$\text{H} + \text{OH} \rightarrow \text{O} + \text{H}_2$	$\text{H} + \text{H}_2\text{S} \rightarrow \text{H}_2 + \text{HS}$
aug-cc-pV(D+d)Z	0.15/-1.28	-0.88/ 0.93	0.44/ 0.00
aug-cc-pV(T+d)Z	0.20/-0.83	-0.76/ 0.67	0.28/ 0.95
aug-cc-pV(Q+d)Z	0.19/-0.09	-0.24/ 0.31	0.43/ 0.74
CBS-limit	0.16/ 0.31	0.05/ 0.02	0.52/ 0.56

The individual errors resulting from CR-CC(2,3) calculations for the various DBH24 reactions, when used in conjunction with the aug-cc-pV($X+Z$)Z basis sets, the corresponding CBS-limit values, and the MSE and MUE values characterizing these calculations are collected in Tables (3.4) and (3.5). Results produced by CR-CC(2,3)/aug-cc-pV(D+d)Z calculations are only included to emphasize the convergence with the basis set. For every reaction considered, improvements in the quality of basis set correlate with improvements in accuracy. For many of these reactions it is clear that the basis set truncation was the main source of the triple-zeta-level error in the calculations reported in Ref. [206]. As shown in Table (3.5), the MSE reduces for both the forward and reverse barrier heights to well beyond chemical accuracy as the CBS-limit of the CR-CC(2,3) energies is approached. The MUEs seem to be converged at the aug-cc-pV(Q+d)Z basis set level since there is almost no difference with the CBS-limit results. Until it is explicitly proven, it can be assumed that additional contributions from the remaining triples, quadruples, and other higher-order connected clusters contribute minimally to the barrier heights considered here and may be disregarded when predicting the activation energies for similar reactions to within a fraction of 1 kcal/mol. In Sects. (3.2.2) and (3.2.3) chemical systems are examined which require at least a partial treatment of quadruple excitations in order to attain results within chemical accuracy. Our interest is in determining if the CR-CC methods, such as CR-CC(2,3)+Q can be used to accurately describe such situations when they arise.

3.2.2 Addition Reactions of Ethylene and Acetylene to Ozone

In the previous section it was shown that the CR-CC(2,3) method is capable of matching the performance of CCSD(T), where both were able to achieve chemical accuracy in con-

Table 3.5: Mean signed errors (MSE) and mean unsigned errors (MUE) resulting from CR-CC(2,3) calculations of all DBH24 forward and reverse barrier heights, $V_f^\ddagger$ and $V_r^\ddagger$ respectively, reported as $\epsilon(V_f^\ddagger) / \epsilon(V_r^\ddagger)$ at varying basis set levels (in kcal/mol).

	MSE	MUE
aug-cc-pV(D+d)Z	-0.28 / 0.96	1.11 / 1.38
aug-cc-pV(T+d)Z	-0.35 / 0.71	0.52 / 0.81
aug-cc-pV(Q+d)Z	-0.15 / 0.54	0.32 / 0.59
CBS-limit	0.07 / 0.38	0.35 / 0.60

junction with medium basis sets when predicting simple thermochemical barrier heights of the predominantly SR nature, and that the CR-CC(2,3) method could even achieve a sub-chemical level of accuracy in the CBS-limit. It was thus clear that for the species included in the DBH24 benchmark database, the correlation effects included in the CCSD(T) and CR-CC(2,3) methods were sufficient to reproduce benchmark data with excellent accuracy. However, not all chemical systems can be described so accurately at the CCSD(T) and CR-CC(2,3) levels. Some systems are more MR and as such, require a balanced description of triply and quadruply excited clusters.

Ozone is one notorious example of a MR system [218–221]. Despite its well known closed-shell singlet electronic structure, ozone exhibits a significant biradical character estimated to be around 33% [222–225]. Ozone is a common reagent in organic chemistry for the generation of ketones, aldehydes, epoxides, peroxides, anhydrides, and polymers via ozonolysis of alkenes and alkynes [226–231]. The mechanism of such processes is generally accepted to proceed by initial formation of a van der Waals (vdW) complex followed by a concerted cycloaddition transition state (TS) before finally reaching the cycloadduct configuration. In this section, cycloadditions of ethylene and acetylene, shown in Fig. (3.1) to the 1,3 termini of ozone

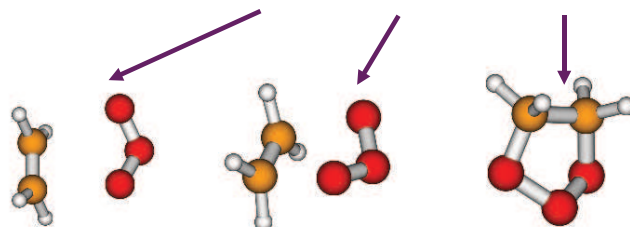
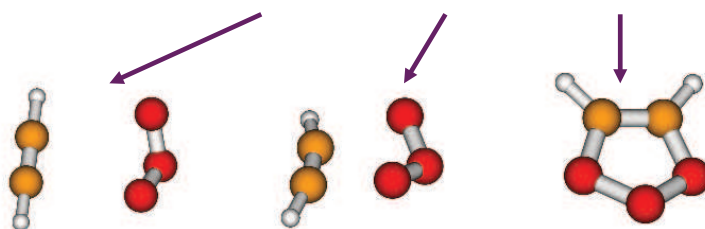


Figure 3.1: Stationary points along the C_2H_2 (top row) and C_2H_4 (bottom row) ozonolysis reaction pathways. In each row, the structures from left to right represent the van der Waals minimum (vdW), transition state (TS), and the cycloadduct, respectively. The oxygen, carbon, and hydrogen atoms are represented by the red, yellow, and grey spheres, respectively. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

Table 3.6: Benchmark values for the C₂H₂ and C₂H₄ ozonolysis reaction pathways.

method	O ₃ +C ₂ H ₂			O ₃ +C ₂ H ₄		
	vdW	TS cycloadduct		vdW	TS cycloadduct	
Wheeler et al. ^a	-1.85	7.74	-63.04	-1.84	3.43	-56.43
CBS CCSDT+(2) _Q ^b	-1.88	7.90	-63.91	-1.94	3.50	-57.15
CBS CCSD(T)+Q ^c	-1.98	7.58	-64.46	-2.03	3.18	-57.86
average = best estimate	-1.90	7.74	-63.80	-1.94	3.37	-57.15

^aFrom Ref. [225], calculated as CCSDT/CBS plus corrections for core correlation (C), adiabatic Born-Oppenheimer terms (A), relativistic effects (R) and a correction for quadruple excitations +(2)_Q estimated at the cc-pVDZ basis set level for C₂H₂ and assumed to be the same for C₂H₄. ^bCBS CCSDT/CBS+CAR from Ref. [225] plus quadruple excitation contributions calculated at the CCSDT(2)_Q/cc-pVDZ level for both C₂H₂ and C₂H₄ separately [242]. ^cCBS CCSD(T)/CBS+CAR from Ref. [225] plus quadruple excitation contributions calculated at the CCSD(TQ)/aug-cc-pVDZ level for both C₂H₂ and C₂H₄ separately [242].

are considered as a more challenging set of thermochemical barrier heights, which have reported literature values from ordinarily reliable theoretical methods ranging from 2 to 18 kcal/mol for ethylene [232–237] and 5 to 22 kcal/mol for acetylene. [238–241] When such large discrepancies are observed between various theoretical methods, it is typical that higher-order effects play a significant role in the description of the species involved and, indeed, work by Wheeler et al. [225] showed that there are non-negligible contributions to the energy from quadruply excited clusters in the ozonolysis of C₂H₂. Table (3.6) collects the best known computational benchmark values for stationary points along the C₂H₂ and C₂H₄ ozonolysis reaction pathways, as reported in Refs. [225] and [242]. Following the strategy of Ref. [242], the averages given in Table (3.6) are the values used for comparison in this discussion.

In a joint study published by the Piecuch and Truhlar groups in 2009 [242], the effect of quadruples on the C₂H₂ ozonolysis was reexamined, as treated with a larger basis set, and

the ozonolysis of C_2H_4 was considered for the first time with quadruple excitations explicitly included. This discussion focuses on the performance of CC methods that start with an iterative CCSD calculation and add noniterative corrections for connected higher-order excitations, including CCSD(T), CCSD(TQ), CR-CC(2,3), and CR-CC(2,3)+Q, although many other CC and non-CC results, which are not discussed here, were reported in Ref. [242] as well. Since calculating corrections to the correlation energy corresponding to quadruples at the aug-cc-pVTZ basis set level was computationally too demanding, a composite method was invented which treats most of the correlation energy with the aug-cc-pVTZ basis set but treats the part of the correlation energy corresponding to quadruples at the aug-cc-pVDZ level. For example, a CR-CC(2,3)/aug-cc-pVTZ calculation augmented by a quadruples correction at the aug-cc-pVDZ level is designated as CR-CC(2,3)/aug-cc-pVTZ+Q(aug-cc-pVDZ). All CR-CC(2,3) energies discussed in this section are the variant D values. The underlying geometries representing stationary points for the reaction pathways were optimized with the M05 density functional [243] and MG3S basis set.

In Table (3.7) the selected conventional CC and CR-CC results are compared, reported as relative energies with respect to the reactants, for each of the stationary points on the reaction pathways. A specific subset of these results was selected to be presented here which illustrates typical problems encountered at various levels of CC theory. We begin by considering the CCSD-level results found using the aug-cc-pVDZ and aug-cc-pVTZ basis sets, found in the first two rows of Table (3.7). It is clear that CCSD, when used with the aug-cc-pVDZ and aug-cc-pVTZ basis sets, produces a proper qualitative description of the two reaction pathways, however the overall accuracies, reported concisely as MUEs of all considered species with respect to the benchmark values in the second to last column, are

quite poor, with reported values of 2.4 and 3.3 kcal/mol, respectively.

The CCSD(T) and CCSD(TQ) results reported here, generated in conjunction with the aug-cc-pVDZ basis set, improve on the overall MUEs of CCSD, producing 1.7 and 1.5 kcal/mol, respectively, but both standard perturbative CC methods fail to give the correct qualitative description of the ethylene ozonolysis reaction pathway, placing the transition state -0.54 and -0.21 kcal/mol below the reactants, respectively. These results, predicting no barrier to reaction at all, are unphysical, disagreeing with high-level theory and experiment [244]. In mechanistic studies such as this, the relative energetics between reactants and the transition states are of particular interest, since the activation energy is often an experimentally derivable quantity, so in the final column of Table (3.7) the MUEs for only the transition states of both reactions are reported. Comparing CCSD(T) and CCSD(TQ) to the other reported methodologies by this measure clearly shows their exceptionally poor performance when used to describe barrier heights when either the reactants or transition state have significant biradical character and the other species do not.

The situation looks much better when the CR-CC methods are used to describe the same reaction pathways. By using the CR-CC(2,3) method rather than CCSD(T) in the same basis set, a lowering of the total MUE from 1.7 to 0.8 kcal/mol is observed, and an appropriate sign is obtained for the barrier in the ethylene ozonolysis. This is already considered sub-chemical accuracy, but not the best result yet. A logical next step toward increasing accuracy is to increase the basis set from aug-cc-pVDZ to aug-cc-pVTZ. It is shown in Table 3.7 that the same MUE of 0.8 kcal/mol is obtained when the aug-cc-pVTZ basis set is employed, but the activation energies move toward the benchmark values. Further improvements are observed when the CR-CC(2,3) results are corrected for quadruples. The incredibly small

total MUE produced by the CR-CC(2,3),D/aug-cc-pVTZ+Q(aug-cc-pVDZ) method (0.6 kcal/mol), and an ever smaller MUE for the activation energies (0.2 kcal/mol) show the ability of the relatively inexpensive CR-CC approaches to produce excellent energetics, even for difficult reaction pathways. In the next section, Sect. (3.2.3), an even more difficult reaction profile is considered.

Table 3.7: Energetics of stationary points relative to reactants, in kcal/mol, produced at various levels of coupled-cluster theory for the ozonolysis of ethylene and acetylene.

method	O ₃ +C ₂ H ₂			O ₃ +C ₂ H ₄			MUE	MUE
	vdW	TS cycloadduct		vdW	TS cycloadduct		All ^a	BHs ^b
CCSD/aug-cc-pVDZ	-2.00	7.47	-69.90	-2.38	2.69	-64.22	2.4	0.5
CCSD/aug-cc-pVTZ	-1.76	9.34	-71.27	-1.89	4.41	-66.36	3.3	1.3
CCSD(T)/aug-cc-pVDZ	-2.50	4.29	-64.21	-3.18	-0.54	-57.58	1.7	3.7
CCSD(TQ)/aug-cc-pVDZ	-2.48	4.73	-63.50	-3.17	-0.21	-56.89	1.5	3.3
CR-CC(2,3),D/aug-cc-pVDZ	-1.11	6.09	-63.46	-1.87	1.44	-57.10	0.8	1.8
CR-CC(2,3),D/aug-cc-pVTZ	-0.58	8.08	-64.06	-0.68	3.55	-58.31	0.8	0.3
CR-CC(2,3),D/aug-cc-pVTZ+Q(aug-cc-pVDZ)	-0.62	8.02	-63.29	-0.74	3.46	-57.52	0.6	0.2
Benchmark values ^c	-1.90	7.74	-63.80	-1.94	3.37	-57.15	—	—

^aAverage mean unsigned errors for all stationary points on the ethylene and acetylene ozonolysis reaction pathways. ^bAverage mean unsigned errors for the transition state barrier heights for ethylene and acetylene ozonolysis. ^cThe best estimates from Table (3.6) obtained in Ref. [242].

3.2.3 Mechanism of the Isomerization of Bicyclobutane to Butadiene

The pericyclic rearrangement of bicyclo[1.1.0]butane (abbreviated as **bicbut**) to *trans*-buta-1,3-diene (abbreviated as **t-but**) was chosen as another test case for the CR-CC methods, both for the intrinsic complexity of the associated isomerization pathways, which involve polyatomic structures with a rapidly varying degree of biradical character that require an accurate and balanced description of the dynamical and non-dynamical correlation effects, as well as for the favorable size of the system, which has several atoms but is still not too large, enabling calculations with larger basis sets to be performed and, in turn, allowing convergence behavior with the size of the basis set to be examined. Let us recall that early experimental studies suggest that the **bicbut**→**t-but** isomerization proceeds by concerted conrotatory movement of the methylene groups [245, 246], as predicted by the Woodward-Hoffman symmetry rules (see Figure (3.2)). Computational studies confirm this, predicting, in addition, that near the end the reaction pathway passes through the *gauche*-buta-1,3-diene configuration (abbreviated as **g-but**) before reaching the final **t-but** configuration [189, 248]. Theoretical studies have also considered the concerted disrotatory [189, 248, 249] and non-concerted [250] pathways, finding the concerted disrotatory TS to be ~ 20 kcal/mol higher in energy than the conrotatory TS and the non-concerted pathway to be much too high in energy to be even considered as a plausible mechanism. The conrotatory TS was found to have a ~ 24 % biradical character while the disrotatory TS was found to have a ~ 90 % biradical character, according to the high level electronic structure calculations reported in Ref. [189]. The current consensus on the mechanism for the isomerization of bicyclo[1.1.0]butane into buta-1,3-diene is that both concerted pathways begin at **bicbut** and, after passing through

the corresponding conrotatory or disrotatory TS (**con-TS** and **dis-TS**, respectively), they converge at the local minimum defining the intermediate **g-but** configuration (see Figure (3.2)). The **g-but** intermediate isomerizes via a low-energy rotational barrier, defined by the TS structure labeled as **gt-TS**, before the final product, **t-but**, is reached. The conrotatory and disrotatory concerted pathways describing the isomerization of bicyclo[1.1.0]butane into buta-1,3-diene, along with their available experimental activation [251] and reaction enthalpies (the latter based on the enthalpies of formation of bicyclo[1.1.0]butane and buta-1,3-diene reported in Ref. [252]; cf. Ref. [189]), and the theoretical enthalpy values obtained in this work, are illustrated in Figure (3.2).

As demonstrated in Ref. [189], the CCSD(T) approach completely fails by placing the disrotatory pathway defined by the strongly biradical **dis-TS** about 20 kcal/mol below the conrotatory pathway, contradicting experiment [245,246] and accurate MR calculations [248] which state that the conrotatory pathway represents a true mechanism. At the same time, the CASSCF and MCQDPT2 approaches, which correctly place the conrotatory pathway below the disrotatory one, provide relatively poor energetics when compared to the available experimental and more accurate electronic structure data [189], although CASSCF produces reasonable geometries of the corresponding stationary points. On the other hand, as was also shown in Ref. [189], the CR-CC(2,3) approach provides an accurate and balanced description of the conrotatory and disrotatory pathways describing the isomerization of bicyclo[1.1.0]butane to *trans*-buta-1,3-diene, and one of the primary goals of this thesis research was to produce CR-CC(2,3) energies for this system which were converged with the basis set [193]. To do this, the set of nuclear geometries optimized at the CASSCF(10,10)/cc-pVDZ level of theory were taken from Ref. [189] and additional CR-CC(2,3) calculations

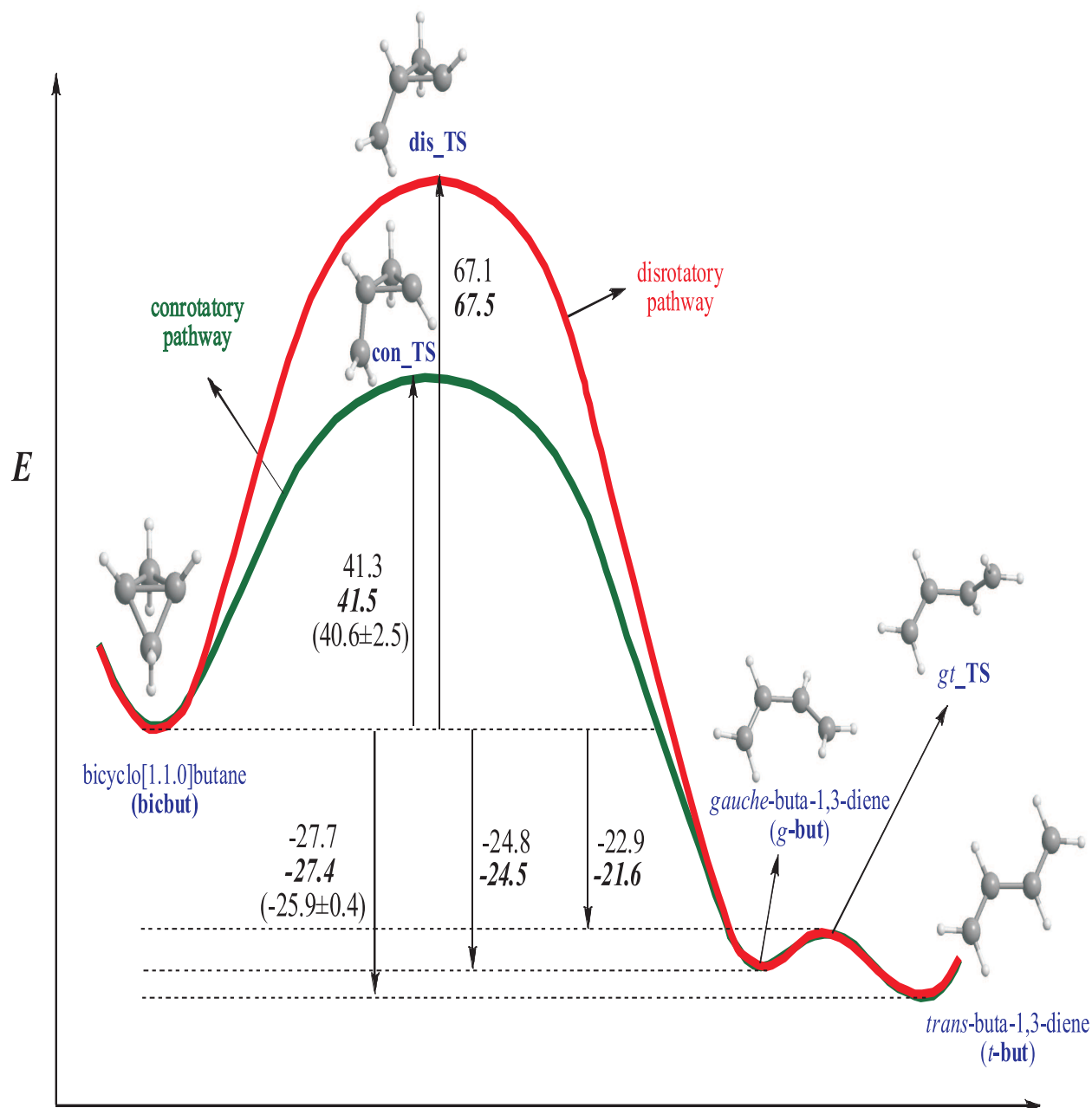


Figure 3.2: The conrotatory and disrotatory pathways describing the isomerization of bicyclo[1.1.0]butane to *trans*-buta-1,3-diene, along with the enthalpy values relative to the reactant at all stationary points obtained in the explicit CR-CC(2,3)/cc-pVQZ//CASSCF(10,10)/cc-pVDZ calculations (numbers in roman) and the enthalpy values obtained with the PES extrapolation procedure using the CR-CC(2,3)/cc-pVDZ, CR-CC(2,3)/cc-pVTZ, and RHF/cc-pVQZ energies at all stationary points resulting from the CASSCF(10,10)/cc-pVDZ optimizations, and the CR-CC(2,3)/cc-pVQZ correlation energy at the reactant geometry (numbers in bold italics). The available experimental enthalpy values are in parentheses. All enthalpies are in kcal/mol.

Table 3.8: Relative enthalpies, in kcal/mol, with respect to reactant as calculated at various levels of theory for the conrotatory and disrotatory **bicbut** $\rightarrow$ **tbut** isomerization pathways.

Method	con_TS	dis_TS	<i>g</i> -but	<i>gt</i> -TS	tbut
CR-CC(2,3)/cc-pVTZ ^a	41.1	66.1	-24.9	-22.1	-27.9
CR-CC(2,3)/cc-pVQZ ^b	41.3	67.1	-24.8	-22.9	-27.7
CR-CC(2,3)/CBS limit ^b	41.5	67.5	-24.5	-21.6	-27.4
CR-CC(2,3)+Q/cc-pVTZ ^b	40.8	67.0	-24.9	-22.1	-27.8
ACSE/6-311G** ^c	41.2	55.7	-23.8	—	—
DMC ^d	40.4(5)	58.6(5)	-25.2(5)	-22.2(5)	-27.9(5)
Experiment ^e	40.6($\pm$ 2.5)	—	—	—	-25.9(4)

^aTaken from Ref. [189]. ^bThe present study (CR-CC(2,3) results taken from Ref. [193]).

^cResults obtained with the anti-Hermitian contracted Schrödinger equation (ACSE) approach taken from Ref. [253]. ^dBenchmark computational results obtained with the diffusion Quantum Monte Carlo (DMC) approach taken from Ref. [254]. ^eBenchmark experimental results taken from Ref. [252].

were performed with the cc-pVQZ basis set and, together with the cc-pVTZ energies from Ref. [189], CBS-limit extrapolated correlation energies were obtained using Eq. (3.98). These and other results are reported in Table (3.8).

Enthalpy values in Table (3.8) are reported relative to reactants as calculated using a variety of approaches. Comparison of the CR-CC(2,3)/cc-pVTZ results from Ref. [189] with the CR-CC(2,3) results obtained through calculations at the cc-pVQZ and CBS-limit levels, as originally reported in Ref. [193], confirms that small improvements were made by saturating the basis set. Later in 2008, a paper was published challenging the accuracy of these calculations. Calculations performed using the anti-Hermitian contracted Schrödinger equation (ACSE) in Ref. [253], reported for comparison in Table (3.8), showed the dis_TS barrier to be much lower than CR-CC(2,3) calculations had predicted. Since the ACSE method is good at accounting for non-dynamical correlation effects, we calculated the +Q

corrections at various basis set levels to assure that we were not neglecting nondynamical correlation effects described by higher-order clusters. Since, as shown in Table (3.8), these corrections adjusted the enthalpies by less than 1 kcal/mol, we assumed that our values for the **dis_TS** barrier were reliable.

An interesting recent development, which affects this discussion, came in 2010 when high-quality diffusion Quantum Monte Carlo (DMC) results became available for the **bicbut** $\rightarrow$ **tbut** isomerization system [254]. By comparing with this theoretical benchmark data it was shown that CR-CC(2,3) and CR-CC(2,3)+Q enthalpy values at all reported basis sets are within ~ 1 kcal/mol for all stationary points except for the **dis_TS** structure. There is a rather large discrepancy at the **dis_TS** stationary point with the highest quality CR-CC data predicting a barrier of around 67 kcal/mol, while the DMC results predict a **dis_TS** barrier around 59 kcal/mol. At this point the issue emerged: what is behind the few kcal/mol difference between the CR-CC and DMC results for the disrotatory transition state? Our group finally found an answer a few weeks ago. Based on a still preliminary study using a new theory being presently developed in our group, abbreviated CC(t;3), that represents a merger of the active-space CCSDt approach and CR-CC(2,3) [255], it appears to be the iterative treatment of the singly and doubly excited clusters in the presence of the leading triple excitations which accounts for the remaining nondynamic and dynamic correlations responsible for this energy lowering, producing a result in perfect agreement with DMC. Thus, while species with extreme biradical character, such as **dis_TS**, may sometimes be difficult to describe to high accuracy, the CR-CC(2,3) method remains the key in obtaining the desired accuracies, since CC(t;3) uses the CR-CC(2,3)-style corrections to correct the CCSDt energies for the triples missing in the CCSDt calculations, while eliminating the

complete failure of CCSD(T).

Excluding the exceptionally difficult **dis_TS** stationary point, which is not part of the energetically favorable reaction profile for the **bicbut** $\rightarrow$ **tbut** isomerization, the CR-CC(2,3) results with or without quadruples are within 1 kcal/mol of the DMC results. It also appears that the best CR-CC results are within the DMC error bars. Meanwhile, the ACSE results significantly underestimate the *g*-but barrier and cannot claim sub-chemical accuracy. We speculate that the ACSE method does significantly better for the **dis_TS** barrier, which has 90 % biradical character, but underestimates the correlation energies for the closed-shell species. This is because the method is very good for describing nondynamical correlation energy, but is less accurate in its description of dynamical correlation effects. On the other hand, unlike CCSD(T), which describes dynamical correlation energy well and nondynamical correlation energy poorly, it is at least a qualitatively predictive method. Meanwhile, the CR-CC(2,3) method and its CR-CC(2,3)+Q and CC(t;3) extensions can, unlike the ACSE approach, give a quantitatively correct description of the energetically favorable **con_TS** reaction profile and, unlike the CCSD(T) method, predict the correct reaction channel.

The success of the ground-state CR-CC formalism in this and previous sections motivates an investigation of the performance of the CR-EOMCC methods for excited states. This is done in the next two sections, with the first considering the performance of CR-EOMCC(2,3) for the description of low-lying excited-state PESs of the water molecule, particularly along the bond-breaking O-H coordinate (in Sect. (3.2.4)), followed by a fully size-intensive description of vertical excitation energies and spectral shifts in weakly bound complexes (in Sect. (3.2.5)).

3.2.4 Excited-State Potential Energy Surfaces for the Dissociation of Water

The motivation for the work reported in the current section is to, for the first time, assess the performance of the CR-EOMCC(2,3) approach in generating excited-state PESs along a single-bond stretching dissociation channel. The water molecule was chosen as our target system since, while there exist in the literature many experimental [256–264] and theoretical [265–269] studies aimed at finding vertical excitation energies out of the ground state (see Ref. [270] for a review of older work), fewer studies have focused on generating PESs other than for the ground and first excited singlet states. Recently, however, a cut of the full ground-state PES corresponding to the asymmetric O-H bond-breaking dissociation was carefully optimized by Li and Paldus using the CCSD theory in conjunction with a relatively large cc-pVTZ basis set and then many low-lying excited-state PES cuts were calculated at the full CI level using a cc-pVTZ basis stripped of polarization functions (referred to as the TZ basis) [271]. Here, we intend to use the set of high-level optimized geometries and corresponding full CI excited-state PES cuts reported in Ref. [271] as benchmark data for testing the performance of the CR-CC(2,3) and CR-EOMCC(2,3) methods. The main objective is to determine whether the CR-CC(2,3) and CR-EOMCC(2,3) methods can overcome the deficiencies of the CCSD and EOMCCSD approaches for the PES cuts of interest, including stretched O-H bond lengths and the $\text{H}_2\text{O} \rightarrow \text{H} + \text{OH}$ dissociation limit, where groups of electronic states of water converge to become degenerate. Since values are available, we will also take this opportunity to compare the PESs obtained in the CR-CC(2,3)/CR-EOMCC(2,3) calculations with those produced by leading MRCC or MRCC-like methods, as taken from Ref. [271]. These include the NR-RMR-CCSD(T), NR-GMS-SU-CCSD, and

(N, M)-CCSD methods, all of which were mentioned in the Introduction.

Following Ref. [271], calculations were performed using a small basis set referred to as TZ, which, as mentioned above, is constructed by removing polarization functions from the cc-pVTZ basis set of Dunning. The geometries associated with the asymmetric dissociation were generated by designating the bond length of the stretched O-H bond and optimizing the H-O-H angle and second O-H bond length at the CCSD/cc-pVTZ level. These geometrical parameters were taken from Ref. [271]. All calculations of singlet ground and excited states were performed using the efficient RHF-based EOMCCSD and CR-EOMCC(2,3) computer codes developed earlier by our group and incorporated in the GAMESS program. Triplet EOMCC calculations were performed using a new EOMCCSD code for ROHF references developed as part of this thesis effort, to be included in a future official GAMESS distribution, the implementation of which is discussed in Sect. (5.1.1). Open-shell ROHF-based CR-EOMCC(2,3) calculations were performed using the spin-orbital computer codes written by our group that are loosely interfaced with GAMESS. In all calculations electrons in the lowest occupied molecular orbital were frozen, spherical contaminants were dropped from the atomic orbital basis sets, and the C_s point group was enforced, distinguishing states as either A' or A'' symmetry. We use the convention for numbering states where the singlet ground state is denoted in a usual way by the symbol X and excited states belonging to the same irreducible representation, i.e. spin and spatial symmetry, are numbered sequentially in order of increasing energy at the equilibrium nuclear configuration.

Ground-state X^1A' PES cuts are reported in Table (3.9), with internuclear separations given in the first column and energies corresponding to various methods given in the remaining columns. The last two rows provide MUEs and non-parallelity errors (NPEs) relative

Table 3.9: Ground-state X^1A' energies for the asymmetric single-bond breaking of H_2O at a series of the O-H bond lengths, R , in atomic units. The last two rows give the mean unsigned errors (MUE) and non-parallelity errors (NPE) relative to the full CI PES obtained in Ref. [271].

R (bohr)	FCI ^a	7R-SU ^b	4R-RMR(T) ^c	CCSD ^d	CR-CC(2,3) ^d
1.3	-1.01567	2.87	0.08	2.84	-0.16
1.6	-1.14894	3.20	0.10	3.17	-0.16
1.809	-1.16847	3.50	0.13	3.46	-0.14
2.0	-1.16417	3.83	0.16	3.77	-0.12
2.4	-1.13058	4.65	0.20	4.59	-0.03
2.8	-1.09255	5.73	0.20	5.72	0.08
3.2	-1.06129	7.15	0.15	7.21	0.21
3.6	-1.03889	8.84	0.13	8.97	0.32
4.0	-1.02451	10.69	0.19	10.81	0.36
4.2	-1.01967	11.61	0.24	11.66	0.33
4.4	-1.01603	12.48	0.29	12.43	0.28
MUE	—	6.68	0.17	6.78	0.20
NPE	—	9.61	0.21	9.59	0.52

^a Full CI energies (E), reported as $(E + 75)$ hartree, taken from Ref. [271]. ^b 7R-GMS-SU-CCSD energies, in millihartree, reported as differences from full CI taken from Ref. [271] ^c 4R-RMR-CCSD(T) energies, reported as differences from full CI taken from Ref. [271] ^d This work. Reported as differences from full CI, in millihartree

to the full CI PES of Li and Paldus [271] for each data set. Let us first consider the errors produced by the CCSD method, which range from 2.84 to 12.43 hartree, with MUE of 6.78 millihartree and NPE of 9.59 millihartree. As expected, errors for all methods increase significantly as the O-H bond is stretched, but it is seen that the addition of triple excitations via the CR-CC(2,3) method reduces absolute errors relative to full CI to no greater than 0.36 millihartree for all points considered. The MUE of 0.20 millihartree and NPE of 0.52 millihartree produced by the black-box CR-CC(2,3) approach compare very well to the expert high-level 4R-RMR-CCSD(T) method, which yields a MUE of 0.17 millihartree and a NPE of 0.21 millihartree.

Results for the low-lying singlet and triplet excited states of the A' symmetry are collected in Tables (3.10) and (3.11), respectively. Each table is broken into three sections, organizing results for the three reported excited states of a given symmetry under the headings $1^1A'$, $2^1A'$, and $3^1A'$ in Table (3.10) and $1^3A'$, $2^3A'$, and $3^3A'$ in Table (3.11). In both tables it is clear that the EOMCCSD method does well for all calculated states in the $R = 1.3 - 2.0$ bohr region, i.e., near the equilibrium geometry on the ground-state PES, but clearly fails at geometries 2.4 bohr and beyond, producing errors relative to full CI of 20-35 millihartree. The CR-EOMCC(2,3) approach improves significantly upon the EOM-CCSD results in the stretched geometry region of each PES cut, reducing the MUEs to no more than 3.75 millihartree. To further assess the quality of the CR-EOMCC(2,3) results, and better appreciate the performance of the “black-box” CR-EOMCC(2,3) theory, comparison with the MRCC results of Ref. [271] is made. The results obtained with the expert 7R-GMS-SU-CCSD approach are worse in all cases, with only a few exceptions. For the $3^1A'$ state, CR-EOMCC(2,3) does significantly better than 7R-GMS-SU-CCSD, while for the $3^3A'$ state 7R-GMS-SU-CCSD does significantly better than CR-EOMCC(2,3). Overall, it may be concluded that the “black-box” CR-EOMCC(2,3) method performs better than the complicated, expert 7R-GMS-SU-CCSD approach for the A' states reported in Tables (3.10) and (3.11).

Now, moving to results for the A'' symmetry states, which are presented for the two lowest-lying singlet states in Table (3.12) and for the two lowest-lying triplet states in Table (3.13), many of the same trends arise as in the case of the A' states, including the failure of EOMCCSD for stretched nuclear geometries and the great improvements offered by the CR-EOMCC(2,3) method. For the $2^1A''$ state results could not be obtained using the CR-

EOMCC(2,3) method for the O-H bond length R of 4.4 bohr, so for this one state, MUEs and NPEs for all methods are calculated using only 10 points, from the $R = 1.3 - 4.2$ bohr region. For all the A'' symmetry states, the 4R-GMS-SU-CCSD and higher-quality (8,4)-CCSD MRCC results can be found in Ref. [271], so both types of calculations are reported in the tables. Comparing the performance of the CR-EOMCC(2,3) method with the (8,4)-CCSD results for the $1^1A''$, $2^1A''$, $1^3A''$, and $2^3A''$ states, it is clear that these methods produce similar results in every case. In a few cases, CR-EOMCC(2,3) does slightly better and in others (8,4)-CCSD is slightly better. For example, looking at the $2^1A''$ state, CR-EOMCC(2,3) produces a MUE relative to full CI of 0.89 millihartree and a NPE of 2.66 millihartree, which is better than the values produced by (8,4)-CCSD, of 3.80 and 9.66 millihartree, respectively, but for the $2^3A''$ state the (8,4)-CCSD method produces slightly better values, with values for MUE and NPE of 2.16 and 1.67 millihartree, respectively, as compared with the MUE of 3.19 millihartree and the NPE of 13.64 millihartree produced by CR-EOMCC(2,3). It may be concluded from comparison with the NR-RMR-CCSD(T), NR-GMS-SU-CCSD, and (N,M) -CCSD results reported in Ref. [271] that the CR-CC(2,3) and CR-EOMCC(2,3) methods produce results of the MRCC or better quality for all of the low-lying PESs of water examined here.

Table 3.10: Same as Table (3.9) for the three lowest-lying $1A'$ states.

R (bohr)	$1\ 1A'$				$2\ 1A'$				$3\ 1A'$			
	FCI ^a	7R-SU ^b	EOMSD ^c	(2,3) ^d	FCI ^a	7R-SU ^b	EOMSD ^c	(2,3) ^d	FCI ^a	7R-SU ^b	EOMSD ^c	(2,3) ^d
1.3	-0.62821	2.07	-0.71	0.27	-0.54185	1.97	-1.49	0.10	-0.39097	3.20	0.64	0.92
1.6	-0.77055	2.49	-0.34	0.49	-0.86730	2.25	-0.63	0.27	-0.57861	3.12	1.29	1.22
1.809	-0.79860	2.93	0.34	0.64	-0.71786	2.55	0.42	0.40	-0.63701	3.36	2.17	1.36
2.0	-0.80575	3.63	1.56	0.76	-0.72455	2.94	1.53	0.52	-0.66800	3.83	3.34	1.46
2.4	-0.81062	5.95	6.14	1.00	-0.71631	2.91	5.29	0.96	-0.69672	4.47	5.46	0.81
2.8	-0.81932	5.55	11.15	0.89	-0.72567	4.00	12.97	1.26	-0.67863	4.72	6.82	1.45
3.2	-0.82860	4.56	15.23	0.39	-0.73203	4.51	20.47	-0.46	-0.66287	8.51	27.97	2.28
3.6	-0.83618	4.17	18.42	-0.34	-0.73289	3.74	27.73	-1.91	-0.66382	19.90	26.69	8.57
4.0	-0.84172	4.04	20.79	-1.20	-0.72999	3.06	32.68	-4.57	-0.66878	28.01	32.64	8.34
4.2	-0.84379	3.98	21.69	-1.64	-0.72738	3.08	33.89	-4.98	-0.66943	29.03	34.21	7.33
4.4	-0.84546	3.92	22.40	-2.07	-0.72415	3.32	34.35	-4.92	-0.66889	28.68	-20.91	7.56
MUE	—	3.94	10.80	0.88	—	3.12	15.59	1.85	—	12.44	14.74	3.75
NPE	—	3.88	23.11	3.07	—	2.54	35.84	6.24	—	25.91	55.12	7.76

^a Full CI energies (E), reported as ($E + 75$) hartree, taken from Ref. [271]. ^b 7R-GMS-SU-CCSD energies, reported as differences from full CI, in millihartree, taken from Ref. [271]. ^c EOMCCSD energies, reported as differences from full CI, in millihartree. ^d CR-EOMCC(2,3) energies, reported as differences from full CI, in millihartree.

Table 3.11: Same as Table (3.9) for the three lowest-lying ${}^3A'$ states.

R (bohr)	$1\ {}^3A'$				$2\ {}^3A'$				$3\ {}^3A'$			
	FCI ^a	7R-SU ^b	EOMSD ^c	(2,3) ^d	FCI ^a	7R-SU ^b	EOMSD ^c	(2,3) ^d	FCI ^a	7R-SU ^b	EOMSD ^c	(2,3) ^d
1.3	-0.65020	2.94	-0.93	0.30	-0.55577	2.92	-1.63	0.18	-0.44370	4.57	0.29	0.84
1.6	-0.79340	3.32	-0.70	0.51	-0.70872	3.76	-0.59	0.42	-0.61947	4.28	1.03	1.21
1.809	-0.82511	3.79	-0.22	0.65	-0.74823	4.68	0.62	0.67	-0.67629	4.04	1.88	1.40
2.0	-0.84272	4.50	0.91	0.81	-0.76263	5.49	1.47	0.92	-0.71482	4.82	2.87	1.30
2.4	-0.89158	5.65	3.76	1.26	-0.78565	5.68	4.85	1.22	-0.73737	7.32	5.29	1.41
2.8	-0.93780	5.57	4.75	1.41	-0.81625	4.96	8.94	1.08	-0.70581	8.02	4.51	1.46
3.2	-0.96789	5.13	4.89	1.41	-0.83261	4.26	12.71	0.82	-0.67660	7.10	-6.01	3.81
3.6	-0.98540	4.56	4.70	1.42	-0.84103	3.79	16.07	0.34	-0.65564	6.22	12.83	3.87
4.0	-0.99516	3.99	4.35	1.45	-0.84553	3.48	18.81	-0.28	-0.64297	7.09	19.65	4.77
4.2	-0.99822	3.75	4.15	1.46	-0.84693	3.37	19.91	-0.62	-0.63921	8.22	33.47	18.41
4.4	-1.00046	3.55	3.94	3.70	-0.84797	3.28	20.82	-0.97	-0.63687	9.63	34.58	17.25
MUE	—	4.36	2.78	1.03	—	4.32	7.42	0.70	—	6.07	7.58	3.75
NPE	—	2.71	5.82	1.16	—	2.76	21.54	1.84	—	4.18	39.48	17.57

^a Full CI energies (E), reported as ($E + 75$) hartree, taken from Ref. [271]. ^b 7R-GMS-SU-CCSD energies, reported as differences from full CI, in millihartree, taken from Ref. [271]. ^c EOMCCSD energies, reported as differences from full CI, in millihartree. ^d CR-EOMCC(2,3) energies, reported as differences from full CI, in millihartree.

Table 3.12: Same as Table (3.9) for the two lowest-lying $1A''$ states.

R (bohr)	$1\ 1A''$					$2\ 1A''$				
	FCI ^a	4R-SU ^b	(8,4)-SD ^c	EOMSD ^d	(2,3) ^e	FCI ^a	4R-SU ^b	(8,4)-SD ^c	EOMSD ^d	(2,3) ^e
1.3	-0.70187	1.80	1.56	-0.71	0.28	-0.61400	1.89	1.53	-1.39	0.21
1.6	-0.85127	2.11	1.73	-0.54	0.53	-0.76861	2.21	1.66	-0.67	0.37
1.809	-0.88566	2.37	1.84	-0.22	0.62	-0.80694	2.46	1.77	0.18	0.44
2.0	-0.90098	2.53	1.91	0.62	0.62	-0.81992	2.78	1.88	0.88	0.47
2.4	-0.92945	2.69	1.87	4.44	0.63	-0.81008	3.50	1.98	1.01	0.41
2.8	-0.95679	2.87	1.65	8.51	0.54	-0.78177	4.47	2.03	0.70	0.40
3.2	-0.97626	3.14	1.34	11.88	0.08	-0.75430	6.21	2.75	0.78	0.58
3.6	-0.98862	3.36	0.94	14.43	-0.64	-0.73276	9.12	5.15	1.57	1.04
4.0	-0.99617	3.47	0.56	16.15	-1.48	-0.71846	13.52	8.89	3.43	2.08
4.2	-0.99870	3.48	0.47	16.73	-1.90	-0.71372	16.02	10.32	4.78	2.87
4.4	-1.00065	3.48	0.66	17.16	-2.30	-0.71037	18.04	11.19	6.35	N.C.
MUE	—	2.85	1.32	8.31	0.87	—	6.22	3.80	1.54	0.89
NPE	—	1.68	1.44	17.44	2.53	—	8.79	9.66	7.74	2.66

^a Full CI energies (E), reported as ($E+75$) hartree, taken from Ref. [271]. ^b 4R-GMS-SU-CCSD energies, reported as differences from full CI, in millihartree, taken from Ref. [271]. ^c (8,4)-CCSD energies, reported as differences from full CI, in millihartree, taken from Ref. [271]. ^d EOMCCSD energies, reported as differences from full CI, in millihartree. ^e CR-EOMCC(2,3) energies, reported as differences from full CI, in millihartree.

Table 3.13: Same as Table (3.9) for the two lowest-lying ${}^3A''$ states.

R (bohr)	$1\ {}^3A''$					$2\ {}^3A''$				
	FCI ^a	4R-SU ^b	(8,4)-SD ^c	EOMSD ^d	(2,3) ^e	FCI ^a	4R-SU ^b	(8,4)-SD ^c	EOMSD ^d	(2,3) ^e
1.3	-0.72170	2.77	1.88	-0.89	0.32	-0.62356	2.75	1.81	-1.19	0.29
1.6	-0.87187	3.18	2.09	-0.73	0.56	-0.78052	3.31	2.13	-0.23	0.53
1.809	-0.90726	3.60	2.29	-0.41	0.66	-0.82109	3.98	2.45	0.80	0.67
2.0	-0.92363	4.08	2.52	0.44	0.68	-0.83571	4.68	2.70	1.61	0.79
2.4	-0.95148	4.87	2.89	4.03	0.84	-0.82785	5.83	2.68	2.02	0.92
2.8	-0.97420	4.72	2.95	7.88	1.00	-0.80175	7.33	2.36	2.59	1.36
3.2	-0.98837	4.31	3.00	11.40	0.88	-0.77863	10.70	2.26	5.19	2.92
3.6	-0.99640	4.00	3.20	14.49	0.50	-0.76634	17.79	2.03	12.44	7.31
4.0	-1.00091	3.81	3.43	16.94	-0.05	-0.76393	25.85	1.36	16.94	12.59
4.2	-1.00236	3.72	3.47	17.89	-0.34	-0.76335	27.93	1.03	26.41	13.93
4.4	-1.00345	3.61	3.25	18.68	0.50	-0.76232	28.62	0.62	29.01	14.51
MUE	—	3.92	2.70	6.46	0.64	—	9.37	2.16	5.83	3.19
NPE	—	2.10	1.59	18.78	1.34	—	25.18	1.67	27.60	13.64

^a Full CI energies (E), reported as ($E+75$) hartree, taken from Ref. [271]. ^b 4R-GMS-SU-CCSD energies, reported as differences from full CI, in millihartree, taken from Ref. [271]. ^c (8,4)-CCSD energies, reported as differences from full CI, in millihartree, taken from Ref. [271]. ^d EOMCCSD energies, reported as differences from full CI, in millihartree. ^e CR-EOMCC(2,3) energies, reported as differences from full CI, in millihartree.

A particularly weak aspect of the SR methods, when used to describe bond-breaking PES cuts, is their erratic behavior as states become degenerate at the dissociation limit. In order to assess the quality of the CC and EOMCC PES cuts considered in this report, a few states have been plotted in Figure (3.3). The states were carefully chosen such that they all converge with at least one other plotted state to become degenerate at the largest internuclear separation. In the plot, these electronic states have been identified as those resulting from combination of the doublet S ground-state of the hydrogen radical with the lowest-energy $^2\Pi$, $^2\Sigma^+$, and $^2\Sigma^-$ states of the hydroxyl radical. It is quite clear from visual inspection of Figure (3.3(a)) that the CCSD/EOMCCSD curves, which should converge to become degenerate as the dissociation occurs, are for the most part still not falling on top of one another as the O-H distance becomes large. This is especially evident for the highest-energy pair of curves, $2^1A'$ and $3^3A''$, which are still separated by 20 millihartree at $R = 4.4$ bohr. Significant improvement is seen in Figure (3.3(b)) where the CR-CC(2,3)/CR-EOMCC(2,3) curves exhibit much better asymptotic behavior compared with the CCSD/EOMCCSD data, and also give excellent agreement with the full CI potentials. However, the most challenging pair of curves for EOMCCSD, $2^1A'$ and $3^3A''$, also give CR-EOMCC(2,3) trouble, as we were not able to obtain CR-EOMCC(2,3) results for the $2^1A'$ state at $R = 4.4$ bohr, although the energies which were obtained for these two PESs showed dramatic overall improvement over the corresponding CCSD/EOMCCSD results.

To summarize, in this section it has been shown that the CR-EOMCC(2,3) method is capable of providing a highly accurate description of excited-state PESs of water involving single bond-breaking. However, excited-state PESs are not directly observable quantities, so next we move toward calculating spectroscopic properties which are observable. In the

next section, the utility of the size-intensive δ -CR-EOMCC(2,3) method in calculations of vertical excitation energies and spectral shifts in an organic chromophore, resulting from the introduction of hydrogen-bonded molecular environments, is examined.

3.2.5 Excitation Energies and Hydrogen-Bonding-Induced Spectral Shifts in Complexes of *cis*-7-Hydroxyquinoline

In spite of being relatively weak, non-covalent interactions with the environment, such as hydrogen bonds, can qualitatively affect the electronic structure and properties of the embedded molecules. Accurately predicting the effect of a hydrogen-bonded environment on the electronic structure of embedded molecules represents a challenge for computational chemistry. Among such properties, electronic excitation energies are of great interest in view of the common use of organic chromophores as probes in various environments [272–275]. Typically, hydrogen bonding results in shifts in the positions of the maxima of the absorption and emission bands anywhere between a few hundred and about 3000 cm^{-1} [276]. Thus, in order to be able to use computer modeling for interpretation of experimental data, the intrinsic errors of the calculated shifts must be very small, on the order of 100 cm^{-1} or less.

We became involved in a joint project with the Wesolowski Geneva group aimed at comparing shifts in the $\pi \rightarrow \pi^*$ excitation energy of the *cis*-7-hydroxyquinoline (*cis*-7HQ) chromophore resulting from the formation of hydrogen-bonded complexes between *cis*-7HQ and a number of small molecules [201]. Our main role was to provide reference excitation energies using EOMCC methods, which were subsequently used to benchmark the so-called frozen-density embedding theory (FDET) approach, developed by Wesolowski et al. [277–281], and the conventional supermolecular form of the time-dependent density functional theory

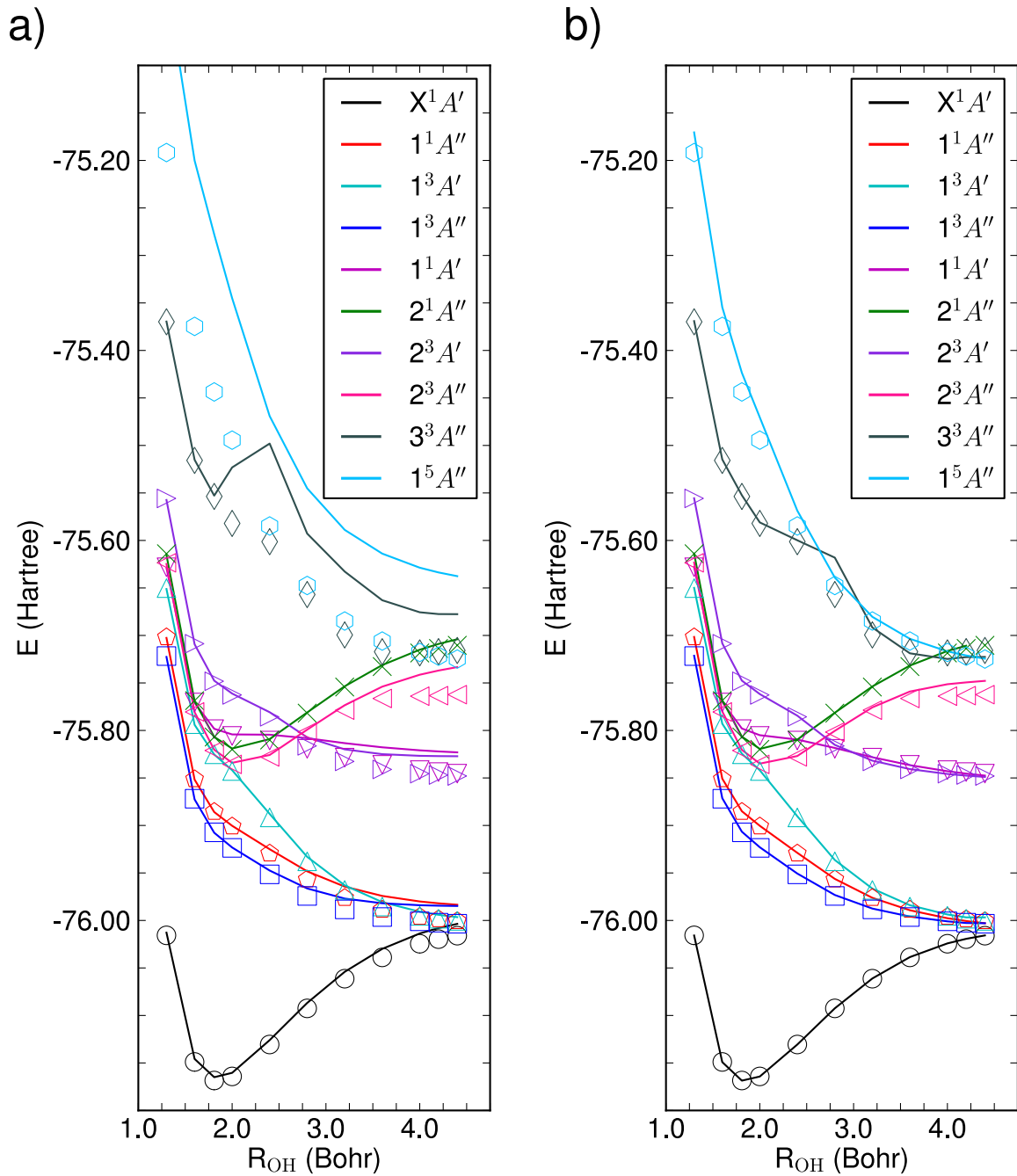


Figure 3.3: Cuts of the PESs for a few ground- and excited-states of a single-bond stretching model of H_2O as obtained with **(a)** the CCSD/EOMCCSD and **(b)** the CR-CC(2,3)/CR-EOMCC(2,3) approaches, and the TZ basis set. Lines are used to represent the CC/EOMCC data, while the corresponding full CI values are represented by points (which are identically replicated in **(a)** and **(b)**).

(TDDFT), frequently used in spectral shift calculations. Since the main goal of this thesis is to examine the performance of the CR-CC/CR-EOMCC and other CC/EOMCC approaches, our discussion below focuses on a comparison of the EOMCCSD and CR-EOMCC(2,3)-level results with the available experimental data for the $\pi \rightarrow \pi^*$ excitation energies and the corresponding environment-induced spectral shifts in the 7HQ chromophore [276]. The environment molecules constituting the eight complexes involving 7HQ that were examined in Ref. [201] included (i) a single water molecule, (ii) a single ammonia molecule, (iii) a water dimer, (iv) a single molecule of methanol, (v) a single molecule of formic acid, (vi) a trimer consisting of ammonia and two water molecules, (vii) a trimer consisting of ammonia, water, and ammonia, and (viii) a trimer consisting of two ammonia and one water molecules (see Figure 3.2.5).

In order to establish the appropriate level of EOMCC theory to serve as a reference for the FDET and supermolecular TDDFT calculations reported in Ref. [201], we first examined the dependence of the vertical excitation energies $\omega_{\pi \rightarrow \pi^*}$ and environment-induced spectral shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ on the basis set. Table (3.14) compares the results of the EOMCCSD calculations obtained with the 6-31+G(d), [282–284] 6-31++G(d,p), [282–284] 6-311+G(d), [284,285] and aug-cc-pVDZ basis sets, as well as with the [5s3p2d/3s2p] basis of Sadlej [286], designated as POL, for the two smallest complexes, 7HQ $\cdots$ H₂O and 7HQ $\cdots$ NH₃, for which we could afford the largest number of computations. The results in Table (3.14) indicate that although the vertical excitation energies $\omega_{\pi \rightarrow \pi^*}$ in the bare cis-7HQ system and its complexes with the water and ammonia molecules vary with the basis set (for the basis sets tested here by as much as about 600 cm⁻¹), the environment-induced shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ are almost insensitive to the basis set choice, varying by at most 27 cm⁻¹ between all basis sets reported. Although we

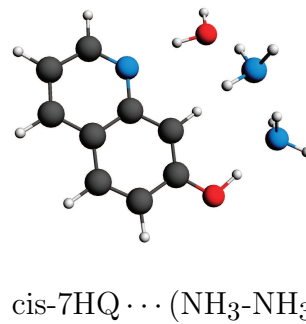
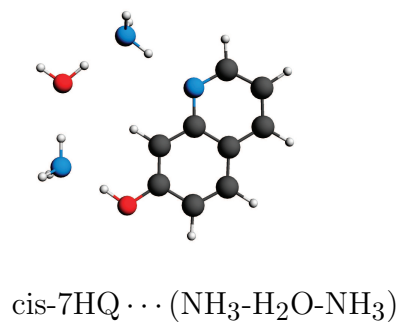
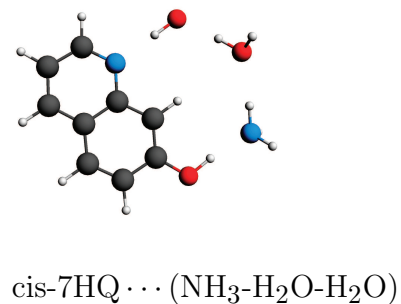
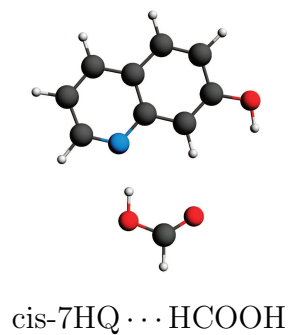
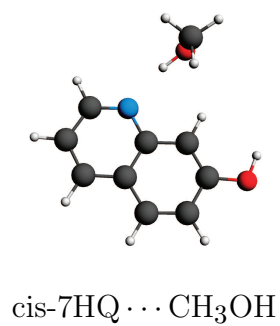
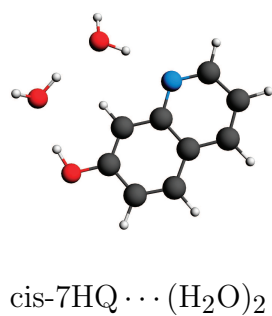
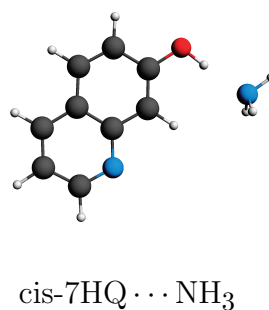
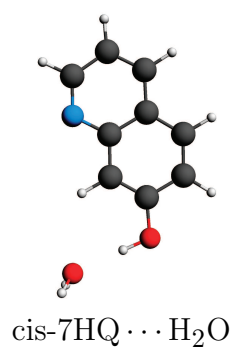


Figure 3.4: The eight hydrogen-bonded complexes of the cis-7HQ molecule.

Table 3.14: The basis-set dependence of the vertical excitation energies $\omega_{\pi \rightarrow \pi^*}$ and the environment-induced shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ (in cm^{-1}) obtained with the EOMCCSD approach corresponding to the lowest $\pi \rightarrow \pi^*$ transition in the cis-7HQ chromophore and its complexes with the water and ammonia molecules.

Basis set	$\omega_{\pi \rightarrow \pi^*}$			$\Delta\omega_{\pi \rightarrow \pi^*}$	
	7HQ	7HQ $\cdots$ H ₂ O	7HQ $\cdots$ NH ₃	7HQ $\cdots$ H ₂ O	7HQ $\cdots$ NH ₃
6-31+G(d)	35171	34643	34396	-528	-775
6-31++G(d,p)	35120	34597	34351	-523	-769
6-311+G(d)	35046	34500	34263	-546	-783
aug-cc-pVDZ	34707	34182	33923	-525	-784
POL	34596	34077	33819	-519	-777

were unable to perform a similarly thorough analysis for other complexes due to prohibitive computer costs, we were able to obtain the EOMCCSD $\omega_{\pi \rightarrow \pi^*}$ and $\Delta\omega_{\pi \rightarrow \pi^*}$ values for all of the complexes examined in this study using the 6-31+G(d) and 6-311+G(d) basis sets. As shown in Table (3.15), the differences between the EOMCCSD/6-31+G(d) and EOMCCSD/6-311+G(d) values of the environment-induced shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ remain small for all complexes of interest, ranging from 8 cm^{-1} in the 7HQ $\cdots$ NH₃ case to 43 cm^{-1} in the case of 7HQ $\cdots$ (H₂O)₂, or 1–3 %. Thus, we can conclude that the choice of the basis set, although important for obtaining the converged $\omega_{\pi \rightarrow \pi^*}$ values, is of little significance when the environment-induced shifts in the vertical excitation energy corresponding to the lowest $\pi \rightarrow \pi^*$ transition in the cis-7HQ chromophore are considered with the EOMCC methods.

Although the EOMCCSD approach is known to provide an accurate description of excited states dominated by one-electron transitions, such as the $\pi \rightarrow \pi^*$ transition in cis-7HQ and its complexes, there have been cases of similar states reported in the literature, where the EOMCCSD level has not been sufficient to obtain high-quality results [287, 288]. Moreover, our interest in this study is in the small energy differences defining the environment-induced

shifts $\Delta\omega_{\pi\rightarrow\pi^*}$, which may be sensitive to the higher-order correlation effects neglected in the EOMCCSD calculations. For this reason, we also examined the effect of triples corrections to EOMCCSD energies on the calculated $\omega_{\pi\rightarrow\pi^*}$ and $\Delta\omega_{\pi\rightarrow\pi^*}$ values by performing the δ -CR-EOMCC(2,3) calculations with the 6-31+G(d) basis set. The results of these calculations, shown in Table (3.15), indicate that triple excitations have a significant effect on the vertical excitation energies $\omega_{\pi\rightarrow\pi^*}$, reducing the 4000-5000 cm^{-1} differences between the EOMCCSD and experimental data to no more than about 800 cm^{-1} , when the δ -CR-EOMCC(2,3),A/6-31+G(d) calculations are performed, and no more than about 500 cm^{-1} when the δ -CR-EOMCC(2,3),D/6-31+G(d) approach is employed, while bringing the $\Delta\omega_{\pi\rightarrow\pi^*}$ values closer to the experimentally observed shifts when compared with EOMCCSD values. Although the differences between the δ -CR-EOMCC(2,3) and EOMCCSD values of the environment-induced shifts $\Delta\omega_{\pi\rightarrow\pi^*}$ resulting from the calculations with the 6-31+G(d) basis set do not exceed 15–16 % of the EOMCCSD values, triples corrections improve the EOMCCSD results and, as such, are useful for the generation of the reference EOMCC data.

Table 3.15: The vertical excitation energies $\omega_{\pi \rightarrow \pi^*}$ and the environment-induced shifts $\Delta\omega_{\pi \rightarrow \pi^*}$ (in cm^{-1}) obtained with the EOMCCSD/6-31+G(d), EOMCCSD/6-311+G(d), δ -CR-EOMCC(2,3), A/6-31+G(d), and δ -CR-EOMCC(2,3), D/6-31+G(d) approaches, and their composite EOMCC corresponding to the lowest $\pi \rightarrow \pi^*$ transition in the cis-7HQ chromophore and its various complexes.

Environment	EOMCCSD/ 6-31+G(d)	EOMCCSD/ 6-311+G(d)	δ -CR-EOMCC(2,3), A/ 6-31+G(d)	δ -CR-EOMCC(2,3), D/ 6-31+G(d)	EOMCC,A ^a	EOMCC,D ^b	Exp. ^c
$\omega_{\pi \rightarrow \pi^*}$							
None	35171	35046	31103	30711	30977	30586	30830
H ₂ O	34643	34500	30558	30199	30415	30056	30240
NH ₃	34396	34263	30291	29922	30157	29788	29925
2H ₂ O	33867	33699	29700	29378	29532	29210	29193
CH ₃ OH	34830	34695	30717	30428	30582	30293	30363
HCOOH	34505	34371	30368	30056	30235	29922	29816
NH ₃ -H ₂ O-H ₂ O	33381	33218	29171	28863	29008	28701	28340
NH ₃ -H ₂ O-NH ₃	33542	33385	29355	29036	29197	28879	28694
NH ₃ -NH ₃ -H ₂ O	33302	33136	29088	28812	28922	28646	28348
$\Delta\omega_{\pi \rightarrow \pi^*}$							
H ₂ O	-528	-546	-544	-512	-562	-530	-590
NH ₃	-775	-783	-812	-789	-820	-797	-905
2H ₂ O	-1304	-1347	-1403	-1333	-1446	-1376	-1637
CH ₃ OH	-341	-351	-386	-283	-396	-292	-467
HCOOH	-666	-675	-734	-655	-743	-664	-1014
NH ₃ -H ₂ O-H ₂ O	-1790	-1828	-1932	-1847	-1969	-1885	-2490
NH ₃ -H ₂ O-NH ₃	-1629	-1661	-1748	-1675	-1780	-1707	-2136
NH ₃ -NH ₃ -H ₂ O	-1869	-1910	-2014	-1899	-2055	-1940	-2482

^a Defined by Eq. (3.99), in which variant A of CR-EOMCC(2,3) is employed. ^b Defined by Eq. (3.99), in which variant D of CR-EOMCC(2,3) is employed. ^c Obtained with the laser resonant two-photon ionization UV spectroscopy [276].

Ideally, one would like to perform the δ -CR-EOMCC(2,3) calculations for basis sets larger than 6-31+G(d), such as 6-311+G(d), but complexes of cis-7HQ examined in this study were too large for performing such calculations on our computers. Thus, in the absence of the δ -CR-EOMCC(2,3) larger basis set data and considering the fact that the triples corrections to the environment-induced shifts $\Delta\omega_{\pi\rightarrow\pi^*}$ are relatively small when compared to the EOMCCSD $\Delta\omega_{\pi\rightarrow\pi^*}$ values, we have decided to combine the EOMCCSD/6-311+G(d) results with the triples corrections to EOMCCSD energies extracted from the δ -CR-EOMCC(2,3)/6-31+G(d) calculations. The final EOMCC values of the vertical excitation energies $\omega_{\pi\rightarrow\pi^*}$ were obtained using a composite approach, in which we augment the EOMCCSD/6-311+G(d) results by the triples corrections to EOMCCSD energies extracted from the δ -CR-EOMCC(2,3)/6-31+G(d) calculations, as in the following formula:

$$\begin{aligned}\omega_{\pi\rightarrow\pi^*}(\text{EOMCC}) &= \omega_{\pi\rightarrow\pi^*}(\text{EOMCCSD}/6\text{-}311\text{+G(d)}) \\ &+ [\omega_{\pi\rightarrow\pi^*}(\delta\text{-CR-EOMCC}(2,3)/6\text{-}31\text{+G(d)}) \\ &- \omega_{\pi\rightarrow\pi^*}(\text{EOMCCSD}/6\text{-}31\text{+G(d)})].\end{aligned}\tag{3.99}$$

As shown in Table (3.15), the resulting composite EOMCC,A and EOMCC,D approaches provide vertical excitation energies $\omega_{\pi\rightarrow\pi^*}$ that are in excellent agreement with the experimental excitation energies, while offering further improvements in the environment-induced shifts $\Delta\omega_{\pi\rightarrow\pi^*}$ when compared with the EOMCCSD/6-311+G(d) and δ -CR-EOMCC(2,3)/6-31+G(d) calculations. Indeed, the EOMCC,A approach, which adds the triples correction extracted from the δ -CR-EOMCC(2,3),A/ 6-31+G(d) calculation to the EOMCCSD/6-311+G(d) energy, gives errors in the calculated excitation energies $\omega_{\pi\rightarrow\pi^*}$ relative to experiment that range between 147 cm^{-1} in the case of the bare cis-7HQ system and 668

cm^{-1} in the case of the $7\text{HQ}\cdots(\text{NH}_3\text{-H}_2\text{O-H}_2\text{O})$ complex, never exceeding 2 % of the experimental excitation energies. The EOMCC,D approach, which adds the triples correction obtained in the $\delta\text{-CR-EOMCC}(2,3)\text{,D}/6\text{-31+G(d)}$ calculation to the EOMCCSD/6-311+G(d) energy, gives errors in the calculated excitation energies $\omega_{\pi\rightarrow\pi^*}$ relative to experiment that range between 17 cm^{-1} in the case of the $7\text{HQ}\cdots(\text{H}_2\text{O})_2$ complex and 361 cm^{-1} for $7\text{HQ}\cdots(\text{NH}_3\text{-H}_2\text{O-H}_2\text{O})$, or no more than 1 % of the experimental values. These results should be compared to the much larger differences between the EOMCCSD/6-311+G(d) and experimental excitation energies that range between 14 and 17 %. The complexation-induced spectral shifts $\Delta\omega_{\pi\rightarrow\pi^*}$ resulting from the EOMCC,A and EOMCC,D calculations agree with their experimental counterparts to within 5–27 % or 15 % on average in the case of EOMCC,A and 10–37 % or 22 % on average in the EOMCC,D case. The EOMCC,D approach, while bringing the excitation energies $\omega_{\pi\rightarrow\pi^*}$ to a closer agreement with experiment than the EOMCCSD/6-311+G(d) calculations, does not offer improvements in the calculated shifts $\Delta\omega_{\pi\rightarrow\pi^*}$. The composite EOMCC,A approach provides additional small improvements in the calculated $\Delta\omega_{\pi\rightarrow\pi^*}$ values, reducing the 7–33 % errors relative to experiment obtained in the EOMCCSD/6-311+G(d) calculations to 5–27 %. The EOMCC,A data, which was generated in this study, served as a computational benchmark for the TDDFT and FDET results and it was successfully shown that the non-relaxed FDET model represents a reasonable alternative to EOMCC methods for calculating environment-induced spectral shifts of excitation energies, but with a much more favorable computational scaling than the EOMCC-type approaches. The non-relaxed FDET results for the $\pi \rightarrow \pi^*$ shifts reported in Refs. [201] agree with our best EOMCC,A and EOMCC,D data to within about 100 cm^{-1} on average. This should be contrasted with the poor performance of the

supermolecular TDDFT approach that gives average errors in the calculated $\Delta\omega_{\pi\rightarrow\pi^*}$ values relative to EOMCC,A or EOMCC,D on the order of 700 cm^{-1} [201]. In the next section the performance of the EA- and IP-EOMCC methods for open-shell systems is considered.

3.2.6 Geometries and Adiabatic Excitation Energies of CNC, C₂N, N₃, and NCO

The accurate determination of geometries and energetics of ground- and excited-state radicals is very difficult, both experimentally, as open-shell species are usually characterized by short lifetimes, and computationally, due to the complexity of the many-electron correlation problem in open-shell systems. This is particularly true for the SR CC and EOMCC methods. Among the main reasons is the fact that the low-lying excited-states of radicals and other open-shell species are often dominated by two-electron and other multi-electron transitions. The goal of this section is the examination of the performance of the full and active-space EA- and IP-EOMCC approaches with up to $3p-2h/3h-2p$ excitations in the calculations of the ground and low-lying excited states of the challenging open-shell CNC, C₂N, NCO, and N₃ molecules [132, 289], as well as the prediction of the corresponding nuclear geometries and adiabatic excitation energies, which was carried out jointly with Mr. Jared Hansen from our group [290]. A series of EA-EOMCCSD($2p-1h$), active-space EA-EOMCCSD($3p-2h$){4}, and full EA-EOMCCSD($3p-2h$) calculations were performed for the CNC and C₂N molecules using the CCSD ground states of the CNC⁺ and C₂N⁺ closed-shell cations to provide the reference wave functions. In addition to the DZP basis, [291] which was employed in a previous study, [132, 289] the cc-pVDZ (all three methods), cc-pVTZ (all three methods), and cc-pVQZ [the EA-EOMCCSD($2p-1h$) and EA-EOMCCSD($3p-2h$){4} approaches] ba-

sis sets were used. The full EA-EOMCCSD($3p-2h$) calculations using the largest cc-pVQZ basis set could not be performed due to the large computer costs of the relevant numerical geometry optimizations [the analytic gradients of the EA-EOMCCSD($3p-2h$) approach are not available]. The active orbital spaces for the EA-EOMCCSD($3p-2h$){4} calculations for CNC and C₂N consisted of the two lowest-energy pairs of unoccupied π molecular orbitals of CNC⁺ and C₂N⁺, respectively. A series of IP-EOMCCSD($2h-1p$), active-space IP-EOMCCSD($3h-2p$){2}, and full IP-EOMCCSD($3h-2p$) calculations were performed for the NCO and N₃ molecules using the CCSD ground states of the NCO⁻ and N₃⁻ closed-shell anions to provide the reference wave functions. Again, in addition to the previously employed [132, 289] DZP basis, the cc-pVDZ (all three methods), cc-pVTZ (all three methods), and cc-pVQZ [the IP-EOMCCSD($2h-1p$) and IP-EOMCCSD($3h-2p$){2} approaches] basis sets were used, and the full IP-EOMCCSD($3h-2p$)/cc-pVQZ calculations were not performed due to the large computer costs of the corresponding numerical geometry optimizations [the analytic gradients of the IP-EOMCCSD($3h-2p$) method are not available]. The active orbital spaces for the IP-EOMCCSD($3h-2p$){2} calculations for NCO and N₃ consisted of the highest-energy pair of occupied orbitals of NCO⁻ and N₃⁻, respectively.

Unlike in earlier work [132, 289], where the nuclear geometries of the ground and excited states of the CNC, C₂N, NCO, and N₃ species were optimized using only one method (SAC-CI-SDT-*R*/PS [292–294]) and one basis set (DZP), in each molecular case and for each electronic state and basis set considered in the present work, the nuclear geometries were optimized at the same level of the EA/IP-EOMCC theory and with the same basis set as those used to evaluate the corresponding total and adiabatic excitation energies. The geometry optimizations using the cc-pVXZ ($X = D, T, \text{ and } Q$) basis sets were constrained

to linear geometries, since the analogous unconstrained optimizations using the DZP basis set and bent initial structures showed that the optimum geometries of the calculated states of CNC, C₂N, NCO, and N₃ are linear. The unconstrained optimizations with the DZP basis set demonstrated that we can assume the $D_{\infty h}$ (in practice, D_{2h}) symmetry for each of the calculated states of CNC and N₃, and that we can use the $C_{\infty v}$ (in practice, C_{2v}) symmetry in the geometry optimizations for C₂N and NCO. In all post-RHF (CCSD and EA/IP-EOMCC) calculations, the lowest-energy core orbitals correlating with the 1s orbitals of the C and N atoms were kept frozen and the spherical components of the d , f , and g functions were employed throughout. In addition to the results of the finite basis set calculations, the total and excitation energies obtained in the EA- and IP-EOMCC calculations for the CNC, C₂N, NCO, and N₃ molecules were extrapolated to the CBS limit. The CBS extrapolations were limited to the EA-EOMCCSD($2p-1h$), IP-EOMCCSD($2h-1p$), active-space EA-EOMCCSD($3p-2h$){4}, and active-space IP-EOMCCSD($3h-2p$){2} calculations, since full EA-EOMCCSD($3p-2h$) and IP-EOMCCSD($3h-2p$) computations including geometry optimizations were too expensive to be run using the cc-pVQZ basis set that would be required to obtain reliable CBS-limit values. In all of the remaining cases, the complete data sets corresponding to the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets were available, enabling reasonably meaningful CBS extrapolations. To verify the numerical stability of our CBS extrapolations, two different extrapolation schemes, referred to as the CBS-A and CBS-B approaches, were utilized.

In the CBS-A scheme, the CBS limit of the ground-state CCSD correlation energy of the closed-shell N -electron reference system relevant to the EA/IP-EOMCC calculations for the $(N \pm 1)$ -electron target species was determined using the well-known extrapolation

formula [217] given in Eq. (3.98) and the cc-pVTZ and cc-pVQZ data. Here, $\Delta E(X)$ in Eq. (3.98) is the CCSD correlation energy obtained with the cc-pVXZ basis set, where X represents the cardinal number of the basis set ($X = 3$ for cc-pVTZ and $X = 4$ for cc-pVQZ), and ΔE_∞ is the CCSD correlation energy in the CBS limit. The resulting extrapolated CBS correlation energy ΔE_∞ was then added to the RHF/cc-pV6Z energy of the N -electron reference system, computed at the optimized geometry of the state of interest resulting from the appropriate EA- or IP-EOMCC/cc-pVQZ calculations. Recalling the well-known and previously discussed fact that the RHF energies converge exponentially with the basis set, it is usually best to simply perform the calculations at the RHF level with a very large correlation consistent basis set, such as cc-pV6Z, if such calculations are affordable. As a check, the level of basis-set convergence was verified by comparing the RHF/cc-pV5Z and RHF/cc-pV6Z data, obtaining differences of about 0.2 millihartree in all of the examined cases. Once the CBS values of the RHF total and CCSD correlation energies of the N -electron reference system were determined, the desired CBS limits of the ground- and excited-state energies of the $(N \pm 1)$ -electron target species corresponding to the EA- or IP-EOMCC calculations of interest were computed using the formula

$$\begin{aligned}
E_{\mu,\infty}^{\text{EA/IP-EOMCC}}(N \pm 1) &= E_{6Z}^{\text{RHF}}(N) + \Delta E_{0,\infty}^{\text{CCSD}}(N) \\
&\quad + E_{\mu,\text{QZ}}^{\text{EA/IP-EOMCC}}(N \pm 1) - E_{0,\text{QZ}}^{\text{CCSD}}(N), \quad (3.100)
\end{aligned}$$

where $E_{\mu,\infty}^{\text{EA/IP-EOMCC}}(N \pm 1)$ is the final extrapolated energy of the $(N \pm 1)$ -electron state $|\Psi_\mu^{(N \pm 1)}\rangle$, $E_{6Z}^{\text{RHF}}(N)$ is the ground-state RHF energy of the closed-shell N -electron reference system obtained with the cc-pV6Z basis set, $\Delta E_{0,\infty}^{\text{CCSD}}(N)$ is the extrapolated CCSD correlation energy of the N -electron reference system obtained using Eq. (3.98),

$E_{\mu, \text{QZ}}^{\text{EA/IP-EOMCC}}(N \pm 1)$ is the total EA/IP-EOMCC energy of the $(N \pm 1)$ -electron state $|\Psi_{\mu}^{(N \pm 1)}\rangle$ obtained with the cc-pVQZ basis set, and $E_{0, \text{QZ}}^{\text{CCSD}}(N)$ is the total CCSD energy of the N -electron reference system obtained using the cc-pVQZ basis set. This method of estimating the CBS values of the total electronic energies of the ground and excited states of the CNC, C₂N, NCO, and N₃ radicals is based on the assumption that the electron-attachment or ionization energies $\omega_{\mu}^{(N \pm 1)}$ obtained with the cc-pVQZ basis set are essentially converged with respect to the basis set, so all one has to do is obtain the CBS limit of the CCSD ground-state energy of the N -electron reference system and add the cc-pVQZ values of the electron-attachment or ionization energies to estimate the CBS energies of the ground and excited states of the corresponding $(N \pm 1)$ -electron target species.

In the second basis set extrapolation method, referred to as the CBS-B approach, the total CBS energy of each $(N \pm 1)$ -electron target state of interest was directly determined using the formula [212]

$$E(X) = E_{\infty} + Be^{-(X-1)} + Ce^{-(X-1)^2}, \quad (3.101)$$

and the cc-pVDZ, cc-pVTZ, and cc-pVQZ data. As in Eq. (3.98), the X variable number entering Eq. (3.101) is again the cardinal number of the cc-pVXZ basis set ($X = 2$ for cc-pVDZ, $X = 3$ for cc-pVTZ, and $X = 4$ for cc-pVQZ), $E(X)$ is the total EA/IP-EOMCC energy computed with the cc-pVXZ basis set, and E_{∞} is the desired CBS limit of the total EA/IP-EOMCC energy for a given electronic state of the $(N \pm 1)$ -electron species. The difference between the CBS-A and CBS-B extrapolation schemes lies in the fact that the latter scheme extrapolates the total EA- or IP-EOMCC energy of each electronic state of the $(N \pm 1)$ -electron target species separately, using Eq. (3.101), whereas the former approach

extrapolates the ground-state correlation energy of the N -electron reference system only using Eq. (3.98) while making an assumption that the electron-attachment and ionization energies resulting from the EA- and IP-EOMCC calculations converge faster with the basis set than the total energies of the $(N \pm 1)$ -electron target species, as reflected in Eq. (3.100).

The results of our EA- and IP-EOMCC calculations, along with the available experimental data [295–297], are reported in Tables (3.16)-(3.21). The EA-EOM-CCSD($2p-1h$), EA-EOMCCSD($3p-2h$){4}, and full EA-EOMCCSD($3p-2h$) results for the CNC and C₂N molecules are reported in Tables (3.16) and (3.17) for the total and adiabatic excitation energies, and (3.18) for the geometries. The IP-EOMCCSD($2h-1p$), IP-EOMCCSD($3h-2p$){2}, and full IP-EOMCCSD($3h-2p$) results for the NCO and N₃ molecules are reported in Tables (3.19) and (3.20) for the total and adiabatic excitation energies, and (3.21) for the geometries. Our discussion is divided into two parts.

We begin with the CNC molecule (see Tables (3.16) and (3.18)). For CNC, the EA-EOMCC optimizations employing the DZP basis set produced results that deviate from the previously reported [289] SAC-CI-SDT- R /PS optimized geometries and previously calculated [132, 289] EA-EOMCC adiabatic excitation energies using the SAC-CI-SDT- R /PS geometries, all obtained with the same DZP basis, by 0.001-0.009 Å and 0.002-0.003 eV, respectively, for all states and all methods considered here. Seeing that our present EA-EOMCC optimizations employing the DZP basis set were able to reproduce the analogous results of the SAC-CI-SDT- R /PS geometry optimizations and the EA-EOMCC excitation energies at the SAC-CI-SDT- R /PS geometries reported in Ref. [289], the effect of the use of the correlation-consistent basis sets of the cc-pVXZ quality on the calculated excitation energies and geometries was investigated.

The ground $X^2\Pi_g$ state of CNC is dominated by $1p$ excitations out of the ground state of the closed-shell reference CNC^+ ion but, the $A^2\Delta_u$ and $B^2\Sigma_u^+$ excited states of the same molecule exhibit a significant two-electron excitation character relative to the $X^2\Pi_g$ state. As shown in Table (3.16), the basic EA-EOM-CCSD($2p-1h$) optimizations produced adiabatic excitation energies that deviate from the experimental values by 3.379-4.022 eV for the $A^2\Delta_u$ and $B^2\Sigma_u^+$ states, demonstrating the same characteristically large errors compared to experiment that are typically seen when the EA-EOMCCSD($2p-1h$) approach is applied to the excited states of radicals dominated by two-electron transitions [147–149, 298]. The full EA-EOMCCSD($3p-2h$) method improves these poor results, reducing the deviations from experiment to 0.336-0.451 eV for both the $A^2\Delta_u$ and $B^2\Sigma_u^+$ states, when the cc-pVDZ and cc-pVTZ basis sets are employed. The reason for this considerable improvement in the data over the EA-EOMCCSD($2p-1h$) method is the explicit inclusion of the $3p-2h$ terms in the $R_\mu^{(N+1)}$ operator in the EA-EOMCCSD($3p-2h$) calculations.

The inclusion of all $3p-2h$ components in the $R_\mu^{(N+1)}$ operator is computationally demanding, particularly when one is interested in numerical gradient optimizations, such as those performed in this work. Thus, it is of great significance to note that the active-space EA-EOMCCSD($3p-2h$){4} optimizations using the cc-pVDZ and cc-pVTZ basis sets, with only four unoccupied orbitals in the active-space, which are only a few times more expensive than the corresponding ground-state CCSD calculations and which require a small fraction of the CPU time, disk, and memory when compared to the parent EA-EOMCCSD($3p-2h$) calculations, reproduce the full EA-EOMCCSD($3p-2h$) optimization results to within 0.019-0.039 eV for the adiabatic excitation energies and 0.8-3.2 millihartree for the total energies of the ground and excited states of CNC examined in this study (see Table (3.16)). The

deviations of the EA-EOMCCSD($3p-2h$){4} results from experiment are 0.312-0.422 eV for all three correlation consistent basis sets used in the EA-EOMCCSD($3p-2h$){4} optimizations. It is interesting and somewhat surprising to note the deviations from experiment increase slightly with the size of the cc-pVXZ basis set for all three EA-EOMCC methods exploited in this work. To help understand this behavior, the EA-EOMCCSD($2p-1h$) and EA-EOMCCSD($3p-2h$){4} results were extrapolated to the CBS limit.

Examining the total energies of the ground and excited states of CNC shown in Table (3.16), it is clear that they are converging with the basis set in a systematic manner. The CBS-A extrapolation scheme is based on the simplifying assumption that the electron-attachment (or, in the IP case, ionization) energies are reasonably well converged with the basis set, when the cc-pVQZ basis set is employed. The data in Table (3.16) show that this is indeed a valid assumption, as the EA-EOMCCSD($2p-1h$) and EA-EOMCCSD($3p-2h$){4} excitation energies do not significantly change when moving from the cc-pVTZ to cc-pVQZ basis, the largest change being 0.109 eV for the less accurate EA-EOMCCSD($2p-1h$) method and only 0.003 eV for the active-space EA-EOMCCSD($3p-2h$){4} approach. Moreover, the CBS extrapolations resulting from the CBS-A and CBS-B schemes produce results that are in good agreement with each other, especially for the higher-order EA-EOMCCSD($3p-2h$){4} method, where the differences in total energies do not exceed 1.3 millihartree, regardless of the electronic state of CNC considered (recall that the CBS-B scheme extrapolates the total energy of each state separately, without any simplifying assumptions). The CBS limits of the excitation energies resulting from the CBS-A and CBS-B extrapolations obtained with the EA-EOMCCSD($3p-2h$){4} approach are essentially identical, deviating by 0.001 eV and 0.005 eV for the $A^2\Delta_u$ and $B^2\Sigma_u^+$ states, respectively. Thus, it may be safely concluded that

the CBS EA-EOMCCSD($3p-2h$){4} results are stable to approximately 1 millihartree for the total energies and 0.005 eV for the adiabatic excitation energies, and can be regarded as converged with the basis set. This suggests that the 0.3–0.4 eV error relative to experiment resulting from the full and active-space EA-EOMCCSD($3p-2h$) calculations for CNC are due to either $4p-3h$ excitations neglected in these calculations or errors in the experimental data.

In analogy to CNC, the ground $X^2\Pi$ state of C_2N is dominated by $1p$ excitations out of the ground state of the closed-shell reference C_2N^+ ion, but the low-lying $A^2\Delta$, $B^2\Sigma^-$, and $C^2\Sigma^+$ excited states have significant $2p-1h$ contributions demonstrating the rather complex MR nature of their corresponding wave functions. The $B^2\Sigma^-$ state also has non-negligible $3p-2h$ contributions, which make this state extremely difficult to describe by the methods used in this study. All of this causes major problems in the EA-EOMCCSD($2p-1h$) calculations. As shown in Table (3.17), even with a large cc-pVQZ basis set, the EA-EOMCCSD($2p-1h$) method incorrectly orders the excited states of C_2N , describing the $C^2\Sigma^+$ state as being lower in energy than the $B^2\Sigma^-$ state. The errors in the EA-EOMCCSD($2p-1h$) results for the adiabatic excitation energies of C_2N relative to experiment are huge. Indeed, our geometry optimizations using the EA-EOMCCSD($2p-1h$) approach produce errors in the calculated adiabatic excitation energies of C_2N relative to experiment of 3.507–3.969 eV for the $A^2\Delta$ state, 4.956–5.511 eV for the $B^2\Sigma^-$ state, and 3.422–3.836 eV for the $C^2\Sigma^+$ state. As with CNC, the full inclusion of the $3p-2h$ components in the electron attaching operator R_μ^{N+1} significantly improves the adiabatic excitation energies relative to the disastrous EA-EOMCCSD($2p-1h$) results, reducing the errors relative to experiment to at most 0.418 eV for the $A^2\Delta$ state, at most 0.915 eV for the $B^2\Sigma^-$ state, and at most 0.524 eV for the $C^2\Sigma^+$ state when the cc-pVDZ and cc-pVTZ basis sets are employed, but the full EA-

EOMCCSD($3p-2h$) are computationally demanding, particularly when larger basis sets have to be examined. Thus, it is important to examine how well the considerably less expensive active-space EA-EOMCCSD($3p-2h$) approach works for the low-lying excited states of C_2N , when the cc-pV x Z basis sets are employed.

As shown in Table (3.17), the results of the active-space EA-EOMCCSD($3p-2h$){4} calculations are almost identical to those obtained with the parent EA-EOM-CCSD($3p-2h$) approach. The adiabatic excitation energies obtained with the full and active-space EA-EOMCCSD($3p-2h$) methods, where the latter approach uses only four unoccupied orbitals in the active-space, calculated using the cc-pVDZ and cc-pVTZ basis sets, differ by 0.023-0.053 eV for all states of C_2N examined here. The total energies obtained in the full EA-EOMCCSD($3p-2h$) and active-space EA-EOMCCSD($3p-2h$){4} calculations employing the cc-pVDZ and cc-pVTZ basis sets differ by 1.6-2.9 millihartree for the $X^2\Pi$ state, 0.5-1.1 millihartree for the $A^2\Delta$ state, 0.5-0.9 millihartree for the $B^2\Sigma^-$ state, and 0.7-1.4 millihartree for the $C^2\Sigma^+$ state. Comparing the EA-EOMCCSD($3p-2h$){4} results with experiment, it can be seen that the adiabatic excitation energies resulting from the EA-EOMCCSD($3p-2h$){4} calculations using the cc-pVQZ basis set differ from the available experimental data by 0.340 eV, 0.844 eV, and 0.454 eV for the $A^2\Delta$, $B^2\Sigma^-$, and $C^2\Sigma^+$ states, respectively, which is a huge error reduction when compared to the corresponding EA-EOMCCSD($2p-1h$)/cc-pVQZ calculations that give the 3.969 eV, 5.511 eV, and 3.836 eV errors for the same three states, in addition to wrong state ordering. The full EA-EOMCCSD($3p-2h$) and active-space EA-EOMCCSD($3p-2h$){4} calculations produce the correct state ordering and relatively small errors for the $A^2\Delta$ and $C^2\Sigma^+$ states, but the discrepancy between the full and active-space EA-EOMCCSD($3p-2h$) results on the one hand and experiment on the other hand for the

$B^2\Sigma^-$ state, on the order of 0.9 eV independent of the basis set, is a problem that needs to be addressed.

The larger deviations with experiment observed in the full EA-EOMCCSD($3p-2h$) and active-space EA-EOMCCSD($3p-2h$){4} calculations for the $B^2\Sigma^-$ state, which do not seem to be decreasing with the basis set and careful geometry optimizations performed in this work, must be related to the presence of the non-negligible $3p-2h$ contributions in the $B^2\Sigma^-$ wave function, which indicate a highly MR character of this state that the EA-EOMCC methods used in the present study cannot capture without incorporating higher-than- $3p-2h$ contributions in the EA-EOMCC considerations. As explained in Ref. [298], the presence of significant $3p-2h$ contributions in the wave function requires an explicit consideration of the $4p-3h$ and, perhaps, higher-than- $4p-3h$ components of the $R_\mu^{(N+1)}$ operator in the EA-EOMCC calculations, neglected at the EA-EOMCCSD($3p-2h$) level. The highly MR character of the $B^2\Sigma^-$ state becomes clear when we examine the CASPT2 and CASSCF-based MRCI calculations reported in Ref. [299]. These calculations are in reasonable agreement with the results of our full and active-space EA-EOMCCSD($3p-2h$) calculations for the $A^2\Delta$ and $C^2\Sigma^+$ states, producing a 0.238 eV error for the $A^2\Delta$ state and 0.219 eV error for the $C^2\Sigma^+$ state when the CASPT2 approach is employed, but the CASPT2 and MRCI results obtained in Ref. [299] for the $B^2\Sigma^-$ state are considerably more accurate than those obtained here with the EA-EOMCCSD($3p-2h$) theory levels. Indeed, the CASPT2 and MRCI calculations for the $B^2\Sigma^-$ state reported in Ref. [299] give errors of 0.225 eV and 0.250 eV, respectively, relative to experiment, as opposed to ~ 0.9 eV obtained with the full EA-EOMCCSD($3p-2h$) and active-space EA-EOMCCSD($3p-2h$){4} methods. Interestingly, the MRCI approach improves the CASPT2 results for the $A^2\Delta$ and $C^2\Sigma^+$ states as well, reduc-

ing the 0.238 eV and 0.219 eV errors obtained in the CASPT2 calculations to 0.060 eV and 0.058 eV, respectively [299], which suggests that the incorporation of the $4p-3h$ and, perhaps, some other higher-order excitations in the EA-EOMCC calculations may be necessary to further improve the description of all three excited states of C_2N examined in this work. Since the calculations reported in the present paper exclude the possibility that the basis set or geometry optimizations may help the EA-EOMCCSD($3p-2h$) results, the next logical step is to examine the role of $4p-3h$ excitations in the EA-EOMCC calculations. One may also have to examine whether the use of the full CCSDT approach rather than the CCSD method in providing the ground-state wave function for the reference C_2N^+ ion plays a role here. These will be the topics of our group’s future work.

As shown in Table (3.17), the adiabatic excitation energies resulting from the EA-EOMCCSD($2p-1h$) and EA-EOMCCSD($3p-2h$){4} calculations with the cc-pVQZ basis set are reasonably well converged with the basis and, although the total energies of the individual electronic states of C_2N are not converged when the cc-pVQZ basis set is employed, they behave in a systematic manner as we go from the cc-pVDZ to cc-pVQZ basis sets, facilitating the CBS extrapolations. Indeed, when going from the cc-pVTZ to the cc-pVQZ basis sets, the changes in the EA-EOMCCSD($3p-2h$){4} excitation energies are very small, at most 0.018 eV. The analogous changes in the EA-EOMCCSD($2p-1h$) excitation energies are somewhat larger (at most 0.120 eV), but can still be viewed as reasonably stable considering the complicated nature of the C_2N excited states that the EA-EOMCCSD($2p-1h$) approach has significant problems with. Overall, the simplifying assumption of the CBS-A extrapolation scheme that one can treat the electron-attachment energies resulting from the EA-EOMCC calculations with the cc-pVQZ basis set as essentially converged values remains valid for

C_2N , so we expect the CBS-A scheme to provide meaningful results. This can be verified by comparing the CBS-A and CBS-B extrapolations. Comparing the EA-EOMCCSD($2p1h$) CBS-A and CBS-B values, the total energies differ by 1.9 millihartree for the $X^2\Pi$ state, 0.1 millihartree for the $A^2\Delta$ state, 0.5 millihartree for the $B^2\Sigma^-$ state, and 0.2 millihartree for the $C^2\Sigma^+$ state. The adiabatic excitation energies resulting from the CBS-A and CBS-B extrapolations of the EA-EOMCCSD($2p1h$) data differ by 0.054 eV for the $A^2\Delta$ state, 0.066 eV for the $B^2\Sigma^-$ state, and 0.047 eV for the $C^2\Sigma^+$ state. The CBS-A and CBS-B results for the EA-EOMCCSD($3p-2h$){4} total energies differ by 6.0 millihartree for the $X^2\Pi$ state, 7.5 millihartree for the $A^2\Delta$ state, 8.2 millihartree for the $B^2\Sigma^-$ state, and 7.6 millihartree for the $C^2\Sigma^+$ state. The differences in the adiabatic excitation energies obtained with the two CBS extrapolation schemes, as applied to the EA-EOMCCSD($3p-2h$){4} data, are 0.042 eV, 0.062 eV, and 0.045 eV for the $A^2\Delta$, $B^2\Sigma^-$ and $C^2\Sigma^+$ states, respectively. We can conclude that our CBS EA-EOMCC results for the C_2N molecule are generally stable to within about 8 millihartree for the total energies and 0.060 eV for the adiabatic excitation energies. The deviations of the CBS-A extrapolated EA-EOMCCSD($3p-2h$){4} results from experiment are 0.371 eV for the $A^2\Delta$ state, 0.893 eV for the $B^2\Sigma^-$ state, and 0.491 eV for the $C^2\Sigma^+$ state. The analogous CBS-B calculations employing the EA-EOMCCSD($3p-2h$){4} data give errors of 0.329 eV, 0.831 eV, and 0.446 eV, respectively. These results indicate once again that higher-than- $3p-2h$ excitations and, perhaps, methods better than CCSD for the description of the ground state of the reference C_2N^+ ion may have to be included in the EA-EOMCC calculations for the low-lying states of the C_2N molecule, particularly in the case of the $B^2\Sigma^-$ state.

Having demonstrated the significance of higher than $2p-1h$ contributions for an accurate

description of the excitation energies in the CNC and C₂N molecules and having established the ability of the active-space EA-EOMCCSD($3p-2h$) approach to capture the most significant $3p-2h$ contributions with only a few active orbitals, independently of the basis set, we turn now to the effectiveness of the EA-EOMCC schemes in describing the equilibrium geometries of the ground and excited states of CNC and C₂N. As seen in Table (3.18), the EA-EOMCCSD($2p-1h$) level of theory gives the C–N bond lengths in CNC designated as $R_{\text{C-N}}$, which deviate from the corresponding experimental values by 0.003-0.015 Å for the $X^2\Pi_g$ state, 0.004-0.009 Å for the $A^2\Delta_u$ state, and 0.001-0.016 Å for the $B^2\Sigma_u^+$ state, when the cc-pVXZ basis sets with $X = \text{D, T, and Q}$ are employed. The full EA-EOMCCSD($3p-2h$) approach employing the cc-pVDZ and cc-pVTZ basis sets produces $R_{\text{C-N}}$ values that deviate from experiment by 0.001-0.017 Å, 0.003-0.012 Å, and 0.004-0.011 Å for the $X^2\Pi_g$, $A^2\Delta_u$, and $B^2\Sigma_u^+$ states, respectively, i.e., results that are of equally high quality and not much different than the low-order EA-EOMCCSD($2p-1h$) data. The analogous active-space EA-EOMCCSD($3p-2h$){4} calculations give $R_{\text{C-N}}$ bond lengths that differ from experiment by 0.001-0.017 Å in the case of the $X^2\Pi_g$ state, 0.003-0.012 Å in the case of the $A^2\Delta_u$ state, and 0.005-0.014 Å when the $B^2\Sigma_u^+$ state is examined. All of this shows that not only is the EA-EOMCCSD($3p-2h$){4} approach able to reproduce the more computationally demanding EA-EOMCCSD($3p-2h$) results for the nuclear geometries of the low-lying states of CNC, but that the high-level EA-EOMCCSD($3p-2h$) values of $R_{\text{C-N}}$ and those obtained with the inexpensive EA-EOMCCSD($2p-1h$) method differ only by 0.002-0.004 Å for the $X^2\Pi_g$ state and 0.001-0.003 Å for the $A^2\Delta_u$ and $B^2\Sigma_u^+$ states, at least when the cc-pVDZ and cc-pVTZ basis sets are employed. The active-space EA-EOMCCSD($3p-2h$){4} approach and the EA-EOMCCSD($2p-1h$) method give $R_{\text{C-N}}$ values that differ by at most 0.004 Å for

all states of CNC and all basis sets examined in this work, confirming the observation that it is sufficient to use the low-level EA-EOMCCSD($2p-1h$) approach to obtain an accurate description of the equilibrium geometries of the low-lying states of CNC.

Similar, but not entirely identical, remarks apply to the C_2N molecule. As shown in Table (3.18), the EA-EOMCCSD($2p-1h$) and EA-EOMCCSD($3p-2h$){4} approaches employing the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets give C–C and C–N bond lengths, R_{C-C} and R_{C-N} , respectively, that differ by at most 0.065 Å when we compare the EA-EOMCCSD($2p-1h$) and the corresponding EA-EOMCCSD($3p-2h$){4} data for all electronic states of C_2N examined in this work, mostly because of the inability of the EA-EOMCCSD($2p-1h$) approach to provide a highly accurate description of the excited-state geometries [the differences between the EA-EOMCCSD($2p-1h$) and EA-EOMCCSD($3p-2h$){4} geometries of the C_2N ’s ground state are less than 0.004 Å]. On the other hand, the differences between the active-space EA-EOM-CCSD($3p-2h$){4} and full EA-EOMCCSD($3p-2h$) values of R_{C-C} and R_{C-N} obtained with the cc-pVDZ and cc-pVTZ basis sets do not exceed 0.001 Å, confirming our earlier remarks about the ability of the active-space EA-EOMCCSD($3p-2h$) approach to capture essentially all correlation effects that are included in the full EA-EOMCCSD($3p-2h$) calculations. We could not find any experimental data for the geometries of the ground and excited states of C_2N , so we cannot comment on the accuracy of our R_{C-C} and R_{C-N} values resulting from the EA-EOMCC calculations in any definitive manner, but, judging by the high quality of the EA-EOMCC results for the geometries of the low-lying states of CNC, we can conclude that the geometries resulting from the full EA-EOMCCSD($3p-2h$) and active-space EA-EOMCCSD($3p-2h$){4} calculations using the cc-pVTZ or cc-pVQZ basis sets are of the similarly high quality. The low-level EA-EOMCCSD($2p-1h$) calculations

seem less accurate than in the CNC case, particularly when the excited states of C_2N are examined, but they are still in reasonable agreement with the high-level full and active-space EA-EOMCCSD($3p-2h$) results.

The above discussion provides us with an important insight about the performance of the EA-EOMCC methods. The EA-EOMCCSD($2p-1h$) approach, while generally inadequate for an accurate description of the excitation energies in open-shell systems, such as the CNC and C_2N molecules examined in this work, is capable of providing reasonably accurate equilibrium geometries, even for excited states that have a significant MR character. On the other hand, it seems to be generally safer to use the active-space EA-EOMCCSD($3p-2h$) approach in geometry optimizations, particularly since it provides results that are virtually identical to the corresponding full EA-EOMCCSD($3p-2h$) data, both for the excitation energies and nuclear geometries.

We now turn to the IP-EOMCC calculations for the NCO and N_3 molecules, which are summarized in Tables (3.19)–(3.21). For both molecules, the IP-EOMCC optimizations employing the DZP basis set, carried out in the present work, produced results that are very similar to the previously reported [132, 289] IP-EOMCC adiabatic excitation energies calculated at the SAC-CI-SDT- R /PS optimized geometries, all obtained with the same DZP basis as that used here. For example, the deviations between the adiabatic excitation energies of NCO and N_3 obtained in the present IP-EOMCC/DZP optimizations and the analogous excitation energies reported in Refs. [132, 289], which used the geometries obtained with the SAC-CI-SDT- R /PS approach, are 0.001-0.078 eV for all states and all methods considered in this study. As in the case of the EA-EOMCC calculations, after seeing that our present IP-EOMCC optimizations for the ground and excited states of NCO and N_3 employing the

DZP basis set were able to reproduce the analogous results calculated at the geometries obtained in the SAC-CI-SDT-*R*/PS calculations, reported in Refs. [132, 289], we moved to the examination of the effect of the use of the correlation-consistent basis sets of the cc-pVXZ quality on the calculated excitation energies and geometries.

Unlike the CNC and C₂N molecules, which are characterized by the presence of low-lying excited states with a significant MR character in their respective electronic spectra, the low-lying states of NCO and N₃ have a predominantly $1h$ excitation character relative to the corresponding NCO⁻ and N₃⁻ reference ions, with only small contributions from higher-than- $1h$ excitations. As a result, it is much easier to describe the low-lying states of NCO and N₃ by the IP-EOMCC methods and already the basic IP-EOMCCSD($2h-1p$) approach performs quite well. For example, as shown in Table (3.19), the deviations from experiment for the adiabatic excitations in NCO resulting from the IP-EOMCCSD($2h-1p$) calculations are only 0.007-0.098 eV for the $A^2\Sigma^+$ state and 0.336-0.381 eV for the $B^2\Pi$ state when the cc-pVXZ basis sets with $X = D, T, \text{ and } Q$ are employed. Inclusion of higher-order ($3h-2p$) correlation effects through the full IP-EOMCCSD($3h-2p$) method offers additional improvements, reducing the overall deviations from experiment to 0.018-0.072 eV in the $A^2\Sigma^+$ case and 0.044-0.085 eV in the $B^2\Pi$ case, when the cc-pVDZ and cc-pVTZ basis sets are employed.

The inexpensive active-space variant of IP-EOMCCSD($3h-2p$) using only two active occupied orbitals, IP-EOMCCSD($3h-2p$){2}, yields similar excitation energy values to those from its more expensive parent scheme, with somewhat larger deviations from experiment of 0.207-0.312 eV for the $A^2\Sigma^+$ state and very small 0.063-0.139 eV deviations for the $B^2\Pi$ state, confirming that one can essentially use any IP-EOMCC approach and obtain a reason-

able description of the low-lying states of NCO, but the deviations between the results of the full and active-space IP-EOMCCSD($3h-2p$) calculations for NCO are somewhat larger than those observed in the EA-EOMCCSD($3p-2h$) computations for CNC and C₂N. This is particularly true for the $A^2\Sigma^+$ state, where they are 0.240 eV for the adiabatic excitation energy and 11.0 millihartree for the total energy when the cc-pVDZ basis set is employed and 0.218 eV and 10.7 millihartree when the cc-pVTZ basis set is used. As pointed out in previous work [289], these larger differences between the full and active-space IP-EOMCCSD($3h-2p$) results for the $A^2\Sigma^+$ state of NCO are likely due to the small active-space used in the latter calculations, which consists of only one pair of highest-energy occupied π orbitals of NCO⁻, and/or from changes in the character of molecular orbitals when going from the NCO⁻ reference ion to the NCO target species. On the other hand, the overall agreement between the full and active-space IP-EOMCCSD($3h-2p$) results for NCO is rather good. For example, the differences between the full and active-space IP-EOMCCSD($3h-2p$) results for the adiabatic excitation energies corresponding to the $B^2\Pi$ state are only 0.019 eV when the cc-pVDZ basis set is employed and 0.023 eV when the cc-pVTZ basis set is used. The differences between the total energies obtained in the full IP-EOMCCSD($3h-2p$) and active-space IP-EOMCCSD($3h-2p$){2} calculations for the $X^2\Pi$ and $B^2\Pi$ states range between 2.1 and 3.5 millihartree when the cc-pVDZ and cc-pVTZ basis set are used, which is an excellent agreement.

Many of the above observations remain valid when the IP-EOMCC methods are applied to N₃. As shown in Table (3.20), the adiabatic excitation energies corresponding to the $B^2\Sigma_u^+$ state obtained with the IP-EOMCCSD($2h-1p$) optimizations employing the cc-pVXZ basis sets with $X = D, T, \text{ and } Q$ differ from the corresponding experimental value by 0.056-0.110

eV. Again, as in the NCO case, the full IP-EOMCCSD($3h-2p$) approach reduces the already small errors in the IP-EOMCCSD($2h-1p$) results for the $B^2\Sigma_u^+$ state of N_3 to the even smaller 0.023-0.049 eV range. The much less expensive active-space IP-EOMCCSD($3h-2p$){2} calculations using the cc-pVXZ basis sets with $X = D, T,$ and Q produce errors of 0.174-0.200 eV, which are larger than those obtained with full IP-EOMCCSD($3h-2p$), but the general agreement between the full and active-space IP-EOMCCSD($3h-2p$) data is reasonable. Indeed, the total energies resulting from the full IP-EOMCCSD($3h-2p$) and active-space IP-EOMCCSD($3h-2p$){2} calculations differ by only 1.4-1.7 millihartree in the $X^2\Pi_g$ case and 6.9-7.3 millihartree in the case of the $B^2\Sigma_u^+$ state. The adiabatic excitation energies corresponding to the $B^2\Sigma_u^+$ state obtained in the full and active-space IP-EOMCCSD($3h-2p$) calculations with the cc-pVDZ and cc-pVTZ basis sets differ by 0.151 eV, which is a reasonable agreement. Again, the somewhat larger differences between the full and active-space IP-EOMCCSD($3h-2p$) data for the $B^2\Sigma_u^+$ state compared to the analogous EA-EOMCC calculations for CNC and C_2N are likely due to the reasons cited above for the NCO molecule.

We now turn our attention to the numerical stability of our IP-EOMCC results for the NCO and N_3 molecules in the CBS limit. As in the EA-EOMCC calculations for CNC and C_2N , the IP-EOMCC total energies of each state of NCO and N_3 shown in Tables (3.19) and (3.20) behave in a systematic manner, as we go from the cc-pVDZ to cc-pVQZ basis sets, showing the initial signs of convergence; and the excitation energies obtained in the IP-EOMCCSD($2h-1p$) and IP-EOMCCSD($3h-2p$){2} calculations with the cc-pVQZ basis set can be regarded as reasonably well converged, which helps the validity of the CBS-A extrapolations. Indeed, the differences between the adiabatic excitation energies calculated with the cc-pVTZ and cc-pVQZ basis sets at the IP-EOMCCSD($2h-1p$) and IP-EOMCCSD($3h-2p$){2}

levels of theory are 0.028-0.029 eV for the $A^2\Sigma^+$ state of NCO, 0.020-0.031 eV for the $B^2\Pi$ state of NCO, and 0.003 eV for the $B^2\Sigma_u^+$ state of N_3 . It is, therefore, not surprising that the CBS-A and CBS-B extrapolations for the ground and excited states of the NCO and N_3 molecules summarized in Tables (3.19) and (3.20) are in good agreement. Indeed, the CBS-A and CBS-B total energies obtained with the IP-EOMCCSD($2h-1p$) data for NCO differ by only 1.5 millihartree for the $X^2\Pi$ and $A^2\Sigma^+$ states and 2.7 millihartree for the $B^2\Pi$ state. The corresponding excitation energies resulting from both CBS extrapolations differ by 0.001 eV for the $A^2\Sigma^+$ state and 0.034 eV for the $B^2\Pi$ state. In consequence, the CBS-A- and CBS-B-extrapolated IP-EOMCCSD($2h-1p$) excitation energies obtained for NCO differ from the corresponding experimental values by 0.009-0.010 eV in the $A^2\Sigma^+$ case and 0.393-0.427 eV in the case of the $B^2\Pi$ state. Similar remarks apply to the IP-EOMCCSD($3h-2p$){2} approach, where the corresponding CBS-A- and CBS-B-extrapolated total energies differ by 2.5, 4.1, and 3.0 millihartree for the $X^2\Pi$, $A^2\Sigma^+$, and $B^2\Pi$ states, respectively, so that the differences in the resulting CBS-A and CBS-B IP-EOMCCSD($3h-2p$){2} excitation energies are 0.045 eV for the $A^2\Sigma^+$ state and 0.014 eV for the $B^2\Pi$ state. As a consequence, the CBS-A- and CBS-B-extrapolated IP-EOMCCSD($3h-2p$){2} adiabatic excitation energies for NCO differ from experiment by 0.189-0.234 eV for the $A^2\Sigma^+$ state and 0.159-0.173 eV for the $B^2\Pi$ state. Much of the above analysis applies to N_3 . Indeed, although the CBS-A and CBS-B extrapolations applied to the IP-EOMCCSD($2h-1p$) and IP-EOMCCSD($3h-2p$){2} total energies produce somewhat larger differences than in the case of NCO (7.6-8.4 millihartree in the case of the $X^2\Pi_g$ state and 6.6-8.6 millihartree in the case of the $B^2\Sigma_u^+$ state), the adiabatic excitation energies resulting from both CBS extrapolations are very stable, to within 0.003 eV for the IP-EOMCCSD($2h-1p$) approach and 0.002 eV for the IP-EOMCCSD($3h-2p$){2} method.

The CBS-A- and CBS-B-extrapolated IP-EOMCCSD($2h-1p$) and IP-EOMCCSD($3h-2p$){2} adiabatic excitation energies corresponding to the $B^2\Sigma_u^+$ state of N_3 are within 0.107-0.194 eV from experiment.

To conclude this discussion, we examine the performance of the IP-EOMCC methods in describing the equilibrium geometries of the ground and low-lying excited states of the NCO and N_3 species. The results of our geometry optimizations for NCO and N_3 are summarized in Table (3.21). In the case of the $X^2\Pi$ state of the NCO molecule, the basic IP-EOMCCSD($2h-1p$) approach produces results that deviate from experiment by 0.016-0.031 Å for the N-C bond length (designated as R_{N-C}) and 0.018-0.033 Å for the C-O bond length (designated as R_{C-O}) when the cc-pVXZ basis sets with $X = D, T$, and Q are employed. The same approach applied to the $A^2\Sigma^+$ state of NCO gives errors of 0.014-0.031 Å for R_{N-C} and 0.018-0.031 Å for R_{C-O} . The IP-EOMCCSD($3h-2p$) results exhibit very similar trends and accuracies, confirming the small role of higher-order contributions neglected in IP-EOMCCSD($2h-1p$) and present in IP-EOMCCSD($3h-2p$). The differences between the IP-EOMCCSD($3h-2p$) and experimental values of R_{N-C} are 0.024-0.038 Å for the $X^2\Pi$ state and 0.021-0.035 Å for the $A^2\Sigma^+$ state. The analogous differences for R_{C-O} are 0.015-0.025 Å for the $X^2\Pi$ state and 0.012-0.022 Å for the $A^2\Sigma^+$ state. Although it may very well be that higher-than- $3h-2p$ contributions neglected in the IP-EOMCCSD($3h-2p$) calculations and high angular momentum functions that are not present in the cc-pVTZ (or cc-pVQZ) basis sets are the sources of the above errors, it is also possible that the experimental geometries of the $X^2\Pi$ and $A^2\Sigma^+$ states of NCO reported in Ref. [295] might be in some error too, since none of the states of NCO examined here is as challenging as some of the states of CNC and C_2N . While there are unexplained differences between the

experimentally N–C and C–O bond lengths in the $X^2\Pi$ and $A^2\Sigma^+$ states of NCO and our theoretical predictions, it is of great interest to note that the differences between the results of the geometry optimizations using the full and active-space IP-EOMCCSD($3h-2p$) approaches are virtually none. Indeed, there is no difference (to within 0.001 Å) between the full IP-EOMCCSD($3h-2p$) and active-space IP-EOMCCSD($3h-2p$){2} results for the N–C bond length in the $X^2\Pi$ state and the corresponding C–O bond lengths differ by 0.002 Å only, when the cc-pVDZ and cc-pVTZ are employed. In the case of the $A^2\Sigma^+$ state, the differences between the full IP-EOMCCSD($3h-2p$) and active-space IP-EOMCCSD($3h-2p$){2} values of $R_{\text{N-C}}$ and $R_{\text{C-O}}$ are 0.004 Å and 0.001-0.002 Å, respectively. In the case of the $B^2\Pi$ state, these differences are 0.003 Å for $R_{\text{N-C}}$ and 0.006-0.008 Å for $R_{\text{C-O}}$. The active-space IP-EOMCCSD($3h-2p$){2} calculations are clearly capable of reproducing the parent IP-EOMCCSD($3h-2p$) data for the N–C and C–O bond lengths in the ground and excited states of NCO to very high accuracy.

Much of the above discussion applies to N_3 . The nearest-neighbor N–N bond lengths, designated as $R_{\text{N-N}}$, resulting from the IP-EOMCCSD($2h-1p$) calculations with the cc-pVXZ basis sets with $X = \text{D, T, and Q}$, differ from the corresponding experimental data by 0.003-0.020 Å for the $X^2\Pi_g$ state and 0.001-0.015 Å for the $B^2\Sigma_u^+$ state. The higher-order IP-EOMCCSD($3h-2p$) optimizations with the cc-pVDZ and cc-pVTZ basis sets produce similar results, errors of 0.002-0.013 Å for the $X^2\Pi_g$ state and 0.007 Å for the $B^2\Sigma_u^+$ state. The active-space IP-EOMCCSD($3h-2p$){2} approach, for which we could also afford the calculations with the cc-pVQZ basis set, produces $R_{\text{N-N}}$ values that deviate from experiment by 0.002-0.016 Å for the $X^2\Pi_g$ state and 0.004-0.012 Å for the $B^2\Sigma_u^+$ state. Again, there is a virtually perfect agreement between the expensive full IP-EOMCCSD($3h-2p$) and in-

expensive active-space IP-EOMCCSD($3h-2p$){2} calculations, where there is no difference (to within 0.001 Å) between the two sets of data in the case of the $X^2\Pi_g$ state and a very small, 0.002-0.003 Å, difference between the full IP-EOMCCSD($3h-2p$) and active-space IP-EOMCCSD($3h-2p$){2} values of R_{N-N} in the case of the $B^2\Sigma_u^+$ state. As in the case of the NCO molecule, the origin of the deviations between the IP-EOMCC calculations employing basis sets as large as cc-pVQZ, which seem numerically quite stable, and experimental R_{N-N} values could lie in the higher-than- $3h-2p$ correlations that we do not consider in this work or in the significance of the high angular momentum functions absent in the cc-pVTZ and cc-pVQZ bases, but one cannot exclude the possibility that the experimental data reported in Ref. [295] may need to be revisited. As in the EA-EOMCC calculations for the CNC and C₂N, it seems to us that the basic IP-EOMCCSD($2h-1p$) method is capable of producing optimized geometries of the ground- and excited-state NCO and N₃ molecules that are comparable to those obtained with the computationally more demanding IP-EOMCCSD($3h-2p$) methods, which is a useful observation from the point of view of other applications of such methods to geometry optimizations in other open-shell species.

Table 3.16: Total and adiabatic excitation energies for the ground and low-lying excited states of CNC, as obtained with the different EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.

Method	Basis	Total Energy (hartree)			Adiabatic Excitation Energy (eV)	
		$X \ ^2\Pi_g$	$A \ ^2\Delta_u$	$B \ ^2\Sigma_u^+$	$A \ ^2\Delta_u - X \ ^2\Pi_g$	$B \ ^2\Sigma_u^+ - X \ ^2\Pi_g$
EA-EOMCCSD($2p-1h$)	DZP	-130.406718	-130.141822	-130.125873	7.208	7.642
	x=D	-130.402813	-130.136443	-130.120048	7.248	7.694
	x=T	-130.502669	-130.220645	-130.204320	7.674	8.118
	x=Q	-130.534268	-130.248264	-130.232033	7.783	8.224
	CBS-A	-130.551020	-130.264878	-130.248586	7.786	8.230
	CBS-B	-130.552172	-130.264020	-130.247849	7.841	8.281
EA-EOMCCSD($3p-2h$)	DZP	-130.411686	-130.260720	-130.238177	4.108	4.721
	x=D	-130.408191	-130.257611	-130.234329	4.097	4.731
	x=T	-130.510334	-130.358548	-130.335201	4.130	4.766
EA-EOMCCSD($3p-2h$){4}	DZP	-130.409784	-130.259560	-130.236779	4.088	4.708
	x=D	-130.406511	-130.256819	-130.233332	4.073	4.712
	x=T	-130.507154	-130.356797	-130.333074	4.091	4.737
	x=Q	-130.538435	-130.388104	-130.364472	4.091	4.734
	CBS-A	-130.554997	-130.404664	-130.380953	4.091	4.736
	CBS-B	-130.556095	-130.405806	-130.382242	4.090	4.731
Experiment ^a					3.761	4.315

^a Taken from Refs. [295, 296].

Table 3.17: Total and adiabatic excitation energies for the ground and low-lying excited states of C₂N, as obtained with the different EA-EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.

Method	Basis	Total Energy (hartree)				Adiabatic Excitation Energy (eV)		
		$X\ ^2\Pi$	$A\ ^2\Delta$	$B\ ^2\Sigma^-$	$C\ ^2\Sigma^+$	$A\ ^2\Delta-X\ ^2\Pi$	$B\ ^2\Sigma^--X\ ^2\Pi$	$C\ ^2\Sigma^+-X\ ^2\Pi$
EA-EOMCCSD($2p-1h$)	DZP	-130.400501	-130.176452	-130.117500	-130.156651	6.097	7.701	6.635
	x=D	-130.400086	-130.174345	-130.115824	-130.152828	6.143	7.735	6.728
	x=T	-130.499176	-130.259914	-130.198940	-130.240134	6.511	8.170	7.049
	x=Q	-130.530280	-130.287558	-130.225643	-130.267832	6.605	8.290	7.142
	CBS-A	-130.546521	-130.303831	-130.241794	-130.283903	6.604	8.292	7.146
	CBS-B	-130.548409	-130.303738	-130.241271	-130.284090	6.658	8.358	7.193
EA-EOMCCSD($3p-2h$)	DZP	-130.405260	-130.292989	-130.270231	-130.265181	3.055	3.674	3.812
	x=D	-130.404842	-130.292610	-130.270337	-130.264086	3.054	3.660	3.830
	x=T	-130.506456	-130.394642	-130.370688	-130.366299	3.043	3.694	3.814
EA-EOMCCSD($3p-2h$){4}	DZP	-130.403651	-130.292385	-130.269696	-130.264361	3.028	3.645	3.791
	x=D	-130.403260	-130.292089	-130.269870	-130.263371	3.025	3.630	3.807
	x=T	-130.503555	-130.393547	-130.369756	-130.364858	2.993	3.641	3.774
	x=Q	-130.533052	-130.423692	-130.399922	-130.394864	2.976	3.623	3.760
	CBS-A	-130.543559	-130.433071	-130.408605	-130.404035	3.007	3.672	3.797
	CBS-B	-130.549517	-130.440555	-130.416854	-130.411633	2.965	3.610	3.752
Experiment ^a						2.636	2.779	3.306

^aTaken from Refs. [295, 297].

Table 3.18: Comparison of the optimized equilibrium geometries for the low-lying states of CNC and C₂N, as obtained with the EA-EOMCC and SAC-CI-SDT-*R*/PS approaches using the DZP[4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets.

Method	Basis	CNC ^a			C ₂ N ^b			
		<i>X</i> ² Π _{<i>g</i>}	<i>A</i> ² Δ _{<i>u</i>}	<i>B</i> ² Σ _{<i>u</i>} ⁺	<i>X</i> ² Π	<i>A</i> ² Δ	<i>B</i> ² Σ [−]	<i>C</i> ² Σ ⁺
SAC-CI-SDT- <i>R</i> /PS	DZP	1.253	1.256	1.259	(1.400, 1.185)	(1.315, 1.207)	(1.302, 1.223)	(1.311, 1.214)
EA-EOMCCSD(2 <i>p</i> -1 <i>h</i>)	DZP	1.259	1.258	1.260	(1.412, 1.196)	(1.372, 1.186)	(1.372, 1.190)	(1.365, 1.192)
	x=D	1.260	1.258	1.260	(1.412, 1.193)	(1.376, 1.182)	(1.376, 1.186)	(1.375, 1.188)
	x=T	1.242	1.245	1.247	(1.389, 1.178)	(1.363, 1.166)	(1.361, 1.170)	(1.356, 1.171)
	x=Q	1.239	1.241	1.243	(1.385, 1.174)	(1.362, 1.162)	(1.360, 1.166)	(1.356, 1.167)
EA-EOMCCSD(3 <i>p</i> -2 <i>h</i>)	DZP	1.261	1.262	1.264	(1.410, 1.198)	(1.329, 1.217)	(1.308, 1.241)	(1.322, 1.224)
	x=D	1.262	1.261	1.263	(1.409, 1.195)	(1.332, 1.212)	(1.313, 1.234)	(1.325, 1.220)
	x=T	1.246	1.246	1.248	(1.388, 1.180)	(1.316, 1.196)	(1.297, 1.215)	(1.308, 1.203)
EA-EOMCCSD(3 <i>p</i> -2 <i>h</i>){4}	DZP	1.262	1.262	1.264	(1.411, 1.197)	(1.329, 1.217)	(1.308, 1.241)	(1.322, 1.224)
	x=D	1.262	1.261	1.264	(1.408, 1.195)	(1.332, 1.212)	(1.313, 1.234)	(1.325, 1.220)
	x=T	1.246	1.246	1.249	(1.389, 1.180)	(1.316, 1.196)	(1.297, 1.216)	(1.308, 1.203)
	x=Q	1.242	1.243	1.245	(1.387, 1.175)	(1.315, 1.190)	(1.295, 1.210)	(1.307, 1.197)
Experiment ^c		1.245	1.249	1.259				

^aThe R_{C-N} bond lengths in Å. The *D*_{2*h*} symmetry was employed. ^bThe numbers in parentheses report the R_{C-C} and R_{C-N} bond lengths, respectively, in Å. The *C*_{2*v*} symmetry was employed. ^cTaken from [295–297].

Table 3.19: Total and adiabatic excitation energies for the ground and low-lying excited states of NCO, as obtained with the different IP EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.

Method	Basis	Total Energy (hartree)			Adiabatic Excitation Energy (eV)	
		$X^2\Pi$	$A^2\Sigma^+$	$B^2\Pi$	$A^2\Sigma^+-X^2\Pi$	$B^2\Pi-X^2\Pi$
IP EOMCCSD($2h-1p$)	DZP	-167.581951	-167.475380	-167.427707	2.900	4.197
	x=D	-167.576116	-167.468912	-167.419125	2.919	4.273
	x=T	-167.718401	-167.613444	-167.560443	2.856	4.298
	x=Q	-167.763112	-167.659168	-167.604432	2.828	4.318
	CBS-A	-167.786913	-167.683604	-167.626529	2.811	4.364
	CBS-B	-167.788412	-167.685072	-167.629275	2.812	4.330
IP EOMCCSD($3h-2p$)	DZP	-167.591701	-167.486508	-167.448441	2.862	3.898
	x=D	-167.587630	-167.481331	-167.441319	2.893	3.981
	x=T	-167.732789	-167.628442	-167.584981	2.839	4.022
IP EOMCCSD($3h-2p$){2}	DZP	-167.589579	-167.476255	-167.446490	3.081	3.891
	x=D	-167.585489	-167.470340	-167.438481	3.133	4.000
	x=T	-167.730109	-167.617771	-167.581463	3.057	4.045
	x=Q	-167.775958	-167.664698	-167.626155	3.028	4.076
	CBS-A	-167.799482	-167.687223	-167.648439	3.055	4.110
	CBS-B	-167.801944	-167.691316	-167.651415	3.010	4.096
Experiment ^a					2.821	3.937

^aTaken from Ref. [295].

Table 3.20: Total and adiabatic excitation energies for the ground and low-lying excited states of N₃, as obtained with the different IP EOMCC approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets and extrapolating to the CBS limit.

Method	Basis	Total Energy (hartree)		Adiabatic Excitation Energy (eV)
		$X\ ^2\Pi_g$	$B\ ^2\Sigma_u^+$	$B\ ^2\Sigma_u^+ - X\ ^2\Pi_g$
IP EOMCCSD($2h-1p$)	DZP	-163.716374	-163.545829	4.641
	x=D	-163.712083	-163.542627	4.611
	x=T	-163.848768	-163.677460	4.662
	x=Q	-163.891293	-163.719861	4.665
	CBS-A	-163.923747	-163.752426	4.662
	CBS-B	-163.915306	-163.743856	4.665
IP EOMCCSD($3h-2p$)	DZP	-163.729782	-163.560803	4.598
	x=D	-163.726673	-163.558437	4.578
	x=T	-163.865416	-163.696218	4.604
IP EOMCCSD($3h-2p$){2}	DZP	-163.728362	-163.554434	4.733
	x=D	-163.725315	-163.551533	4.729
	x=T	-163.863766	-163.689041	4.755
	x=Q	-163.907325	-163.732710	4.752
	CBS-A	-163.939566	-163.764109	4.747
	CBS-B	-163.931977	-163.757469	4.749
Experiment ^a				4.555

^aTaken from Ref. [295].

Table 3.21: Comparison of the optimized equilibrium geometries for the low-lying states of N₃ and NCO, as obtained with the IP EOMCC and SAC-CI-SDT-*R*/PS approaches using the DZP [4s2p1d] and cc-pVXZ (X = D, T, Q) basis sets.

Method	Basis	N ₃ ^{<i>a</i>}		NCO ^{<i>b</i>}		
		<i>X</i> ² Π _{<i>g</i>}	<i>B</i> ² Σ _{<i>u</i>} ⁺	<i>X</i> ² Π	<i>A</i> ² Σ ⁺	<i>B</i> ² Π
SAC-CI-SDT- <i>R</i> /PS	DZP	1.188	1.185	(1.230, 1.193)	(1.191, 1.190)	(1.220, 1.309)
IP EOMCCSD(2 <i>p</i> -1 <i>h</i>)	DZP	1.195	1.191	(1.232, 1.196)	(1.197, 1.192)	(1.225, 1.318)
	x=D	1.185	1.181	(1.231, 1.188)	(1.196, 1.184)	(1.223, 1.313)
	x=T	1.171	1.169	(1.219, 1.177)	(1.182, 1.175)	(1.206, 1.306)
	x=Q	1.168	1.165	(1.216, 1.173)	(1.179, 1.171)	(1.202, 1.304)
IP EOMCCSD(3 <i>p</i> -2 <i>h</i>)	DZP	1.200	1.196	(1.240, 1.198)	(1.200, 1.198)	(1.235, 1.328)
	x=D	1.190	1.187	(1.238, 1.191)	(1.200, 1.190)	(1.233, 1.322)
	x=T	1.175	1.173	(1.224, 1.181)	(1.186, 1.180)	(1.216, 1.312)
IP EOMCCSD(3 <i>p</i> -2 <i>h</i>){2}	DZP	1.200	1.194	(1.240, 1.196)	(1.196, 1.196)	(1.239, 1.319)
	x=D	1.190	1.184	(1.238, 1.189)	(1.196, 1.189)	(1.236, 1.314)
	x=T	1.175	1.171	(1.224, 1.179)	(1.182, 1.178)	(1.219, 1.306)
	x=Q	1.172	1.168	(1.222, 1.174)	(1.179, 1.174)	(1.214, 1.303)
Experiment ^{<i>c</i>}		1.188	1.180	(1.200, 1.206)	(1.165, 1.202)	

^{*a*}The R_{N-N} bond lengths in Å. The *D*_{2*h*} symmetry was employed. ^{*b*}The numbers in parentheses report the R_{N-C} and R_{C-O} bond lengths, respectively, in Å. The *C*_{2*v*} symmetry was employed. ^{*c*}Taken from Ref. [295].

Chapter 4

Potential Energy Surface

Extrapolation Schemes

4.1 Motivation

The primary goal of this dissertation so far has been the examination of *ab initio* electronic structure methods which allow for the efficient generation of highly accurate molecular PESs or chemical reaction pathways. Unfortunately, under the constraints of current computing capabilities and algorithms, the range of applicability of the electronic structure methods discussed in Sect. (3) is limited to small- to medium-sized systems if local correlation, fragmentation, or other similar techniques are not exploited, and those have additional intrinsic errors. This is not only because of characteristic steep scalings of computer costs of typical high-accuracy *ab initio* methods with the system size, but also because of the enormous numbers of points typically associated with PESs of larger molecules. Under the conventional procedure, one usually follows to obtain a PES, the calculation of m^{τ} points is required,

where m is the number of nuclear geometries required to represent a one-dimensional PES cut (typically, m is on the order of 10) and τ is the PES dimension (for the one-dimensional PES cuts along the relevant intrinsic reaction coordinates, $\tau = 1$; for the global multidimensional PESs, $\tau = 3\mathcal{M} - 6$, where $\mathcal{M}$ is the number of atomic nuclei in the molecular system of interest). Meanwhile, the CPU time associated with the accurate *ab initio* electronic structure calculation of a single point of the PES scales at least as kn_u^4 , where the prefactor k is a polynomial function of the number of occupied orbitals n_o . For most medium to large systems this can cause even a single point energy calculation to become prohibitively expensive when using a basis set of realistic size. The total CPU time required to generate a PES for a given molecular system scales as $km^\tau n_u^4$, a scaling so poor that studies involving larger polyatomic systems must be limited to small basis sets and small numbers of points on the PES for the calculations to remain computationally feasible. The focus of this chapter is to develop and test numerical techniques which can help reduce the enormous computer costs associated with the conventional procedure for generating PESs.

Considerable progress has been made toward alleviating these large computer costs with the proposal of an *ab initio* extrapolation scheme, described in Ref. [184], that predicts a PES corresponding to a larger basis set from the results of smaller basis set calculations by scaling electron correlation energies. In the PES extrapolation scheme suggested in Ref. [184], a universal correlation energy scaling factor is determined at a single nuclear geometry, called a pivot geometry, over a series of basis sets of growing size. The scaling factor is then applied to electron correlation energies calculated using smaller basis sets at the remaining geometries to obtain the entire PES at the desired (larger) basis set level. The original work [184], as well as more recent effort by us [193, 300], and Varandas [301, 302] have demonstrated

the effectiveness of this procedure, where in each studied case one could generate the target PESs to within, on average, fractions of a millihartree of the true calculated energies, while effectively reducing the number of points which must be calculated on the high-level PES of interest from m^τ to one or a few. The general principles behind the PES extrapolation methodology of Refs. [184, 193, 300] are outlined in Sect. (4.2), while Sect. (4.3) elaborates on the the so-called single-level PES extrapolation schemes based on the ideas laid down in Refs [184, 193, 300]. In Sect. (4.3.1) the details of the single-level PES extrapolation scheme are described, while Sect. (4.3.2) deals with an extension allowing one to perform PES extrapolations to the CBS-limit in an inexpensive way. Section (4.3.4) surveys a few of the user-defined parameters, which must be chosen using the PES extrapolation schemes examined in this work, and Sect. (4.3.3) demonstrates the performance of a few different single-level PES extrapolation schemes in the first-ever practical application involving a complex polyatomic system reported in Refs. [193, 300], namely the **bicbut** $\rightarrow$ **t-but** isomerization examined in Sect. (3.2.3).

Despite the significant improvements in computational expense offered by the single-level PES extrapolation schemes of Refs. [184, 193, 300], performing even a single high-level *ab initio* calculation with a larger basis set may sometimes be too taxing. In all such cases, the PES extrapolation scheme of Refs. [184, 193, 300] that requires one large-basis set high-level *ab initio* calculation cannot be of much help. Rather than be forced to resort to using less accurate methods to describe the associated PESs, an additional flexibility can be utilized within the framework of the PES extrapolation scheme of Refs. [184, 193, 300] which enables one to predict the correlation energy scaling factor for calculations using a higher-order methodology from scaling factors calculated with lower-order methods. Using this so-called

dual-level PES extrapolation scheme, introduced for the first time in Sect. (4.4), one can obtain a surface at the quality of a very accurate method and large basis set without having to calculate even a single point at the target *ab initio* level of interest. This new approach allows results to be obtained much more affordably for larger systems, as any explicit calculation at the desired level of theory is completely circumvented. To demonstrate the potential cost savings this implies, relative computational costs associated with performing typical calculations at the HF and selected MBPT and CC levels are collected in Table (4.1). The utility of replacing a CR-CC(2,3) calculation by, say, a CCSD or MP4SDQ calculation is immediately apparent after examination of Table (4.1), as the computational effort can be reduced by a factor of 7 or 40, respectively. If the dual-level PES extrapolation scheme could be used to produce a PES which is virtually identical to the results of explicit large basis set CR-CC(2,3) calculations using only the calculations performed with lower-level methods, such as CCSD or MP4SDQ and small basis set CR-CC(2,3) computations, it would offer incredible savings in the required computational effort.

To test the accuracy of the dual-level PES extrapolation scheme proposed in this thesis, where a number of different lower-order *ab initio* methodologies combined with small basis set CR-CC(2,3) calculations are used to approximate large basis set CR-CC(2,3) results, several chemical systems were chosen for benchmark studies. The systems considered here include the asymmetric stretch of the H₂O molecule, the bond stretching of the F₂ and HCl molecules, and, once again, the **bicbut**→*t*-**but** isomerization of Sect. (3.2.3). The three single-bond breaking potential energy curves have been well studied and serve our purposes particularly well because they have regions clearly dominated by dynamical correlation effects near the equilibrium bond lengths and regions with significant nondynamical correlation

Table 4.1: Computer costs of typical *ab initio* wave function calculations at the aug-cc-pVTZ basis set level, taken from Ref. [204]

Method	Scalings of CPU steps ^a		CPU time ^b
	Iterative	Noniterative	
CR-CC(2,3)	$\mathcal{N}^6$	$\mathcal{N}^7$	574 ^c
CCSD(T)	$\mathcal{N}^6$	$\mathcal{N}^7$	287
CCSD	$\mathcal{N}^6$	—	86
MP4SDQ	—	$\mathcal{N}^6$	15
MP3	—	$\mathcal{N}^6$	12
MP2	—	$\mathcal{N}^5$	3
HF	$\mathcal{N}^4$	—	1

^a $\mathcal{N}$ is a measure of the system size. ^b The CPU time for each method is reported as the time required by an energy gradient calculation for phosphinomethanol divided by the computer time characterizing the corresponding HF/aug-cc-pVTZ energy gradient calculation with the same software on the same computer. Although such costs depend to some extent (for example, 15%) on the machine, the program, and the computer load, they still provide a useful indication of computer resource demand. ^c The cost of the CR-CC(2,3) method was not measured explicitly, but rather approximated by doubling the CPU time of the CCSD(T) calculation (the most expensive steps of CR-CC(2,3) are approximately twice as expensive as those of CCSD(T)).

effects as relevant bond lengths approach the fully dissociated limits. The **bicbut** $\rightarrow$ **t-but** isomerization is also reexamined, as it provides a good example of a reaction profile for a polyatomic molecule which is composed of stationary points with strongly varying biradical character. The general theory associated with the dual-level PES extrapolation scheme is presented in Sect. (4.4.1) and results of the applications for the various di- and tri-atomic systems and the isomerization pathways of bicyclo[1.1.0]butane to *trans*-butadiene are given in Sects. (4.4.2) and (4.4.3), respectively. Finally, in Sect. (4.4.4), comparisons are made regarding relative cost and accuracy under a number of different combinations of user-defined choices required by the PES extrapolation schemes considered in this work.

4.2 Theory

The PES extrapolation scheme proposed in Ref. [184] and further developed in Refs. [193, 300–302] focuses on extrapolating the difference $\Delta E^{(A)}$ between the total electronic energy, $E^{(A)}$ obtained with some correlated approach A , and the base energy, $E^{(\text{base})}$, that one should be able to calculate with any basis set. For most applications discussed in this dissertation the base energy $E^{(\text{base})}$ is set equivalent to the RHF reference energy $E^{(\text{RHF})}$, so that the extrapolated energy component $\Delta E^{(A)}$ is the total correlation energy, but one can envision other ways of decomposing the total energy $E^{(A)}$ into $E^{(\text{base})}$ and $\Delta E^{(A)}$. For example, $E^{(\text{base})}$ could be the CASSCF energy and $E^{(A)}$ the MRCI energy, in which case we describe the non-dynamical correlation effects exactly within the CASSCF approximation and extrapolate the difference $\Delta E^{(A)} = E^{(A)} - E^{(\text{base})}$ describing dynamical correlations. In this chapter, we focus on two specific choices of $E^{(\text{base})}$ and $\Delta E^{(A)}$ both related to the choice of target method A as CR-CC(2,3), namely, (i) $E^{(\text{base})} = E^{(\text{RHF})}$ and $\Delta E^{(A)} =$

$E^{(\text{CR-CC}(2,3))} - E^{(\text{RHF})} \equiv \Delta E^{(\text{CR-CC}(2,3))}$, so that $E^{(\text{base})}$ is the RHF energy, which is easy to calculate, and $\Delta E^{(A)}$ is the CR-CC(2,3) correlation energy, which is the expensive part we want to extrapolate, and (ii) $E^{(\text{base})} = E^{(\text{CCSD})}$ and $\Delta E^{(A)} = E^{(\text{CR-CC}(2,3))} - E^{(\text{CCSD})} \equiv \delta(2,3)$, so that $E^{(\text{base})}$ is the CCSD energy, which we can often calculate even when large basis sets are employed, and $\Delta E^{(A)}$ is the triples correction of CR-CC(2,3), which is the most expensive component for the CR-CC(2,3) energy that we want to extrapolate. In addition to the target method A , we introduce the auxiliary correlated approach B , which in single-level PES extrapolation techniques of Refs. [184,193,300–302] equals A , and in dual-level schemes is some other correlated method, less expensive than A . We use the auxiliary approach B to determine the approximate correlation energy scaling factor, allowing us to rescale the desired $\Delta E^{(A)}$ energy part from smaller to larger basis sets, as described below.

Suppose a set of correlated PES calculations are performed with correlated methods A and B , as described below, using smaller basis sets indexed by formal numbers $m-1$ and m . Let us designate the resulting PESs obtained for basis sets $m-1$ and m using method A as $E_{m-1}^{(A)}(\mathbf{R})$ and $E_m^{(A)}(\mathbf{R})$, respectively, and using method B as $E_{m-1}^{(B)}(\mathbf{R})$ and $E_m^{(B)}(\mathbf{R})$, respectively. Using the PES extrapolation scheme that interests us here, the extrapolated PES for the $E_{m+1}^{(A)}(\mathbf{R})$ target $(m+1)$ -th basis set is obtained as

$$E_{m+1}^{(A)}(\mathbf{R}) = E_{m+1}^{(\text{base})}(\mathbf{R}) + \chi_{m+1,m}^{(B)}(\mathbf{R}) \Delta E_m^{(A)}(\mathbf{R}). \quad (4.1)$$

Here, $E_{m+1}^{(\text{base})}(\mathbf{R})$ is the energy of the base method calculated with the $(m+1)$ -th basis set, $\Delta E_m^{(A)}(\mathbf{R}) = E_m^{(A)}(\mathbf{R}) - E_m^{(\text{base})}(\mathbf{R})$, $\mathbf{R}$ denotes the τ -dimensional vector of the nuclear space coordinates defining the PES, and the scaling factor $\chi_{m+1,m}^{(B)}(\mathbf{R})$ is defined as [184,193,300–

$$\chi_{m+1,m}^{(B)}(\mathbf{R}) = 1 + \left[\frac{S(\mathbf{R})_{m,m-1}^{(B)} - 1}{S(\mathbf{R}_e)_{m,m-1}^{(B)} - 1} \right] [S(\mathbf{R}_e)_{m+1,m}^{(B)} - 1], \quad (4.2)$$

with

$$S(\mathbf{R})_{m,m-1}^{(B)} = \Delta E_m^{(B)}(\mathbf{R}) / \Delta E_{m-1}^{(B)}(\mathbf{R}). \quad (4.3)$$

Thus, the desired high-level PES $E_{m+1}^{(A)}(\mathbf{R})$, obtained with the largest basis set $m + 1$, is extrapolated from the PESs $E_{m-1}^{(B)}(\mathbf{R})$ and $E_m^{(B)}(\mathbf{R})$ obtained in smaller basis set calculations, the base energies $E_{m-1}^{(\text{base})}(\mathbf{R})$, $E_m^{(\text{base})}(\mathbf{R})$, and $E_{m+1}^{(\text{base})}(\mathbf{R})$, and a single correlated energy, $E_{m+1}^{(B)}(\mathbf{R}_e)$, calculated at the pivot geometry $\mathbf{R}_e$.

It is easy to see that Eq. (4.1) represents the simplest mathematical expression one can propose to extrapolate the energies $E_{m+1}^{(A)}(\mathbf{R})$ from the smaller basis sets $(m - 1)$ and m to the larger basis set $(m + 1)$. Indeed, the scaling function $\chi_{m+1,m}^{(B)}(\mathbf{R})$ satisfies the following desirable properties: (i) $\chi_{m+1,m}^{(B)}(\mathbf{R}) \rightarrow 1$ for all values of $\mathbf{R}$ when $m \rightarrow \infty$, and (ii) $\chi_{m+1,m}^{(B)}(\mathbf{R}_e) = E_{m+1}^{(B)}(\mathbf{R}_e) / E_m^{(B)}(\mathbf{R}_e)$, so that when $B = A$, $E_{m+1}^{(A)}(\mathbf{R}_e) = E_{m+1}^{(\text{base})}(\mathbf{R}_e) + \Delta E_{m+1}^{(A)}(\mathbf{R}_e)$, as one would like to have. Of course, we hope that $\chi_{m+1,m}^{(B)}(\mathbf{R})$ determined using the information obtained with smaller basis sets $(m - 1)$ and m , and the “correlation” energy $\Delta E_{m+1}^{(B)}(\mathbf{R}_e)$ obtained with target basis set $(m + 1)$ at a reference (pivot) geometry $\mathbf{R}_e$ only is universal enough to extrapolate the $\Delta E_{m+1}^{(A)}(\mathbf{R})$ values at the remaining points $\mathbf{R}$ on the PES, even when method B is an approximation to the target approach A .

4.3 Single-Level Potential Energy Surface Extrapolation Schemes

In the original proposal of the PES extrapolation scheme of Ref. [184], it was assumed that the scaling factor $\chi_{m+1,m}^{(B)}(\mathbf{R})$ and the correlation energy it is applied to, $\Delta E_m^{(A)}(\mathbf{R})$, were required to be generated using the same electronic structure method. While this constraint, i.e., that $A = B$ in Eqs. (4.1)–(4.3), will be a helpful simplifying assumption for the initial discussion of the PES extrapolation scheme considered here, it will be eventually shown in Sect. (4.4), that the theory can work equally well with $A \neq B$, where B is only an approximation to A , resulting in the so-called dual-level PES extrapolation scheme. In this section, we begin a discussion of several ways to employ the basic equations defining the PES extrapolation scheme, Eqs. (4.1)–(4.3), with the single-level PES extrapolation to larger basis sets and to the CBS-limit in Sects. (4.3.1) and (4.3.2), respectively. Following this, a discussion of the role of pivot geometries, R_e , and base wave functions, $\Psi^{(\text{base})}$, which are behind base energies $E^{(\text{base})}$, is given in Sect. (4.3.3), and an application to the bicyclobutane isomerization pathways of Sect. (3.2.3) is shown in Sect. (4.3.4).

4.3.1 Potential Energy Surface Extrapolation to Larger Basis Sets

Beginning with a simple example to demonstrate how the PES extrapolation scheme is typically used, consider a case where $A = B$ and $E^{(\text{base})} = E^{(\text{RHF})}$, in which a PES corresponding to the cc-pVQZ basis set (previously $m + 1$) is to be extrapolated from PESs obtained from cc-pVDZ ($m - 1$) and cc-pVTZ (m) basis set calculations. In this example, the following quantities must be collected to perform a single-level PES extrapolation: cor-

relation energies for every point calculated at the cc-pVDZ and cc-pVTZ basis set levels using method A , base energies for every point calculated at the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis set levels, and a single correlated energy calculated with the cc-pVQZ basis set at the pivot geometry $\mathbf{R}_e$. For simplicity, in all discussions in this dissertation except for those involving CBS-limit extrapolation, the cc-pVDZ, cc-pVTZ, and cc-pVQZ or the aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis set series are employed, meaning that for all finite basis set extrapolations considered here, the index m used in Eqs. (4.1)–(4.3) may be equated with the cardinal number 3 of the cc-pVXZ basis sets, corresponding to cc-pVTZ or aug-cc-pVTZ, and thus $m-1$ and $m+1$ correspond to cc-pVDZ or aug-cc-pVDZ and cc-pVQZ or aug-cc-pVQZ, respectively. Although PES extrapolations in this dissertation focus on extrapolating CR-CC(2,3) correlation energies across the (aug-)cc-pVXZ basis sets of Dunning and co-workers, which are very systematic in terms of the dependence of the angular momentum functions on the cardinal number X , it must be emphasized that the PES extrapolation method is not tailored to fit any single class of electronic structure approaches nor is it limited to any specific family of basis sets. This was demonstrated in the original study, Ref. [184], using the MRCI approach and a series of Pople-type basis sets.

4.3.2 Potential Energy Surface Extrapolation to the Complete Basis Set Limit

The theory for PES extrapolation schemes presented in Sect. (4.2) is not limited to basis sets immediately sequential in size either, as in, for example, the cc-pVDZ, cc-pVTZ, and cc-pVQZ sequence. In fact, the target basis set $m+1$ in Eqs. (4.1)–(4.3) could be of infinite size enabling us to extrapolate the PES $E_{m+1}^{(A)}(\mathbf{R})$, where $(m+1)$ represents the CBS limit,

from the finite basis set calculations. This requires knowledge of the CBS limits of the base energies, $E_{m+1}^{(\text{base})}(\mathbf{R}) \equiv E_{\infty}^{(\text{base})}(\mathbf{R})$, at each point $\mathbf{R}$ on the PES of interest and the CBS-limit of the correlation energy at the pivot geometry $\mathbf{R}_e$, $\Delta E_{m+1}^{(\text{base})}(\mathbf{R}) \equiv \Delta E_{\infty}^{(\text{base})}(\mathbf{R})$, which is needed to define the $\chi_{m+1,m}(\mathbf{R})$ scaling factor in a situation where $(m+1)$ represents the infinite basis set.

As an example, consider the PES extrapolation for a sequence of basis sets of cc-pVXZ quality to the CBS-limit in a situation where the target method A is CR-CC(2,3) and where the base energy originates from RHF calculations. The CBS-limit base PES $E_{m+1}^{(\text{base})}(\mathbf{R})$ needed in Eq. (4.1) can be determined via relatively inexpensive RHF/cc-pV6Z level calculations (recall from Sect. (3.2.1) that at the RHF level, cc-pV6Z results are an excellent approximation to the corresponding CBS limit). One can then determine the CBS limit of the correlation energy $\Delta E_{\infty}^{(A)}(\mathbf{R}_e)$ at the pivot geometry $\mathbf{R}_e$ using, for example, correlation energies obtained with the cc-pVTZ and cc-pVQZ basis sets and one of the CBS extrapolation laws, such as Eq. (3.98). The resulting CBS-limit value of $\Delta E_{\infty}^{(A)}(\mathbf{R}_e)$ can be used as $\Delta E_{m+1}^{(A)}(\mathbf{R}_e)$ to determine the $S(\mathbf{R}_e)_{m+1,m}^{(A)}$ ratio and, ultimately, the scaling factor $\chi_{m+1,m}^{(A)}(\mathbf{R})$ (recall that $B = A$ here). Thus, the necessary quantities for extrapolation of the CR-CC(2,3) PES to the CBS-limit are the RHF/cc-pVDZ and RHF/cc-pVTZ values at all points $\mathbf{R}$ of interest, the CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ values at all points $\mathbf{R}$ of interest, the CBS-limit base energies $E_{m+1}^{(\text{base})}(\mathbf{R})$ at all points $\mathbf{R}$ of interest, obtained, for example, in the RHF/cc-pV6Z calculations, and a CBS-limit extrapolated correlation energy $\Delta E_{\infty}^{(A)}(\mathbf{R}_e)$ corresponding to a single pivot geometry $\mathbf{R}_e$, which serves as the $\Delta E_{m+1}^{(A)}(\mathbf{R}_e)$ value in Eqs. (4.1)–(4.3). We could carry out a similar CBS PES extrapolation procedure for methods other than CR-CC(2,3) and other sequences of the correlation

consistent basis sets. An actual application of this procedure, as applied to the CR-CC(2,3) approach, will be discussed in Sect. (4.3.4), but before we do this, we examine the role of pivot geometries $\mathbf{R}_e$ and base wave functions used to define $E_{m+1}^{(\text{base})}(\mathbf{R})$ in Eq. (4.1).

4.3.3 The Role of Pivot Geometries and Base Wave Functions in Single-Level PES Extrapolations

A few inherent flexibilities of the PES extrapolation scheme should be discussed before moving on to applications. First, the pivot geometry, which appears in Eq. (4.2) as $\mathbf{R}_e$, can be chosen as the equilibrium geometry, as the geometry corresponding to the reactants or products, or in principle, as any other single point on the PES of interest. It is even possible to choose more than one pivot geometry in PES extrapolations [301,302], although the benefits of doing so may be minimal since, as was shown in Ref. [193], the accuracy of results does not depend on the choice of the pivot geometry $\mathbf{R}_e$. This fact will be illustrated in the next section. One also has a choice of generalizing Eqs. (4.1)–(4.3) to several pivots $\mathbf{R}_e$ that may provide an adequate sampling of the PES of interest for performing an extrapolation. The idea of multiple pivots $\mathbf{R}_e$ has been explored with considerable success, in Refs. [301,302].

The other inherent flexibility which will be investigated numerically in the next section is the choice of base wave function defining the base energies in Eqs. (4.1)–(4.3). In the specific case of the PES calculations performed in Sect. (4.3.4), where the total electronic energies of interest are those obtained using the CR-CC(2,3) approach, two alternative definitions of the base energy $E^{(\text{base})}$ are considered, $E^{(\text{base})} = E^{(\text{RHF})}$ and $E^{(\text{base})} = E^{(\text{CCSD})}$. In the former case, as already alluded to above, we decompose the CR-CC(2,3) energy into the RHF base energy and the CR-CC(2,3) total correlation energy, which we want to extrapolate across

the entire PES. In the latter case, the CR-CC(2,3) energy is decomposed into the CCSD base energy and the triples correction of CR-CC(2,3) which we want to extrapolate across the entire PES. The relative performance of the PES extrapolation procedure based on Eqs. (4.1)–(4.3) under each of these two definitions of the base wave function is reported in the next section.

4.3.4 Application to the Isomerization of Bicyclobutane to Butadiene

Since the excellent performance of the single-level PES extrapolation scheme had already been established for di- and tri-atomic molecular PESs in Ref. [184], in 2008 we published a study aimed at assessing the potential usefulness of the PES extrapolation scheme in studies of chemical reaction pathways involving polyatomic molecules [193]. In that work, the PES extrapolation scheme was applied to the stationary points defining the conrotatory and disrotatory paths characterizing the isomerization of bicyclo[1.1.0]butane to buta-1,3-diene, which we discussed earlier in Sect. (3.2.3). There were a number of reasons for conducting the study in Ref. [193]. The first reason was to examine the basis set dependence of the CR-CC(2,3) results for both isomerization pathways, The second reason was to examine whether the PES extrapolation scheme was capable of recovering the results of the laborious point-wise CR-CC(2,3)/cc-pVQZ calculations from the PESs obtained in the CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ calculations, the base energies (RHF or CCSD) obtained in the cc-pVDZ, cc-pVTZ, and cc-pVQZ calculations, and a single CR-CC(2,3) energy obtained with the cc-pVQZ basis set at the pivot geometry $\mathbf{R}_e$ on the relevant polyatomic reaction pathway. Finally, the third reason for the study in Ref. [193] was to demonstrate the effectiveness of

our PES extrapolation procedure to the CBS-limit.

As already mentioned, two choices of the base wave function for defining the ΔE values that enter Eqs. (4.1)–(4.3) were considered, namely, the RHF energy ($E^{(\text{base})} = E^{(\text{RHF})}$) and the CCSD energy ($E^{(\text{base})} = E^{(\text{CCSD})}$). The results obtained with both choices of the reference energy and for the pivot geometry defined by the **bicbut** reactant are collected in Table (4.2). Different choices of the pivot geometry, $\mathbf{R}_e$, were also considered, each corresponding to one of the stationary points that define the two isomerization pathways. These results are shown in Tables (4.3)–(4.7). It can be seen in Tables (4.2)–(4.7) that the resulting maximum differences between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies characterizing both pathways are virtually independent of the choice of the pivot geometry $\mathbf{R}_e$. For this reason, in the following discussion we mainly focus on one specific choice of $\mathbf{R}_e$, namely, the geometry of the **bicbut** reactant (see Figure 3.2), as shown in Table (4.2), while mentioning numerical results produced using other choices of $\mathbf{R}_e$ very briefly.

As one can see in Table (4.2), independent of the choice of the base energies (RHF or CCSD), there is virtually no difference between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies. The differences between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies characterizing both isomerization pathways do not exceed 0.631 millihartree when $E^{(\text{base})} = E^{(\text{RHF})}$, and 0.277 millihartree when $E^{(\text{base})} = E^{(\text{CCSD})}$, when the geometry of the **bicbut** reactant is used as the pivot geometry $\mathbf{R}_e$. To appreciate the small magnitude of these extrapolation errors, the changes in the total electronic energies when going from the cc-pVTZ to cc-pVQZ basis sets are also included in Table (4.2), which are values on the order of 42–44 millihartree. Similar remarks are true for the other pivot ge-

Table 4.2: The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **bicbut** reactant defines the pivot geometry for the PES extrapolations.

RHF Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695497	-44.149	0.000
con_TS	-155.424792	-155.581666	-155.625392	-155.625147	-43.727	0.245
dis_TS	-155.388352	-155.540034	-155.582481	-155.581850	-42.447	0.631
g-but	-155.533086	-155.689327	-155.733236	-155.732754	-43.909	0.482
gt_TS	-155.528145	-155.684308	-155.728073	-155.727663	-43.764	0.410
t-but	-155.537788	-155.694080	-155.738043	-155.737480	-43.963	0.563
CCSD Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695497	-44.149	0.000
con_TS	-155.424792	-155.581666	-155.625392	-155.625212	-43.727	0.180
dis_TS	-155.388352	-155.540034	-155.582481	-155.582364	-42.447	0.117
g-but	-155.533086	-155.689327	-155.733236	-155.732971	-43.909	0.265
gt_TS	-155.528145	-155.684308	-155.728073	-155.727894	-43.764	0.179
t-but	-155.537788	-155.694080	-155.738043	-155.737766	-43.963	0.277

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal CR-CC(2,3)/cc-pVQZ energies, in hartree, resulting from the PES extrapolation procedure discussed in the text. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^dDeviations, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies.

Table 4.3: The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **con_TS** transition state defines the pivot geometry for the PES extrapolations.

RHF Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695750	-44.149	-0.253
con_TS	-155.424792	-155.581666	-155.625392	-155.625392	-43.727	0.000
dis_TS	-155.388352	-155.540034	-155.582481	-155.582077	-42.447	0.404
g-but	-155.533086	-155.689327	-155.733236	-155.733001	-43.909	0.235
gt_TS	-155.528145	-155.684308	-155.728073	-155.727908	-43.764	0.164
tbut	-155.537788	-155.694080	-155.738043	-155.737726	-43.963	0.317
CCSD Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695698	-44.149	-0.202
con_TS	-155.424792	-155.581666	-155.625392	-155.625392	-43.727	0.000
dis_TS	-155.388352	-155.540034	-155.582481	-155.582510	-42.447	-0.029
g-but	-155.533086	-155.689327	-155.733236	-155.7323159	-43.909	0.077
gt_TS	-155.528145	-155.684308	-155.728073	-155.728082	-43.764	-0.009
tbut	-155.537788	-155.694080	-155.738043	-155.737954	-43.963	0.090

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal CR-CC(2,3)/cc-pVQZ energies, in hartree, resulting from the PES extrapolation procedure discussed in the text. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^dDeviations, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies.

Table 4.4: The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **dis_TS** transition state defines the pivot geometry for the PES extrapolations.

RHF Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.696200	-44.149	-0.704
con_TS	-155.424792	-155.581666	-155.625392	-155.625828	-43.727	-0.436
dis_TS	-155.388352	-155.540034	-155.582481	-155.582481	-42.447	0.000
g-but	-155.533086	-155.689327	-155.733236	-155.733439	-43.909	-0.203
gt_TS	-155.528145	-155.684308	-155.728073	-155.728345	-43.764	-0.272
tbut	-155.537788	-155.694080	-155.738043	-155.738163	-43.963	-0.120
CCSD Base						
E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d	
bicbut	-155.493284	-155.651348	-155.695497	-155.695659	-44.149	-0.162
con_TS	-155.424792	-155.581666	-155.625392	-155.625357	-43.727	0.036
dis_TS	-155.388352	-155.540034	-155.582481	-155.582481	-42.447	0.000
g-but	-155.533086	-155.689327	-155.733236	-155.733122	-43.909	0.114
gt_TS	-155.528145	-155.684308	-155.728073	-155.728045	-43.764	0.028
tbut	-155.537788	-155.694080	-155.738043	-155.737917	-43.963	0.127

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal CR-CC(2,3)/cc-pVQZ energies, in hartree, resulting from the PES extrapolation procedure discussed in the text. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^dDeviations, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies.

Table 4.5: The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **g-but** intermediate defines the pivot geometry for the PES extrapolations.

RHF Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695992	-44.149	-0.495
con_TS	-155.424792	-155.581666	-155.625392	-155.625626	-43.727	-0.234
dis_TS	-155.388352	-155.540034	-155.582481	-155.582294	-42.447	0.187
g-but	-155.533086	-155.689327	-155.733236	-155.733236	-43.909	0.000
gt_TS	-155.528145	-155.684308	-155.728073	-155.728142	-43.764	-0.070
tbut	-155.537788	-155.694080	-155.738043	-155.737960	-43.963	0.083
CCSD Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695781	-44.149	-0.285
con_TS	-155.424792	-155.581666	-155.625392	-155.625467	-43.727	-0.074
dis_TS	-155.388352	-155.540034	-155.582481	-155.582570	-42.447	-0.089
g-but	-155.533086	-155.689327	-155.733236	-155.733236	-43.909	0.000
gt_TS	-155.528145	-155.684308	-155.728073	-155.728159	-43.764	-0.086
tbut	-155.537788	-155.694080	-155.738043	-155.738031	-43.963	0.012

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal CR-CC(2,3)/cc-pVQZ energies, in hartree, resulting from the PES extrapolation procedure discussed in the text. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^dDeviations, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies.

Table 4.6: The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **gt_TS** transition state defines the pivot geometry for the PES extrapolations.

RHF Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695920	-44.149	-0.423
con_TS	-155.424792	-155.581666	-155.625392	-155.625557	-43.727	-0.164
dis_TS	-155.388352	-155.540034	-155.582481	-155.582229	-42.447	0.252
g-but	-155.533086	-155.689327	-155.733236	-155.733166	-43.909	0.070
gt_TS	-155.528145	-155.684308	-155.728073	-155.728073	-43.764	0.000
tbut	-155.537788	-155.694080	-155.738043	-155.737891	-43.963	0.153
CCSD Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695689	-44.149	-0.192
con_TS	-155.424792	-155.581666	-155.625392	-155.625384	-43.727	0.009
dis_TS	-155.388352	-155.540034	-155.582481	-155.582503	-42.447	-0.022
g-but	-155.533086	-155.689327	-155.733236	-155.733150	-43.909	0.086
gt_TS	-155.528145	-155.684308	-155.728073	-155.728073	-43.764	0.000
tbut	-155.537788	-155.694080	-155.738043	-155.737945	-43.963	0.099

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal CR-CC(2,3)/cc-pVQZ energies, in hartree, resulting from the PES extrapolation procedure discussed in the text. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^dDeviations, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies.

Table 4.7: The calculated CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ energies and the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **t-but** product defines the pivot geometry for the PES extrapolations.

RHF Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.696077	-44.149	-0.580
con_TS	-155.424792	-155.581666	-155.625392	-155.625709	-43.727	-0.316
dis_TS	-155.388352	-155.540034	-155.582481	-155.582370	-42.447	0.110
g-but	-155.533086	-155.689327	-155.733236	-155.733319	-43.909	-0.083
gt_TS	-155.528145	-155.684308	-155.728073	-155.728225	-43.764	-0.152
tbut	-155.537788	-155.694080	-155.738043	-155.738043	-43.963	0.000
CCSD Base						
	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_4^{(extr)b}$	ΔE_{43}^c	ϵ_4^d
bicbut	-155.493284	-155.651348	-155.695497	-155.695795	-44.149	-0.298
con_TS	-155.424792	-155.581666	-155.625392	-155.625479	-43.727	-0.086
dis_TS	-155.388352	-155.540034	-155.582481	-155.582579	-42.447	-0.099
g-but	-155.533086	-155.689327	-155.733236	-155.733248	-43.909	-0.013
gt_TS	-155.528145	-155.684308	-155.728073	-155.728172	-43.764	-0.099
tbut	-155.537788	-155.694080	-155.738043	-155.738043	-43.963	0.000

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal CR-CC(2,3)/cc-pVQZ energies, in hartree, resulting from the PES extrapolation procedure discussed in the text. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^dDeviations, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies.

ometries examined in Tables (4.3)–(4.7), which correspond to the remaining stationary points along the conrotatory and disrotatory pathways defining the **bicbut**→**t-but** isomerization. As shown in Tables (4.3)–(4.7), the differences between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies characterizing both isomerization pathways do not exceed 0.704 millihartree, when $E^{(\text{base})} = E^{(\text{RHF})}$, and 0.298 millihartree, when $E^{(\text{base})} = E^{(\text{CCSD})}$, and are often much smaller, independent of the choice of the pivot geometry $\mathbf{R}_e$. Thus, the single-level PES extrapolation procedure, originally proposed in Ref. [184] and further developed in Refs. [193,300], reproduces changes in the total energies when going from the cc-pVTZ to cc-pVQZ basis sets to within 1.5 %, when the RHF energies are used as the base energies, and 0.6 %, when the CCSD energies are employed instead to define the base energies $E^{(\text{base})}$. As one might expect and as shown in Tables (4.2)–(4.7), the use of the CCSD wave function in determining the base energies reduces the observed differences between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies. However, this is to be expected since the use of the correlated base energies, such as those obtained in the CCSD calculations, reduces the fraction of the correlation energy to be extrapolated, which in turn decreases the magnitude of the extrapolation errors. On the other hand, it is quite remarkable that the use of uncorrelated RHF base energies in determining the ΔE values that enter Eqs. (4.1)–(4.3) leads to extrapolated CR-CC(2,3)/cc-pVQZ PESs which are identical to the calculated CR-CC(2,3)/cc-pVQZ PESs to within ~ 0.7 millihartree, independent of the choice of $\mathbf{R}_e$. These small differences between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ electronic energies are also reflected in the small differences between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ enthalpy values characterizing the six stationary points along the conrotatory and disrotatory pathways that define the isomerization of bicyclo[1.1.0]butane to

buta-1,3-diene, shown in Figure (3.2), which are on the order of 1 kcal/mol or, in most cases, even less.

Returning now to the third reason for the study of the **bicbut**→*t*-**but** reaction published in Ref. [193], this system was used to test the accuracy which may be obtained using single-level PES extrapolation techniques to obtain a PES at the CBS-limit. In the case in Ref. [193], the cc-pVDZ basis set was chosen as basis level $m-1$, the cc-pVTZ basis set was chosen as basis level m , and the infinite basis set limit was chosen as basis set $m+1$ in Eqs. (4.1)–(4.3). Thus, results of CR-CC(2,3) calculations at the cc-pVTZ and cc-pVQZ basis set levels were used to determine the CBS limit of the CR-CC(2,3) correlation energy at $\mathbf{R}_e$ using Eq. (3.98), and the resulting CBS value of the CR-CC(2,3) correlation energy was subsequently used to define $\Delta E_{m+1}^{(A)}(\mathbf{R}_e)$ in Eqs. (4.1)–(4.3) (recall again that $A = B$ in the present considerations). Then, the $\Delta E_{m+1}^{(A)}(\mathbf{R}_e)$ correlation energy at $\mathbf{R}_e$ determined in this way and the CR-CC(2,3)/cc-pVDZ and CR-CC(2,3)/cc-pVTZ correlation energies, $\Delta E_{m-1}^{(A)}(\mathbf{R})$ and $\Delta E_m^{(A)}(\mathbf{R})$, respectively, at all geometries $\mathbf{R}$ of interest were used to obtain the scaling factor $\chi_{m+1,m}^{(A)}(\mathbf{R})$ from the cc-pVTZ basis set to the CBS limit according to Eqs. (4.2) and (4.3). Once the scaling factor $\chi_{m+1,m}^{(A)}(\mathbf{R})$ at each $\mathbf{R}$ was established, we used it to determine the CBS limit of the CR-CC(2,3) PES, $E_{m+1}^{(A)}(\mathbf{R})$, at all geometries of interest by multiplying the CR-CC(2,3) correlation energies $\Delta E_m^{(A)}(\mathbf{R})$ obtained with the cc-pVTZ basis set by $\chi_{m+1,m}^{(A)}(\mathbf{R})$ and by adding the resulting energies to the base energies $E_{m+1}^{(\text{base})}(\mathbf{R})$ obtained in the RHF/cc-pV6Z calculations which, as already explained, are practically equivalent to the RHF energies in the CBS limit. By avoiding the point-wise CBS extrapolations of the CR-CC(2,3) correlation energies from the cc-pVTZ and cc-pVQZ basis sets at all geometries $\mathbf{R}$ (we had to perform the CBS extrapolation of the CR-CC(2,3) energy only at the pivot

geometry $\mathbf{R}_e$) and by using the cc-pV6Z values of the RHF energies, we saved a lot of CPU cycles, while producing the smooth CR-CC(2,3)-level PESs corresponding to the CBS limit, since the base RHF PESs obtained with a large, cc-pV6Z basis set and the scaling factor $\chi_{m+1,m}^{(A)}(\mathbf{R})$ are smooth functions of $\mathbf{R}$.

Indeed, as shown in Table 4.8, the differences between the energies resulting from the CBS extrapolation scheme that combines Eqs. (4.1)–(4.3) with Eq. (3.98) and the energies resulting from the conventional point-wise CBS extrapolations using the CR-CC(2,3)/cc-pVTZ and CR-CC(2,3)/cc-pVQZ data at each stationary point defining the conrotatory and disrotatory pathways shown in Figure (3.2) do not exceed 1.092 millihartree, i.e., they are on the same order as the intrinsic error of the CBS extrapolations based on Eq. (3.98). This method provides a significant advantage over conventional CBS-limit extrapolation of PESs when it is considered that combining Eqs. (4.1)–(4.3) with Eq. (3.98) requires only one CR-CC(2,3)/cc-pVQZ calculation (at the pivot geometry), whereas the conventional point-wise CBS extrapolation method requires the CR-CC(2,3)/cc-pVQZ calculations at each point on the PES of interest, which represents an enormous increase in computer cost.

Table 4.8: The calculated CR-CC(2,3)/cc-pVDZ, CR-CC(2,3)/cc-pVTZ, and CR-CC(2,3)/cc-pVQZ energies and the CBS values of the CR-CC(2,3) energies obtained using the point-wise extrapolations exploiting Eq. (3.98) and the CBS extrapolation procedure combining Eqs. (4.1)–(4.3) with Eq. (3.98) discussed in the text at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **bicbut** reactant was used to define the pivot geometry for the PES extrapolations based on Eqs. (4.1)–(4.3).

	E_2^a	E_3^a	$E_4^{(calc)a}$	$E_\infty^{(calc)b}$	$E_\infty^{(extr)c}$	$\Delta E_{\infty,3}^d$	ϵ_∞^e
bicbut	-155.493284	-155.651348	-155.695497	-155.723202	-155.723202	-71.854	0.000
con_TS	-155.424792	-155.581666	-155.625392	-155.652777	-155.652353	-71.112	0.424
dis_TS	-155.388352	-155.540034	-155.582481	-155.608902	-155.607810	-68.868	1.092
g-but	-155.533086	-155.689327	-155.733236	-155.760795	-155.759962	-71.468	0.833
gt_TS	-155.528145	-155.684308	-155.728073	-155.755491	-155.754781	-71.182	0.710
t-but	-155.537788	-155.694080	-155.738043	-155.765622	-155.764647	-71.541	0.975

^aTotal energies, in hartree, calculated at the CR-CC(2,3)/cc-pVDZ (E_2), CR-CC(2,3)/cc-pVTZ (E_3), and CR-CC(2,3)/cc-pVQZ (E_4) levels. ^bTotal energies, in hartree, obtained by adding the RHF/cc-pV6Z energies to the CBS correlation energies resulting from the point-wise CBS extrapolations employing the CR-CC(2,3)/cc-pVTZ and CR-CC(2,3)/cc-pVQZ calculations and Eq. (3.98). ^cTotal energies, in hartree, obtained by adding the RHF/cc-pV6Z energies to the CBS values of the correlation energy resulting from the extrapolation procedure combining Eqs. (4.1)–(4.3) with Eq. (3.98), as discussed in the text. ^dDifferences, in millihartree, between the CR-CC(2,3)/CBS energies resulting from the point-wise CBS extrapolations employing Eq. (3.98) and the corresponding CR-CC(2,3)/cc-pVTZ energies. ^eDifferences, in millihartree, between the CR-CC(2,3)/CBS energies resulting from the point-wise CBS extrapolations employing Eq. (3.98) and the CBS PES extrapolation combining Eqs. (4.1)–(4.3) with Eq. (3.98).

4.4 Dual-Level Potential Energy Surface Extrapolation Schemes

The main idea of the PES extrapolation scheme is to scale the difference, ΔE , between the total electronic energy E , and the energy of some base wave function $E^{(\text{base})}$. The difference, ΔE , representing the correlation energy or some fraction of it, is scaled to the quality of a larger basis set by applying a scaling factor which predicts the change in the correlation energy with the size of the basis set at a particular nuclear configuration. Since the basis set dependence of the correlation energy may be similar for related electronic structure methods, after the original PES extrapolation scheme was proposed in Ref. [184] and further developed and tested in Ref. [193], additional flexibility was introduced by allowing the scaling factor to be generated using a method different from the method used to calculate the surface of interest. This is reflected in Eqs. (4.1) – (4.3) by the designation of method A , corresponding to the desired level of theory of the target PES, and method B , corresponding to the level of theory used in the generation of the scaling factor, $\chi_{m+1,m}^{(B)}(\mathbf{R})$. A discussion of the ideas behind the dual-level PES extrapolation scheme is given in Sect. (4.4.1). This is followed by a few applications in Sects. (4.4.2) and (4.4.3), where the CR-CC(2,3) method is used to calculate the PESs of interest using smaller basis sets, where computer costs remain low, in combination with the dual-level PES extrapolation scheme, in which the correlation scaling factor $\chi_{m+1,m}^{(B)}(\mathbf{R})$ is generated with lower-order methods, to obtain the PESs corresponding to the massively more expensive large basis set CR-CC(2,3) calculations. The resulting PESs are obtained using only a fraction of the computational resources required by the single-level PES extrapolation scheme, not to mention the point-by-point CR-CC(2,3) computations

using larger bases. A detailed analysis of the huge computational savings which the single- and dual-level PES extrapolation schemes can offer will be presented in Sect. (4.4.4).

4.4.1 Using Lower-Order Methods to Produce Correlation Energy Scaling Factors in High-Level Calculations

The dual-level extrapolation scheme presented here closely follows the original procedure, proposed in Ref. [184] and described in Sect. (4.3), differing only in the definition of the scaling factor $\chi_{m+1,m}^{(B)}(\mathbf{R})$, which may now be obtained with a different, less expensive, methodology, B , than the target methodology A of interest. By allowing A and B to differ, the approximate high-level PES $E_{m+1}^{(A)}(\mathbf{R})$ can be generated by extrapolating the PES $E_m^{(A)}(\mathbf{R})$ obtained with smaller basis set m , using the base surfaces $E_{m-1}^{(\text{base})}(\mathbf{R})$, $E_m^{(\text{base})}(\mathbf{R})$, and $E_{m+1}^{(\text{base})}(\mathbf{R})$, the correlated surfaces $E_{m-1}^{(B)}(\mathbf{R})$ and $E_m^{(B)}(\mathbf{R})$, and a single correlated energy calculated at the pivot geometry, $E_{m+1}^{(B)}(\mathbf{R}_e)$, all obtained with method B , which is, by choice, less expensive than the target method A . The most significant advantage over the original PES extrapolation scheme, where $\chi_{m+1,m}^{(A)}(\mathbf{R})$ was used, is that the calculation of the single high-level energy $E_{m+1}^{(A)}(\mathbf{R}_e)$, which accounts for the majority of the expense of the original procedure of Ref. [184], is avoided entirely and replaced instead by the calculation of $E_{m+1}^{(B)}(\mathbf{R}_e)$, using a quantum-chemistry method B which is less expensive than A .

The conventional low-order MBPT or CC approximations are among the approaches which may be considered as lower-order methods for the extrapolations of the high-level PESs resulting from higher-order CC calculations. The systematically improvable hierarchy of MBPT approaches is especially appealing in this study because it provides a series of “built in” lower-order methods that can facilitate an investigation of the lowest levels of correlation

treatment required to properly reproduce the correlation scaling factors for the higher-level CC methodology, CR-CC(2,3). The lowest-level MBPT methods, including MP2, MP3, and MP4X ($X = D, DQ$, and SDQ), are all much less expensive than the CR-CC(2,3) theory, with formal noniterative scaling steps of only $\mathcal{N}^5$ for MP2 and $\mathcal{N}^6$ for MP3 and MP4SDQ. For comparison, the CPU-time determining steps of CR-CC(2,3) scale as $\mathcal{N}^6$ in the iterative CCSD part and $\mathcal{N}^7$ in the triples correction part. In fact, even CCSD offers considerable savings compared to CR-CC(2,3), so this would be another candidate for determining the correlation energy scaling factor for extrapolating the CR-CC(2,3) PES.

As an example, if, say $A = \text{CR-CC}(2,3)$ and $B = \text{CCSD}$, a CR-CC(2,3)/cc-pVQZ PES $\Delta E_4^{(\text{CR-CC}(2,3))}(\mathbf{R})$ may be obtained by extrapolating a CR-CC(2,3)/cc-pVTZ PES $\Delta E_3^{(\text{CR-CC}(2,3))}(\mathbf{R})$ by applying $\chi_{4,3}^{(\text{CCSD})}(\mathbf{R})$, a correlation energy scaling factor constructed from CCSD/cc-pVDZ energies, CCSD/cc-pVTZ energies, and a single CCSD/cc-pVQZ energy at the pivot geometry $\mathbf{R}_e$ (rather than from the analogous set of energies obtained with CR-CC(2,3), as would be the case in the previously discussed single-level PES extrapolation scheme) and adding the resulting $\chi_{4,3}^{(\text{CCSD})}(\mathbf{R})\Delta E_3^{(\text{CR-CC}(2,3))}(\mathbf{R})$ term to an explicitly calculated RHF/cc-pVQZ PES that provides the base $E_4^{(\text{base})}(\mathbf{R})$ term in Eq. (4.1). Similarly, if $A = \text{CR} - \text{CC}(2,3)$ and $B = \text{MP4SDQ}$, we can determine the CR-CC(2,3)/cc-pVQZ-level PES $E_4^{(\text{CR-CC}(2,3))}(\mathbf{R})$ by adding the $\chi_{4,3}^{(\text{MP4SDQ})}(\mathbf{R})\Delta E_3^{(\text{CR-CC}(2,3))}(\mathbf{R})$ term, where $\chi_{4,3}^{(\text{MP4SDQ})}(\mathbf{R})$ is a correlation energy scaling factor obtained from the MP4SDQ calculation using the cc-pVDZ and cc-pVTZ basis sets and a single MP4SDQ/cc-pVQZ point calculated at $\mathbf{R}_e$, to the RHF/cc-pVQZ base energy $E_4^{(\text{base})}(\mathbf{R})$. Other examples of the auxiliary method B used in determining the correlation energy scaling factor for extrapolating the CR-CC(2,3) PES examined in this work include the MP2, MP3, MP4D, and MP4DQ ap-

proaches. We apply the dual-level PES extrapolation procedure, as described above, to predict the CR-CC(2,3)/cc-pVQZ or CR-CC(2,3)/aug-cc-pVQZ PESs characterizing a few single-bond dissociations (Sect. (4.4.2)) and the previously examined bicyclobutane isomerization pathways (Sect. (4.4.3)).

4.4.2 Application to Single Bond-Breaking Potential Energy Curves

In this section, single bond-breaking potential energy curves of the H_2O , HCl , and F_2 molecules are considered as benchmark cases to test the accuracy of the dual-level PES extrapolation procedure. In all PES extrapolations discussed in this section, the pivot geometry $\mathbf{R}_e$ is taken to be the equilibrium geometry, the base wavefunction is chosen to be the RHF wavefunction, and the CR-CC(2,3)/aug-cc-pVTZ PESs are extrapolated to the CR-CC(2,3)/aug-cc-pVQZ level by applying a correlation energy scaling factor generated from the following quantities: RHF/aug-cc-pVDZ, RHF/aug-cc-pVTZ and RHF/aug-cc-pVQZ base surfaces, B /aug-cc-pVDZ and B /aug-cc-pVTZ surfaces, and a single B /aug-cc-pVQZ energy at the pivot geometry $\mathbf{R}_e$, where the lower-order method B is MP2, MP3, MP4X ($X = \text{D, DQ, SDQ}$), or CCSD.

In Table (4.9), extrapolation errors are collected for selected points on the $\text{H}_2\text{O} \rightarrow \text{OH} + \text{H}$ dissociation curve. In the first two columns of this table, benchmark CR-CC(2,3)/aug-cc-pVQZ energies, $E_4^{(\text{CR-CC}(2,3))}(\mathbf{R})$ are given for selected points $\mathbf{R}$ on the potential energy curve, along with the energy differences, $\Delta E_{4,3}(\mathbf{R})$, between $E_4^{(\text{CR-CC}(2,3))}(\mathbf{R})$ and the CR-CC(2,3)/aug-cc-pVTZ energies $E_3^{(\text{CR-CC}(2,3))}(\mathbf{R})$. In the following columns, extrapolation errors are reported, in millihartree, relative to the corresponding true $E_4^{(\text{CR-CC}(2,3))}(\mathbf{R})$ values, organized according to the correlation energy scaling $\chi_{4,3}^{(B)}(\mathbf{R})$ that was used in their

generation. Each set of extrapolation errors should be compared to those reported in the far right column of Table (4.9), which represent the errors found using the single-level formulation of the PES extrapolation scheme, appropriately designated by the scaling factor $\chi_{4,3}^{(\text{CR-CC}(2,3))}(\mathbf{R})$ found in the corresponding column heading. While not all lower-order methods can be used to accurately reproduce all points on the CR-CC(2,3)/aug-cc-pVQZ PES, it is notable that relatively small extrapolation errors are observed when all of the lower-order methods (except for MP2) are employed to generate the near-equilibrium region of the $\text{H}_2\text{O} \rightarrow \text{OH} + \text{H}$ curve. This implies that MP3, the various variants of MP4, and CCSD yield correlation energies which have a very similar correlation energy scaling to the CR-CC(2,3) energies in this region. If one is only concerned with the near-equilibrium region of this bond-stretching curve, all of these methods except MP2 can be used successfully to scale a CR-CC(2,3) correlation energy to the quality of a larger basis sets. Unfortunately, the MP3 and various MP4X methods fail to produce the proper scaling at the largest internuclear distances, shown by steeply rising extrapolation errors beyond $2\mathbf{R}_e$. As a specific example, when $\chi_{4,3}^{(\text{MP4SDQ})}(\mathbf{R})$ is used, the largest reported error in the extrapolation of the CR-CC(2,3)/aug-cc-pVQZ PES in the region from $0.75\mathbf{R}_e - 2\mathbf{R}_e$ is only 0.338 millihartree, which represents a recovery of about 98% of the correlation energy change when going from aug-cc-pVTZ to aug-cc-pVQZ, identified in Table (4.9) as $\Delta E_{4,3}^{(\text{CR-CC}(2,3))}$. However, at $4\mathbf{R}_e$, the same error rises sharply to 4.798 millihartree, which is 26% of $\Delta E_{4,3}^{(\text{CR-CC}(2,3))}$ at this geometry. It is quite remarkable, though, to observe such tiny errors while extrapolating the CR-CC(2,3)/aug-cc-pVQZ PES in the relatively large $R = 0.75\mathbf{R}_e - 2\mathbf{R}_e$ region, on the order of small fractions of a millihartree, particularly when we realize that the CPU time of the MP4SDQ calculations is approximately the same as the cost of a single CCSD iteration.

To predict an accurate PES at the highly stretched geometries, it is clear from the extrapolation errors reported in Table (4.9) that CCSD is the only lower-order method which accurately and consistently predicts the correct correlation energy scaling in both the equilibrium and bond-breaking regions of the CR-CC(2,3) surface. When $\chi_{4,3}^{(\text{CCSD})}(\mathbf{R})$ is used, the maximum reported deviation from the true CR-CC(2,3)/aug-cc-pVQZ curve in the $\mathbf{R} = 0.75\mathbf{R}_e - 4\mathbf{R}_e$ region is only 1.124 millihartree, with extrapolation errors in the $R = 0.75\mathbf{R}_e - 2\mathbf{R}_e$ region not exceeding 0.394 millihartree. For comparison, when the CR-CC(2,3) methodology is used to generate the correlation energy scaling factor, as in the single-level scheme, the largest error found for the extrapolated CR-CC(2,3)/aug-cc-pVQZ H₂O PES is 0.760 millihartree in the $R = 0.75\mathbf{R}_e - 2\mathbf{R}_e$ region and 0.314 millihartree when $\mathbf{R}$ does not exceed $2\mathbf{R}$. From these results it is clear that for the H₂O system, $\chi_{4,3}^{(\text{CCSD})}(\mathbf{R})$ and $\chi_{4,3}^{(\text{CR-CC(2,3)})}(\mathbf{R})$ are virtually identical for all reported geometries. The final observation from Table (4.9) is that the MP2 method is found to consistently overestimate $\chi_{4,3}^{(\text{CR-CC(2,3)})}(\mathbf{R})$, producing relatively large negative extrapolation errors on the order of (-4)–(-2) millihartree, even in the near-equilibrium region. Thus, the MP2 method should not be used in conjunction with the dual-level PES extrapolation scheme to describe the scaling of CR-CC(2,3) with the basis set.

Analogous sets of extrapolation errors for the HCl and F₂ bond-stretching surfaces are reported in Tables (4.10) and (4.11), respectively, and similar trends are observed therein. For both the HCl and F₂ systems the MP3 and MP4X methods are again shown to provide the proper correlation energy scaling in the equilibrium region, where large percentages (between 93–100%) of $\Delta E_{4,3}^{(\text{CR-CC(2,3)})}(\mathbf{R})$ are consistently recovered and errors relative to true CR-CC(2,3)/aug-cc-pVQZ calculations are on the order of a millihartree or less. Also,

as before, extrapolation errors grow rapidly as the internuclear separation approaches the $\mathbf{R} > 2\mathbf{R}_e$ region, with CCSD appearing to be the only lower-order method which may be used successfully at all internuclear separations. This is especially evident for the F_2 curve where results based on correlation energy scaling factors obtained with MP2, MP3, MP4D, MP4DQ, and MP4SDQ diverge badly at $4\mathbf{R}_e$ from the true CR-CC(2,3)/aug-cc-pVQZ curve, and while the use of CCSD instead leads to reasonable behavior. In fact, the maximum reported extrapolation errors are again found to be very close at all internuclear distances when the CCSD and CR-CC(2,3) correlation energy scaling factors are used to scale the CR-CC(2,3)/aug-cc-pVTZ HCl and F_2 PESs. For HCl, they are 1.891 and 1.664 millihartree, respectively, each corresponding to a recovery of about 90% of $\Delta E_{4,3}^{(\text{CR-CC}(2,3))}(\mathbf{R})$. For F_2 , extrapolation errors found using $\chi_{4,3}^{(\text{CCSD})}(\mathbf{R})$ show a recovery of 93–99% of the corresponding $\Delta E_{4,3}^{(\text{CR-CC}(2,3))}(\mathbf{R})$ values, which may be compared to 95–100% obtained with $\chi_{4,3}^{(\text{CR-CC}(2,3))}$.

Table 4.9: A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies of the H₂O molecule in which one of the two O-H bonds (**R**) is stretched, while keeping the other O-H bond at the equilibrium length and the H-O-H angle fixed at 104.5°. The equilibrium geometry defines the pivot geometry **R**_e and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest orbital, correlating with the 1s shell of the oxygen atom was kept frozen.

R/R _e ^a	<i>E</i> ₄ ^b	$\Delta E_{4,3}$ ^c	$\epsilon_4(\chi_{4,3}^{(B)})^d$						
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD	CR-CC(2,3)
0.75	-76.259133	-22.676	-2.993	-0.763	-0.611	-0.257	-0.073	-0.186	-0.140
0.90	-76.351620	-21.605	-2.759	-0.573	-0.425	-0.046	0.121	0.031	0.012
1.00	-76.363051	-21.059	-2.742	-0.576	-0.428	-0.033	0.131	0.054	0.000
1.10	-76.355931	-20.641	-2.723	-0.580	-0.430	-0.019	0.145	0.085	0.009
1.25	-76.329620	-20.278	-2.624	-0.509	-0.349	0.080	0.253	0.225	0.133
1.50	-76.276456	-19.952	-2.639	-0.513	-0.316	0.130	0.338	0.394	0.314
2.00	-76.199626	-19.168	-3.613	-1.217	-0.788	-0.359	-0.071	0.294	0.244
3.00	-76.162662	-18.479	-4.086	-0.601	0.729	0.774	1.450	0.912	0.744
4.00	-76.161060	-18.303	-2.232	1.965	3.688	2.953	4.798	1.124	0.760

^aThe equilibrium value of **R** used here is **R**_e = 0.95785 Å. ^bThe calculated CR-CC(2,3)/aug-cc-pVQZ total energies in hartree. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/aug-cc-pVQZ and CR-CC(2,3)/aug-cc-pVTZ energies.

^dDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies, where the latter energies were generated by applying the correlation energy scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with *B* = MP2, MP3, MP4D, MP4DQ, MP4SDQ, CCSD, and CR-CC(2,3). The choice of *B* = CR-CC(2,3) is equivalent to the single-level PES extrapolation scheme.

Table 4.10: A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies for several internuclear separations $R_{\text{H-Cl}}$ of the HCl molecule. The equilibrium geometry defines the pivot geometry $\mathbf{R}_e$ and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest five orbitals, correlating with the 1s, 2s, and 2p shells of Cl, were kept frozen.

$\mathbf{R}/\mathbf{R}_e^a$	E_4^b	$\Delta E_{4,3}^c$	$\epsilon_4(\chi_{4,3}^{(B)})^d$						
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD	CR-CC(2,3)
0.75	-460.245772	-22.053	-4.017	-1.415	-1.254	-0.834	-0.854	-0.896	-1.273
0.90	-460.351216	-20.979	-3.055	-0.537	-0.387	0.033	0.022	-0.003	-0.410
1.00	-460.364178	-20.704	-2.610	-0.127	0.017	0.439	0.437	0.427	0.000
1.10	-460.356856	-20.550	-2.330	0.128	0.268	0.692	0.703	0.713	0.270
1.25	-460.329720	-20.327	-2.178	0.261	0.407	0.834	0.868	0.926	0.469
1.50	-460.277589	-19.823	-2.429	0.065	0.276	0.697	0.793	0.991	0.580
2.00	-460.212229	-18.747	-3.755	-0.828	-0.293	0.055	0.456	0.966	0.730
3.00	-460.192428	-18.345	-2.889	0.849	2.050	1.675	4.602	1.891	1.664
4.00	-460.192532	-17.809	-0.558	2.958	3.966	1.670	9.747	1.573	1.049

^aThe equilibrium value of $\mathbf{R}$ used here is $\mathbf{R}_e = 1.27455 \text{ \AA}$. ^bThe calculated CR-CC(2,3)/aug-cc-pVQZ total energies in hartree. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/aug-cc-pVQZ and CR-CC(2,3)/aug-cc-pVTZ energies. ^dDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies, where the latter energies were generated by applying scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, CCSD, and CR-CC(2,3)}$. The choice of $B=\text{CR-CC(2,3)}$ is equivalent to the single-level extrapolation scheme.

Table 4.11: A comparison of calculated and extrapolated CR-CC(2,3)/aug-ccpVQZ energies for several internuclear separations $R_{\text{F-F}}$ of the F_2 molecule. The equilibrium geometry defines the pivot geometry $\mathbf{R}_e$ and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest two orbitals, correlating with the 1s shells of the F atoms, were kept frozen.

$\mathbf{R}/\mathbf{R}_e^a$	E_4^b	$\Delta E_{4,3}^c$	$\epsilon_4(\chi_{4,3}^{(B)})^d$						
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD	CR-CC(2, 3)
0.75	-199.196897	-56.537	-5.098	-2.516	-1.915	-1.208	-0.689	-1.120	-1.011
0.90	-199.349657	-53.009	-3.991	-1.752	-1.132	-0.444	0.065	-0.197	-0.258
1.00	-199.364732	-51.743	-3.835	-1.737	-1.028	-0.378	0.191	0.036	0.000
1.10	-199.356794	-51.172	-3.871	-1.793	-0.929	-0.349	0.284	0.236	0.227
1.25	-199.333764	-50.464	-4.209	-2.010	-0.775	-0.389	0.313	0.396	0.455
1.50	-199.307123	-49.741	-4.370	-1.655	0.587	0.339	1.119	0.944	0.965
2.00	-199.296727	-49.300	-1.232	3.454	8.189	5.241	6.107	2.453	2.146
3.00	-199.297648	-49.368	6.748	13.398	18.543	0.429	4.979	3.239	2.460
4.00	-199.297985	-49.422	12.149	18.901	22.880	30.269	31.933	3.585	2.673

^aThe equilibrium value of $\mathbf{R}$ used here is $\mathbf{R}_e = 0.988351 \text{ \AA}$. ^bThe calculated CR-CC(2,3)/aug-cc-pVQZ total energies in hartree. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/aug-cc-pVQZ and CR-CC(2,3)/aug-cc-pVTZ energies. ^dDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies, where the latter energies were generated by applying the correlation energy scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, CCSD, and CR-CC(2,3)}$. The choice of $B = \text{CR-CC(2,3)}$ is equivalent to the single-level PES extrapolation scheme.

In summary, MP2 consistently fails to predict the proper information about the CR-CC(2,3) correlation energy scaling, while MP3 typically does much better. This leads to the conclusion that at least third-order correlation energy effects must be included for the proper description of the correlation energy scaling in the near-equilibrium PES regions. In regions further away from equilibrium, the third and even the partial fourth-order perturbative corrections still cannot produce the proper correlation energy scaling factors that would be compatible with those of CR-CC(2,3), which is a consequence of the inability of MBPT to describe bond-breaking, but one can use the CCSD approach instead, which is qualitatively correct at larger internuclear separations in single-bond breaking situations, providing a reasonable estimate of the CR-CC(2,3) correlation energy scaling with the basis set.

4.4.3 Application to the Isomerization of Bicyclobutane to Butadiene

The results of dual-level PES extrapolations on the reaction profiles for the isomerization of bicyclobutane to butadiene are given in Table (4.12). The format of this table is similar to that of Tables (4.9)–(4.11), except that here the geometries of interest are the stationary points along the conrotatory and disrotatory pathways shown in Figure (3.2) and the cc-pVXZ rather than aug-cc-pVXZ basis sets are employed throughout. From the results presented in Table (4.12) it can be seen that the MP3 and various MP4 methods may be successfully used to probe the CR-CC(2,3) correlation energy scaling for the energetically favored conrotatory reaction profile, which consists entirely of species with correlation energy dominated by lower-order excitations, but in every case the MBPT methods produce a significantly larger error for the highly biradical dis_TS geometry. On the other hand, when

$\chi_{4,3}^{(\text{CCSD})}(\mathbf{R})$ is used to extrapolate the CR-CC(2,3)/cc-pVQZ PES, the reported extrapolation errors remain within a millihartree of the explicitly calculated CR-CC(2,3)/cc-pVQZ energy values at every stationary point, rivaling the sub-millihartree accuracies found using the single-level PES extrapolation scheme where $\chi_{4,3}^{(\text{CR-CC}(2,3))}(\mathbf{R})$ is employed instead. This is another case where a quasi-degeneracy, in this case resulting from the biradical nature of the **dis_TS** configuration, inhibits the MBPT methods from producing the correct correlation energy scaling information. It is clear, of the methods considered here, that only the CCSD approach can offer a correlation energy scaling factor compatible with CR-CC(2,3), although the extrapolation errors obtained with MP4SDQ, which are 1.950 millihartree or ~ 1.5 kcal/mol for the strongly biradical **dis_TS** structure (located over 60 kcal/mol above the reactant) and less than 1 millihartree for the remaining structures, are excellent as well. Also, once again, MP2 correlation energies do not contain sufficient information to model the CR-CC(2,3) correlation energy scaling at any geometry.

Table 4.12: A comparison of calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the **bicbut**→**t-but** isomerization. The **bicbut** reactant defines the pivot geometry $\mathbf{R}_e$ and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest four orbitals, correlating with the 1s shells of the carbon atoms, were kept frozen.

Structure	E_4^a	$\Delta E_{4,3}^b$	$\epsilon_4(\chi_{4,3}^{(B)})^c$						
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD	CR-CC(2, 3)
bicbut	-155.695497	-44.149	-9.521	-0.860	-0.397	0.845	0.760	0.582	0.000
con_TS	-155.625392	-43.727	-10.169	-1.578	-0.993	0.292	0.397	0.466	0.245
dis_TS	-155.582481	-42.447	-13.436	-4.643	-3.890	-2.671	-1.950	0.096	0.631
<i>g</i> -but	-155.733236	-43.909	-9.730	-0.879	-0.325	0.899	0.906	0.927	0.482
<i>gt</i> _TS	-155.728073	-43.764	-9.815	-0.934	-0.374	0.844	0.837	0.876	0.410
<i>t</i> -but	-155.738043	-43.963	-9.614	-0.786	-0.234	0.972	0.987	1.008	0.563

^aThe calculated CR-CC(2,3)/cc-pVQZ total energies in hartree. ^bDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^cDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies, where the latter energies were generated by applying the correlation energy scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, CCSD, and CR-CC(2,3)}$. The choice of $B = \text{CR-CC(2,3)}$ is equivalent to the single-level extrapolation scheme.

4.4.4 The Role of Pivot Geometries and Base Wave Functions in Dual-Level PES Extrapolations

In Tables (4.9)–(4.12) it was clear that regardless of the correlation energy scaling factor used in the PES extrapolation scheme based on Eqs. (4.1) – (4.3), the extrapolation errors are largest in regions where the nature of the electron correlation effects differ most from those found at the pivot geometry. To see how much the reported errors could be reduced by employing additional pivot geometries, the same sets of PES extrapolations were considered, but this time with all nuclear configurations treated as pivot geometries. In addition to introducing a new extrapolation method which should yield improved accuracies, this approach provides a direct measure of the error introduced when lower-order correlation energy scaling factors are employed to predict the results of higher-order calculations with a larger basis set, since any error due to the earlier assumption of the approximate transferability of the scaling factor from one geometry to another is eliminated. Additional calculations required to perform these extrapolations, when compared to those required for the extrapolations of the previous section (done with a single pivot geometry), consist of the remaining calculations required to obtain each PES using the aug-cc-pVQZ basis for the bond-breaking curves or the cc-pVQZ basis for the **bicbut**→*t*-**but** isomerization pathways using the lower-level methodology *B*. This is still relatively inexpensive, since we never have to perform the high-level CR-CC(2,3) calculation with the largest basis set employed at any geometry.

A comparison of Tables (4.9) and (4.13), (4.10) and (4.14), (4.11) and (4.15), and (4.12) and (4.16) demonstrates the full extent of the accuracy which may be gained by using additional pivot geometries for each of the systems considered in this study. The most interesting detail to note in these comparisons is that when the same method is used to

predict the correlation energy scaling factor, increasing the number of pivot geometries does not necessarily result in an overall improvement in the quality of the extrapolated surface. Extrapolation errors are, for the most part, improved when the number of pivot geometries is increased, especially when MP4SDQ and CCSD are used to generate χ , but there are no significant benefits from switching from the previously discussed single-point approach to its multi-point analog.

As stated before, the many options inherently included in the PES extrapolation scheme based on Eqs. (4.1) – (4.3) allow one to tailor it to make it more accurate or more affordable, as required by a given application. To demonstrate clearly and concisely the different levels of accuracy and savings in the computer effort which may be obtained using different tiers of the PES extrapolation scheme, the results of four different PES extrapolations examined here, using the **bicbut** $\rightarrow$ **t-but** isomerization as an example, are collected in Table (4.17). We recall that the goal of each extrapolation in this case is to predict the CR-CC(2,3)/cc-pVQZ PES (as represented by six structures on the corresponding conrotatory and disrotatory pathways) from the results of lower-level calculations. The most efficient PES extrapolation considered for this table, where a lower-order MP4SDQ scaling factor is used to extrapolate the CR-CC(2,3)/cc-pVTZ PES to the level of CR-CC(2,3)/cc-pVQZ calculations using one pivot geometry (**bicbut**) and RHF base energies, requires only 5% of the CPU time of the conventional method and the mean unsigned error (MUE) is slightly below 1 millihartree. The MUE is reduced by 0.314 millihartree when the correlation energy scaling factor obtained with MP4SDQ is replaced by that produced by CCSD, but the required CPU time is also increased to 10% of that required by the conventional CR-CC(2,3)/cc-pVQZ calculations. The results obtained with the original single-level PES extrapolation procedure are also

presented, where the time required to perform these calculations is 22% of that required conventionally, but the mean unsigned error falls to below 0.4 millihartree. Finally, when the base energy is obtained in CCSD calculations and $\chi_{4,3}^{(\text{CR-CC}(2,3))}(\mathbf{R})$ is used to scale the remaining correlation energy, an MUE of 0.170 is attained, but now the computational time savings amounts only to about one third of the conventional procedure. It can be seen in Table (4.17) that regardless of the PES extrapolation approach used, the MUE remains below 1 millihartree, which is a relatively insignificant loss in accuracy compared to the conventional, point-wise CR-CC(2,3)/cc-pVQZ calculations. By using the dual-level PES extrapolation scheme to extrapolate CR-CC(2,3)/cc-pVQZ energies, we have reduced the time required to accurately construct a PES for this problematic polyatomic isomerization by more than an order of magnitude. These computational savings would only grow larger if another system were considered which contained a greater number of electrons or a larger number of points on the PES were considered.

Table 4.13: A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies of the H₂O molecule, in which one of the two O-H bonds $\mathbf{R}$ is stretched, while keeping the other O-H bond at the equilibrium length and the H-O-H angle fixed at 104.5°. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest orbital, correlating with the 1s orbital of the oxygen atom, was kept frozen.

$\mathbf{R}/\mathbf{R}_e^a$	E_4^b	$\Delta E_{4,3}^c$	$\epsilon_4(\chi_{4,3}^{(B)})^d$					
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD
0.75	-76.259133	-22.676	-2.766	-0.504	-0.360	-0.014	0.129	0.059
0.90	-76.351620	-21.605	-2.737	-0.534	-0.391	-0.016	0.138	0.063
1.00	-76.363051	-21.059	-2.742	-0.576	-0.428	-0.033	0.131	0.054
1.10	-76.355931	-20.641	-2.769	-0.636	-0.475	-0.060	0.114	0.041
1.25	-76.329620	-20.278	-2.835	-0.733	-0.538	-0.096	0.097	0.034
1.50	-76.276456	-19.952	-3.072	-0.951	-0.668	-0.190	0.033	0.021
2.00	-76.199626	-19.168	-4.107	-1.651	-1.073	-0.591	-0.301	-0.080
3.00	-76.162662	-18.479	-5.201	-1.501	0.009	0.207	0.856	0.101
4.00	-76.161060	-18.303	-3.619	1.054	3.143	2.604	4.400	0.164

^aThe equilibrium value of $\mathbf{R}$ used here is $\mathbf{R}_e = 0.95785 \text{ \AA}$. ^bThe calculated CR-CC(2,3)/aug-cc-pVQZ total energies in hartree. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/aug-cc-pVQZ and CR-CC(2,3)/aug-cc-pVTZ energies. ^dDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies, with the latter energies generated by applying the correlation energy scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, and CCSD}$.

Table 4.14: A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies for several internuclear separations $\mathbf{R}$ of the HCl molecule. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest five orbitals, correlating with the 1s, 2s, and 2p shells of Cl, were kept frozen.

$\mathbf{R}/\mathbf{R}_e^a$	E_4^b	$\Delta E_{4,3}^c$	$\epsilon_4(\chi_{4,3}^{(B)})^d$					
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD
0.75	-460.245772	-22.053	-2.707	-0.044	0.123	0.516	0.491	0.447
0.90	-460.351216	-20.979	-2.625	-0.087	0.056	0.466	0.453	0.424
1.00	-460.364178	-20.704	-2.610	-0.127	0.017	0.439	0.437	0.427
1.10	-460.356856	-20.550	-2.680	-0.220	-0.062	0.373	0.384	0.401
1.25	-460.329720	-20.327	-2.885	-0.404	-0.204	0.249	0.284	0.362
1.50	-460.277589	-19.823	-3.370	-0.739	-0.426	0.046	0.146	0.386
2.00	-460.212229	-18.747	-5.151	-1.906	-1.204	-0.736	-0.265	0.336
3.00	-460.192428	-18.345	-5.225	-0.655	0.927	0.985	4.375	0.933
4.00	-460.192532	-17.809	-2.748	2.016	3.643	1.988	11.510	0.533

^aThe equilibrium value of $\mathbf{R}$ used here is $\mathbf{R}_e = 1.27455 \text{ \AA}$. ^bThe calculated CR-CC(2,3)/aug-cc-pVQZ total energies in hartree. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/aug-cc-pVQZ and CR-CC(2,3)/aug-cc-pVTZ energies. ^dDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies, with the latter energies generated by applying the correlation energy scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, and CCSD}$.

Table 4.15: A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies for several internuclear separations $\mathbf{R}$ of the F_2 molecule. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest two orbitals, correlating with the 1s orbitals on the fluorine atoms, were kept frozen.

$\mathbf{R}/\mathbf{R}_e^a$	E_4^b	$\Delta E_{4,3}^c$	$\epsilon_4(\chi_{4,3}^{(B)})^d$					
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD
0.75	-199.196897	-56.537	-3.415	-1.068	-0.764	-0.128	0.310	0.190
0.90	-199.349657	-53.009	-3.570	-1.389	-0.877	-0.215	0.290	0.149
1.00	-199.364732	-51.743	-3.835	-1.737	-1.028	-0.378	0.191	0.036
1.10	-199.356794	-51.172	-4.116	-2.030	-1.063	-0.466	0.176	0.009
1.25	-199.333764	-50.464	-4.758	-2.539	-1.062	-0.646	0.107	-0.115
1.50	-199.307122	-49.741	-5.496	-2.676	-0.019	-0.237	0.664	-0.119
2.00	-199.296727	-49.300	-3.680	1.008	6.381	3.741	4.751	0.250
3.00	-199.297648	-49.368	4.354	11.655	18.778	-4.021	2.203	0.723
4.00	-199.297985	-49.422	9.504	17.002	22.892	30.905	32.328	0.876

^aThe equilibrium value of $\mathbf{R}$ used here is $\mathbf{R}_e = 0.988351 \text{ \AA}$. ^bThe calculated CR-CC(2,3)/aug-cc-pVQZ total energies in hartree. ^cDifferences, in millihartree, between the actual CR-CC(2,3)/aug-cc-pVQZ and CR-CC(2,3)/aug-cc-pVTZ energies. ^dDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies, with the latter energies generated by applying scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, and CCSD}$.

Table 4.16: A comparison of calculated and extrapolated CR-CC(2,3)/aug-cc-pVQZ energies at the stationary points defining the conrotatory and disrotatory pathways characterizing the bicbut→t-but isomerization. Each geometry serves as its own pivot geometry and RHF defines the base wave function for the PES extrapolations. In all post-RHF calculations, the lowest four orbitals, correlating with the 1s orbitals of the carbon atoms, were kept frozen.

Structure	E_4^a	$\Delta E_{4,3}^b$	$\epsilon_4(\chi_{4,3}^{(B)})^c$					
			MP2	MP3	MP4D	MP4DQ	MP4SDQ	CCSD
bicbut	-155.695497	-44.149	-9.521	-0.860	-0.397	0.845	0.760	0.582
con_TS	-155.625392	-43.727	-10.455	-1.614	-1.036	0.322	0.373	0.295
dis_TS	-155.582481	-42.447	-14.074	-4.825	-4.164	-2.824	-2.131	-0.529
gbut	-155.733236	-43.909	-10.245	-1.009	-0.473	0.781	0.728	0.699
gt_TS	-155.728073	-43.764	-10.366	-1.085	-0.548	0.702	0.652	0.641
tbut	-155.738043	-43.963	-10.220	-1.007	-0.472	0.776	0.729	0.703

^aThe calculated CR-CC(2,3)/cc-pVQZ total energies in hartree. ^bDifferences, in millihartree, between the actual CR-CC(2,3)/cc-pVQZ and CR-CC(2,3)/cc-pVTZ energies. ^cDifferences, in millihartree, between the calculated and extrapolated CR-CC(2,3)/cc-pVQZ energies, with the latter energies generated by applying the correlation energy scaling factors $\chi_{4,3}^{(B)}(\mathbf{R})$ obtained with $B = \text{MP2, MP3, MP4D, MP4DQ, MP4SDQ, and CCSD}$.

Table 4.17: A summary of the necessary calculations and the corresponding computer resources required to utilize different tiers of the PES extrapolation scheme based on Eqs. (4.1) – (4.3) to scale the bicyclobutane isomerization pathway from the CR-CC(2,3)/cc-pVTZ level of theory to the CR-CC(2,3)/cc-pVQZ level, along with the corresponding extrapolation errors. The **bicbut** structure is used to provide the pivot geometry.

Base Energy ^a	Correlation Energy Scaling Factor ^b	cc-pVQZ Calculations Required ^c			CPU Time (t/t_{conv}) ^d	Mean Unsigned Error (millihartree) ^e
		RHF	CCSD	CR-CC(2,3)		
RHF	MP4SDQ	6	0	0	0.05	0.973
RHF	CCSD	6	1	0	0.10	0.659
RHF	CR-CC(2,3)	6	1	1	0.22	0.389
CCSD	CR-CC(2,3)	6	6	1	0.63	0.170
Conventional Calculation ^f		6	6	6	1	—

^aThe method used to generate the base energy. ^bThe method B used to generate the correlation energy scaling factor $\chi_{4,3}^{(B)}(\mathbf{R})$ in Eq. (4.1). ^cThe number of cc-pVQZ basis set calculations which must be performed using a given base wave function and a given correlation energy scaling factor to extrapolate the CR-CC(2,3)/cc-pVQZ PES. ^dThe CPU time needed to perform the necessary calculations for each PES extrapolation type relative to the time needed to generate the true CR-CC(2,3)/cc-pVQZ PES. ^eThe mean unsigned error representing an extrapolated CR-CC(2,3)/cc-pVQZ reaction pathway generated using the designated base energy and correlation energy scaling factor. ^fCharacteristics of the true PES calculation, in which each stationary point on the conrotatory and disrotatory pathways of the **bicbut** $\rightarrow$ **t-but** isomerization is calculated at the CR-CC(2,3)/cc-pVQZ level.

Chapter 5

Development of Computer Codes for the GAMESS software Package

A large portion of the current doctoral effort was devoted to writing EOMCC computer codes for the GAMESS software package. In this section, we discuss the highly efficient GAMESS implementations of the open-shell EOMCCSD and IP-EOMCCSD($2h-1p$) methods developed as part of this thesis project, based on theory discussed in Sect. (3.1.2) and the corresponding factorized equations, in terms of recursively generated intermediates that lead to the vectorized computer codes through the use of fast matrix multiplication routines from the BLAS library. The open-shell EOMCCSD and IP-EOMCCSD($2h-1p$) codes were interfaced with previously existing ROHF and RHF/ROHF integral routines, respectively, available in the GAMESS software package [194], as well as the CC programs and routines for the generation of matrix elements of the similarity-transformed Hamiltonian of CCSD originally developed for GAMESS by the Piecuch group at Michigan State University. In Sect. (5.1), we begin our discussion of the implementation of these programs, with specific

details outlined for the open-shell EOMCCSD and IP-EOMCCSD($2h-1p$) methods in Sects. (5.1.1) and (5.1.2), respectively.

5.1 Key Details of Efficient Computer Implementations

Both the open-shell EOMCCSD and IP-EOMCCSD($2h-1p$) codes must begin by solving the usual CCSD equations for the ground-state of the N -electron reference system in order to obtain the singly and doubly excited cluster amplitudes, t_a^i and t_{ab}^{ij} , respectively. In both cases, this is done using the general ROHF-based CCSD codes included in GAMESS that work for closed- and open-shells, developed by the Piecuch group and described in [130].

Following the CCSD calculation, the converged t_a^i and t_{ab}^{ij} amplitudes are used to construct the one- and two-body matrix elements of the CCSD similarity-transformed Hamiltonian $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$, $\bar{h}_p^q$ and $\bar{h}_{pq}^{rs}$, respectively, which define the one- and two-body components of $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$ within the second quantized formalism,

$$\bar{H}_1^{(\text{CCSD})} = \bar{h}_p^q a^p a_q, \quad (5.1)$$

and

$$\bar{H}_2^{(\text{CCSD})} = \frac{1}{4} \bar{h}_{pq}^{rs} N[a^p a^q a_s a_r], \quad (5.2)$$

respectively. $N[\dots]$ is the normal product of the operators between the brackets and the Einstein summation convention over repeated upper and lower indices is assumed throughout. The explicit equations defining these matrix elements in terms of the matrix elements of the Hamiltonian in the normal-ordered form $f_p^q = \langle p|f|q\rangle$ and $v_{pq}^{rs} = \langle pq|v|rs\rangle - \langle pq|v|sr\rangle$, and CCSD cluster amplitudes t_a^i and t_{ab}^{ij} , are given in Table (5.1). Once these common

initial steps are complete, the appropriate expressions for solving the EOMCCSD and IP-EOMCCSD($2h-1p$) eigenvalue problems have to be constructed. These steps, which are specific to the open-shell EOMCCSD and IP-EOMCCSD($2h-1p$) codes considered here, are outlined in Sects. (5.1.1) and (5.1.2), respectively. Once the suitable EOMCCSD and IP-EOMCCSD($2h-1p$) equations are constructed, another common feature in our implementation of the EOMCCSD and IP-EOMCCSD($2h-1p$) approaches is that we rely on the Hirao-Nakatsuji generalization [303] of the Davidson diagonalization algorithm [304] to solve the resulting non-Hermitian eigenvalue problems, Eqs. (3.24) and (3.25) in Sect. (3.1.2.1).

5.1.1 Standard Equation-of-Motion Coupled-Cluster Theory with Singles and Doubles for Open-Shell Systems

The left-hand sides of the EOMCCSD equations are calculated by projecting $[\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C|\Phi\rangle$ onto the subspace of all singly- and doubly-excited determinants, $\langle\Phi_{i\alpha}^{a\alpha}|$, $\langle\Phi_{i\beta}^{a\beta}|$, $\langle\Phi_{i\alpha j\alpha}^{a\alpha b\alpha}|$, and $\langle\Phi_{i\alpha j\beta}^{a\alpha b\beta}|$, to obtain the following expressions:

$$\begin{aligned}
\langle\Phi_{i\alpha}^{a\alpha}|[\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C|\Phi\rangle &= \bar{h}_{a\alpha}^{e\alpha} r_{e\alpha}^{i\alpha} - \bar{h}_{m\alpha}^{i\alpha} r_{a\alpha}^{m\alpha} + \bar{h}_{m\alpha}^{e\alpha} r_{a\alpha e\alpha}^{i\alpha m\alpha} \\
&\quad + \bar{h}_{m\beta}^{e\beta} r_{a\beta e\beta}^{i\alpha m\alpha} + \bar{h}_{a\alpha m\alpha}^{i\alpha e\alpha} r_{e\alpha}^{m\alpha} + \bar{h}_{a\alpha m\beta}^{i\alpha e\beta} r_{e\beta}^{m\beta} \\
&\quad - \frac{1}{2} \bar{h}_{m\alpha n\alpha}^{i\alpha f\alpha} r_{a\alpha f\alpha}^{m\alpha n\alpha} + \frac{1}{2} \bar{h}_{a\alpha n\alpha}^{e\alpha f\alpha} r_{e\alpha f\alpha}^{i\alpha n\alpha} \\
&\quad - \bar{h}_{m\alpha n\beta}^{i\alpha f\beta} r_{a\alpha f\beta}^{m\alpha n\beta} + \bar{h}_{a\alpha n\beta}^{e\alpha f\beta} r_{e\alpha f\beta}^{i\alpha n\beta}, \tag{5.3}
\end{aligned}$$

Table 5.1: Explicit algebraic expressions for the one- and two-body matrix elements of $\bar{H}_{N,\text{open}}^{(\text{CCSD})}$ ($\bar{h}_p^q$ and $\bar{h}_{pq}^{rs}$, respectively) taken from Refs. [126, 132].

Intermediate	Expression ^a
$\bar{h}_i^a$	$f_i^a + v_{im}^{ae} t_e^m$
$\bar{h}_i^j$	$f_i^j + v_{im}^{je} t_e^m + \frac{1}{2} v_{mi}^{ef} t_{ef}^{mj} + \bar{h}_i^e t_e^j$
$\bar{h}_a^b$	$I_a^b - \bar{h}_m^b t_a^m$
$\bar{h}_{ai}^{bc}$	$v_{ai}^{bc} - v_{mi}^{bc} t_a^m$
$\bar{h}_{ij}^{ka}$	$v_{ij}^{ka} + v_{ij}^{ea} t_e^k$
$\bar{h}_{ab}^{cd}$	$v_{ab}^{cd} + \frac{1}{2} v_{mn}^{cd} t_{ab}^{mn} - \bar{h}_{am}^{cd} t_b^m + v_{bm}^{cd} t_a^m$
$\bar{h}_{ij}^{kl}$	$v_{ij}^{kl} + \frac{1}{2} v_{ij}^{ef} t_{ef}^{kl} - \bar{h}_{ij}^{le} t_e^k + v_{ij}^{ke} t_e^l$
$\bar{h}_{ia}^{jb}$	$I_{ia}^{jb} - v_{im}^{eb} t_{ea}^{jm} - \bar{h}_{im}^{jb} t_a^m$
$\bar{h}_{ab}^{ic}$	$v_{ab}^{ic} + v_{ab}^{ec} t_e^i - \bar{h}_{mb}^{ic} t_a^m + I_{ma}^{ic} t_b^m - \bar{h}_m^c t_{ab}^{im} + \bar{h}_{bm}^{ce} t_{ae}^{im} - v_{am}^{ce} t_{be}^{im} + \frac{1}{2} \bar{h}_{nm}^{ic} t_{ab}^{nm}$
$\bar{h}_{ia}^{jk}$	$v_{ia}^{jk} + \bar{h}_{mi}^{jk} t_a^m - v_{ia}^{ke} t_e^j + \mathcal{A}^{jk} \bar{h}_{im}^{je} t_{ae}^{km} + \bar{h}_i^{ej} t_{ea}^{jk} + I_{ia}^{je} t_e^k - \frac{1}{2} v_{ai}^{ef} t_{ef}^{jk}$
I_a^{tb}	$f_a^b + v_{am}^{be} t_e^m$
I_a^b	$I_a^b - \frac{1}{2} v_{mn}^{eb} t_{ea}^{mn}$
I_{ia}^{jb}	$v_{ia}^{jb} + v_{ia}^{eb} t_e^j$

^a Summation over repeated upper and lower indices is assumed. $f_p^q = \langle p|f|q \rangle$ and $v_{pq}^{rs} = \langle pq|v|rs \rangle - \langle pq|v|sr \rangle$ are the one- and two-body matrix elements of the Hamiltonian in the normal-ordered form (one- and two-electron integrals), and the t_a^i and t_{ab}^{ij} are the singly and doubly excited cluster amplitudes defining the ground-state CCSD wave function of the N -electron reference system. The antisymmetrizer $\mathcal{A}_{pq} = 1 - (pq)$ operator is also used, where (pq) is the transposition of indices p and q .

$$\begin{aligned}
\langle \Phi_{i\beta}^{a\beta} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle &= \bar{h}_{a\beta}^{e\beta} r_{e\beta}^{i\beta} - \bar{h}_{m\beta}^{i\beta} r_{a\beta}^{m\beta} + \bar{h}_{m\beta}^{e\beta} r_{a\beta}^{i\beta m\beta} \\
&+ \bar{h}_{m\alpha}^{e\alpha} r_{a\beta e\alpha}^{i\beta m\alpha} + \bar{h}_{a\beta m\beta}^{i\beta e\beta} r_{e\beta}^{m\beta} + \bar{h}_{a\beta m\alpha}^{i\beta e\alpha} r_{e\alpha}^{m\alpha} \\
&- \frac{1}{2} \bar{h}_{m\beta n\beta}^{i\beta f\beta} r_{a\beta f\beta}^{m\beta n\beta} + \frac{1}{2} \bar{h}_{a\beta n\beta}^{e\beta f\beta} r_{e\beta f\beta}^{i\beta n\beta} \\
&- \bar{h}_{m\beta n\alpha}^{i\beta f\alpha} r_{a\beta f\alpha}^{m\beta n\alpha} + \bar{h}_{a\beta n\alpha}^{e\beta f\alpha} r_{e\beta f\alpha}^{i\beta n\alpha}, \tag{5.4}
\end{aligned}$$

$$\begin{aligned}
\langle \Phi_{i\alpha j\alpha}^{a\alpha b\alpha} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle &= -\mathcal{A}_{i\alpha j\alpha} \bar{h}_{m\alpha}^{i\alpha} r_{a\alpha b\alpha}^{m\alpha j\alpha} + \mathcal{A}_{a\alpha b\alpha} \bar{h}_{a\alpha}^{e\alpha} r_{e\alpha b\alpha}^{i\alpha j\alpha} \\
&- \mathcal{A}_{a\alpha b\alpha} \bar{h}_{m\alpha b\alpha}^{i\alpha j\alpha} r_{a\alpha}^{m\alpha} + \mathcal{A}_{i\alpha j\alpha} \bar{h}_{a\alpha b\alpha}^{e\alpha j\alpha} r_{e\alpha}^{i\alpha} \\
&+ \frac{1}{2} \bar{h}_{a\alpha b\alpha}^{e\alpha f\alpha} r_{e\alpha f\alpha}^{i\alpha j\alpha} + \frac{1}{2} \bar{h}_{m\alpha n\alpha}^{i\alpha j\alpha} r_{a\alpha b\alpha}^{m\alpha n\alpha} \\
&- \mathcal{A}_{i\alpha j\alpha} \mathcal{A}_{ab} \bar{h}_{m\alpha a\alpha}^{i\alpha e\alpha} r_{e\alpha b\alpha}^{m\alpha j\alpha} \\
&+ \mathcal{A}_{i\alpha j\alpha} \mathcal{A}_{ab} \bar{h}_{a\alpha m\beta}^{i\alpha e\beta} r_{e\beta b\alpha}^{m\beta j\alpha} \\
&- \bar{h}_{m\alpha a\alpha b\alpha}^{i\alpha e\alpha j\alpha} r_{e\alpha}^{m\alpha} + \bar{h}_{a\alpha b\alpha m\beta}^{i\alpha j\alpha e\beta} r_{e\beta}^{m\beta} \\
&- \frac{1}{2} \mathcal{A}_{a\alpha b\alpha} \bar{h}_{m\alpha b\alpha n\alpha}^{i\alpha j\alpha f\alpha} r_{a\alpha f\alpha}^{m\alpha n\alpha} \\
&- \mathcal{A}_{a\alpha b\alpha} \bar{h}_{m\alpha b\alpha n\beta}^{i\alpha j\alpha f\beta} r_{a\alpha f\beta}^{m\alpha n\beta} \\
&+ \frac{1}{2} \mathcal{A}_{i\alpha j\alpha} \bar{h}_{a\alpha b\alpha n\alpha}^{e\alpha j\alpha f\alpha} r_{e\alpha f\alpha}^{i\alpha n\alpha} \\
&- \mathcal{A}_{i\alpha j\alpha} \bar{h}_{a\alpha b\alpha n\beta}^{e\alpha j\alpha f\beta} r_{e\alpha f\beta}^{i\alpha n\beta}, \tag{5.5}
\end{aligned}$$

$$\begin{aligned}
\langle \Phi_{i\beta j\beta}^{a\beta b\beta} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle = & -\mathcal{A}_{i\beta j\beta} \bar{h}_{m\beta}^{i\beta} r_{a\beta b\beta}^{m\beta j\beta} + \mathcal{A}_{a\beta b\beta} \bar{h}_{a\beta}^{e\beta} r_{e\beta b\beta}^{i\beta j\beta} \\
& -\mathcal{A}_{a\beta b\beta} \bar{h}_{m\beta b\beta}^{i\beta j\beta} r_{a\beta}^{m\beta} + \mathcal{A}_{i\beta j\beta} \bar{h}_{a\beta b\beta}^{e\beta j\beta} r_{e\beta}^{i\beta} \\
& + \frac{1}{2} \bar{h}_{a\beta b\beta}^{e\beta f\beta} r_{e\beta f\beta}^{i\beta j\beta} + \frac{1}{2} \bar{h}_{m\beta n\beta}^{i\beta j\beta} r_{a\beta b\beta}^{m\beta n\beta} \\
& -\mathcal{A}_{a\beta b\beta} \mathcal{A}_{i\beta j\beta} \bar{h}_{m\beta a\beta}^{i\beta e\beta} r_{e\beta b\beta}^{m\beta j\beta} \\
& + \mathcal{A}_{a\beta b\beta} \mathcal{A}_{i\beta j\beta} \bar{h}_{a\beta m\alpha}^{i\beta e\alpha} r_{e\alpha b\beta}^{m\alpha j\beta} \\
& -\bar{h}_{m\beta a\beta b\beta}^{i\beta e\beta j\beta} r_{e\beta}^{m\beta} + \bar{h}_{a\beta b\beta m\alpha}^{i\beta j\beta e\alpha} r_{e\alpha}^{m\alpha} \\
& -\frac{1}{2} \mathcal{A}_{a\beta b\beta} \bar{h}_{m\beta b\beta n\beta}^{i\beta j\beta f\beta} r_{a\beta f\beta}^{m\beta n\beta} \\
& -\mathcal{A}_{a\beta b\beta} \bar{h}_{m\beta b\beta n\alpha}^{i\beta j\beta f\alpha} r_{a\beta f\alpha}^{m\beta n\alpha} \\
& + \frac{1}{2} \mathcal{A}_{i\beta j\beta} \bar{h}_{a\beta b\beta n\beta}^{e\beta j\beta f\beta} r_{e\beta f\beta}^{i\beta n\beta} \\
& -\mathcal{A}_{i\beta j\beta} \bar{h}_{a\beta b\beta n\alpha}^{e\beta j\beta f\alpha} r_{e\beta f\alpha}^{i\beta n\alpha}, \tag{5.6}
\end{aligned}$$

and

$$\begin{aligned}
\langle \Phi_{i\alpha j\beta}^{a\alpha b\beta} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle = & -\bar{h}_{m\alpha}^{i\alpha} r_{a\alpha b\beta}^{m\alpha j\beta} - \bar{h}_{n\beta}^{j\beta} r_{a\alpha b\beta}^{i\alpha n\beta} + \bar{h}_{a\alpha}^{e\alpha} r_{e\alpha b\beta}^{i\alpha j\beta} \\
& + \bar{h}_{b\beta}^{f\beta} r_{a\alpha f\beta}^{i\alpha j\beta} - \bar{h}_{m\alpha b\beta}^{i\alpha j\beta} r_{a\alpha}^{m\alpha} - \bar{h}_{a\alpha n\beta}^{i\alpha j\beta} r_{b\beta}^{n\beta} \\
& + \bar{h}_{a\alpha b\beta}^{e\alpha j\beta} r_{e\alpha}^{i\alpha} + \bar{h}_{a\alpha b\beta}^{i\alpha f\beta} r_{f\beta}^{j\beta} \\
& + \bar{h}_{a\alpha b\beta}^{e\alpha f\beta} r_{e\alpha f\beta}^{i\alpha j\beta} + \bar{h}_{m\alpha n\beta}^{i\alpha j\beta} r_{a\alpha b\beta}^{m\alpha n\beta} \\
& - \bar{h}_{a\alpha n\beta}^{e\alpha j\beta} r_{e\alpha b\beta}^{i\alpha n\beta} - \bar{h}_{m\alpha b\beta}^{i\alpha f\beta} r_{a\alpha f\beta}^{m\alpha j\beta} \\
& + \bar{h}_{a\alpha n\beta}^{i\alpha f\beta} r_{f\beta b\beta}^{n\beta j\beta} + \bar{h}_{m\alpha b\beta}^{e\alpha j\beta} r_{e\alpha a\alpha}^{m\alpha i\alpha} \\
& - \bar{h}_{m\alpha a\beta b\beta}^{i\alpha e\alpha j\beta} r_{e\alpha}^{m\alpha} - \bar{h}_{a\alpha n\beta b\beta}^{i\alpha j\beta f\beta} r_{f\beta}^{n\beta} \\
& - \bar{h}_{m\alpha a\alpha}^{i\alpha e\alpha} r_{e\alpha b\beta}^{m\alpha j\beta} - \bar{h}_{b\beta n\beta}^{f\beta j\beta} r_{a\alpha f\beta}^{i\alpha n\beta} \\
& + \bar{h}_{a\alpha b\beta m\alpha}^{i\alpha j\beta e\alpha} r_{e\alpha}^{m\alpha} + \bar{h}_{a\alpha b\beta n\beta}^{i\alpha j\beta f\beta} r_{f\beta}^{n\beta} \\
& - \frac{1}{2} \bar{h}_{m\alpha b\beta n\alpha}^{i\alpha j\beta f\alpha} r_{a\alpha f\alpha}^{m\alpha n\alpha} - \bar{h}_{m\alpha b\beta n\beta}^{i\alpha j\beta f\beta} r_{a\alpha f\beta}^{m\alpha n\beta} \\
& - \bar{h}_{a\alpha n\beta m\alpha}^{i\alpha j\beta e\alpha} r_{e\alpha b\beta}^{m\alpha n\beta} - \frac{1}{2} \bar{h}_{a\alpha n\beta m\beta}^{i\alpha j\beta e\beta} r_{e\beta b\beta}^{m\beta n\beta} \\
& + \frac{1}{2} \bar{h}_{a\alpha b\beta n\alpha}^{e\alpha j\beta f\alpha} r_{e\alpha f\alpha}^{i\alpha n\alpha} + \bar{h}_{a\alpha b\beta n\beta}^{e\alpha j\beta f\beta} r_{e\alpha f\beta}^{i\alpha n\beta} \\
& + \bar{h}_{a\alpha b\beta m\alpha}^{i\alpha f\beta e\alpha} r_{f\beta e\alpha}^{j\beta m\alpha} \\
& + \frac{1}{2} \bar{h}_{a\alpha b\beta m\beta}^{i\alpha f\beta e\beta} r_{f\beta e\beta}^{j\beta m\beta} ,
\end{aligned} \tag{5.7}$$

where the antisymmetrizer $\mathcal{A}_{pq} = 1 - (pq)$ is used, with (pq) designating the transposition of indices p and q , and

$$\begin{aligned}
\bar{h}_{m_\alpha b_\beta n_\alpha}^{i_\alpha j_\beta f_\alpha} &= v_{m_\alpha n_\alpha}^{e_\alpha f_\alpha} t_{e_\alpha b_\beta}^{i_\alpha j_\beta}, \\
\bar{h}_{m_\alpha b_\beta n_\beta}^{i_\alpha j_\beta f_\beta} &= v_{m_\alpha n_\beta}^{e_\alpha f_\beta} t_{e_\alpha b_\beta}^{i_\alpha j_\beta}, \\
\bar{h}_{a_\alpha n_\beta m_\alpha}^{i_\alpha j_\beta e_\alpha} &= v_{m_\alpha n_\beta}^{e_\alpha f_\beta} t_{a_\alpha f_\beta}^{i_\alpha j_\beta}, \\
\bar{h}_{a_\alpha n_\beta m_\beta}^{i_\alpha j_\beta e_\beta} &= v_{m_\beta n_\beta}^{e_\beta f_\beta} t_{a_\alpha f_\beta}^{i_\alpha j_\beta}, \\
\bar{h}_{a_\alpha b_\beta n_\alpha}^{e_\alpha j_\beta f_\alpha} &= v_{m_\alpha n_\alpha}^{e_\alpha f_\alpha} t_{a_\alpha b_\beta}^{m_\alpha j_\beta}, \\
\bar{h}_{a_\alpha b_\beta n_\beta}^{e_\alpha j_\beta f_\beta} &= v_{m_\alpha n_\beta}^{e_\alpha f_\beta} t_{a_\alpha b_\beta}^{m_\alpha j_\beta}, \\
\bar{h}_{a_\alpha b_\beta m_\alpha}^{i_\alpha f_\beta e_\alpha} &= v_{m_\alpha n_\beta}^{e_\alpha f_\beta} t_{a_\alpha b_\beta}^{i_\alpha n_\beta},
\end{aligned}$$

and

$$\bar{h}_{a_\alpha b_\beta m_\beta}^{i_\alpha f_\beta e_\beta} = v_{m_\beta n_\beta}^{e_\beta f_\beta} t_{a_\alpha b_\beta}^{i_\alpha n_\beta}, \tag{5.8}$$

where the μ subscript was dropped from the r amplitudes for clarity. By substituting the three-body components of the similarity transformed Hamiltonian of CCSD given in Eq. (5.8) and factorizing the resulting equations, the open-shell EOMCCSD equations projected on doubly excited determinants, Eqs. (5.5), (5.6), and (5.7), may be rewritten in the following

way:

$$\begin{aligned}
\langle \Phi_{i\alpha j\alpha}^{a\alpha b\alpha} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle &= \mathcal{A}_{i\alpha j\alpha} \mathcal{A}_{a\alpha b\alpha} \left(-\frac{1}{2} \bar{h}_{m\alpha}^{i\alpha} r_{a\alpha b\alpha}^{m\alpha j\alpha} + \frac{1}{2} \bar{h}_{a\alpha}^{e\alpha} r_{e\alpha b\alpha}^{i\alpha j\alpha} \right. \\
&\quad - \frac{1}{2} \bar{h}_{m\alpha b\alpha}^{i\alpha j\alpha} r_{a\alpha}^{m\alpha} + \frac{1}{2} \bar{h}_{a\alpha b\alpha}^{e\alpha j\alpha} r_{e\alpha}^{i\alpha} \\
&\quad + \frac{1}{8} \bar{h}_{a\alpha b\alpha}^{e\alpha f\alpha} r_{e\alpha f\alpha}^{i\alpha j\alpha} + \frac{1}{8} \bar{h}_{m\alpha n\alpha}^{i\alpha j\alpha} r_{a\alpha b\alpha}^{m\alpha n\alpha} \\
&\quad - \bar{h}_{m\alpha a\alpha}^{i\alpha e\alpha} r_{e\alpha b\alpha}^{m\alpha j\alpha} + \bar{h}_{a\alpha m\beta}^{i\alpha e\beta} r_{e\beta b\alpha}^{m\beta j\alpha} \\
&\quad \left. - \frac{1}{2} \chi_{i\alpha}^{m\alpha} t_{a\alpha b\alpha}^{m\alpha j\alpha} + \frac{1}{2} \chi_{a\alpha}^{e\alpha} t_{e\alpha b\alpha}^{i\alpha j\alpha} \right) \quad (5.9)
\end{aligned}$$

$$\begin{aligned}
\langle \Phi_{i\beta j\beta}^{a\beta b\beta} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle &= \mathcal{A}_{a\beta b\beta} \mathcal{A}_{i\beta j\beta} \left(-\frac{1}{2} \bar{h}_{m\beta}^{i\beta} r_{a\beta b\beta}^{m\beta j\beta} + \frac{1}{2} \bar{h}_{a\beta}^{e\beta} r_{e\beta b\beta}^{i\beta j\beta} \right. \\
&\quad - \frac{1}{2} \bar{h}_{m\beta b\beta}^{i\beta j\beta} r_{a\beta}^{m\beta} + \frac{1}{2} \bar{h}_{a\beta b\beta}^{e\beta j\beta} r_{e\beta}^{i\beta} \\
&\quad + \frac{1}{8} \bar{h}_{a\beta b\beta}^{e\beta f\beta} r_{e\beta f\beta}^{i\beta j\beta} + \frac{1}{8} \bar{h}_{m\beta n\beta}^{i\beta j\beta} r_{a\beta b\beta}^{m\beta n\beta} \\
&\quad - \bar{h}_{m\beta a\beta}^{i\beta e\beta} r_{e\beta b\beta}^{m\beta j\beta} + \bar{h}_{a\beta m\alpha}^{i\beta e\alpha} r_{e\alpha b\beta}^{m\alpha j\beta} \\
&\quad \left. - \frac{1}{2} \chi_{m\beta}^{i\beta} t_{a\beta b\beta}^{m\beta j\beta} + \frac{1}{2} \chi_{a\beta}^{e\beta} t_{e\beta b\beta}^{i\beta j\beta} \right) \quad (5.10)
\end{aligned}$$

and

$$\begin{aligned}
\langle \Phi_{i\alpha j\beta}^{a\alpha b\beta} | [\bar{H}_{N,\text{open}}^{(\text{CCSD})}(R_{\mu,1} + R_{\mu,2})]_C | \Phi \rangle = & -\bar{h}_{m\alpha}^{i\alpha} r_{a\alpha b\beta}^{m\alpha j\beta} - \bar{h}_{n\beta}^{j\beta} r_{a\alpha b\beta}^{i\alpha n\beta} + \bar{h}_{a\alpha}^{e\alpha} r_{e\alpha b\beta}^{i\alpha j\beta} \\
& + \bar{h}_{b\beta}^{f\beta} r_{a\alpha f\beta}^{i\alpha j\beta} - \bar{h}_{m\alpha b\beta}^{i\alpha j\beta} r_{a\alpha}^{m\alpha} - \bar{h}_{a\alpha n\beta}^{i\alpha j\beta} r_{b\beta}^{n\beta} \\
& + \bar{h}_{a\alpha b\beta}^{e\alpha j\beta} r_{e\alpha}^{i\alpha} + \bar{h}_{a\alpha b\beta}^{i\alpha f\beta} r_{f\beta}^{j\beta} \\
& + \bar{h}_{a\alpha b\beta}^{e\alpha f\beta} r_{e\alpha f\beta}^{i\alpha j\beta} + \bar{h}_{m\alpha n\beta}^{i\alpha j\beta} r_{a\alpha b\beta}^{m\alpha n\beta} \\
& - \bar{h}_{a\alpha n\beta}^{e\alpha j\beta} r_{e\alpha b\beta}^{i\alpha n\beta} - \bar{h}_{m\alpha b\beta}^{i\alpha f\beta} r_{a\alpha f\beta}^{m\alpha j\beta} \\
& + \bar{h}_{a\alpha n\beta}^{i\alpha f\beta} r_{f\beta b\beta}^{n\beta j\beta} + \bar{h}_{m\alpha b\beta}^{e\alpha j\beta} r_{e\alpha a\alpha}^{m\alpha i\alpha} \\
& - \bar{h}_{m\alpha a\alpha}^{i\alpha e\alpha} r_{e\alpha b\beta}^{m\alpha j\beta} - \bar{h}_{b\beta n\beta}^{f\beta j\beta} r_{a\alpha f\beta}^{i\alpha n\beta} \\
& - \chi_{m\alpha}^{i\alpha} t_{a\alpha b\beta}^{m\alpha j\beta} + \chi_{a\alpha}^{e\alpha} t_{e\alpha b\beta}^{i\alpha j\beta} \\
& - \chi_{n\beta}^{j\beta} t_{a\alpha b\beta}^{i\alpha n\beta} + \chi_{b\beta}^{f\beta} t_{a\alpha f\beta}^{i\alpha j\beta}
\end{aligned} \tag{5.11}$$

where

$$\begin{aligned}
\chi_{m\alpha}^{i\alpha} = & -v_{n\alpha m\alpha}^{i\alpha f\alpha} r_{f\alpha}^{n\alpha} + v_{m\alpha n\beta}^{i\alpha f\beta} r_{f\beta}^{n\beta} - v_{n\alpha m\alpha}^{e\alpha f\alpha} r_{f\alpha}^{n\alpha} t_{e\alpha}^{i\alpha} \\
& + v_{m\alpha n\beta}^{e\alpha f\beta} r_{f\beta}^{n\beta} t_{e\alpha}^{i\alpha} + \frac{1}{2} v_{m\alpha n\alpha}^{e\alpha f\alpha} r_{e\alpha f\alpha}^{i\alpha n\alpha} + v_{m\alpha n\beta}^{e\alpha f\beta} r_{e\alpha f\beta}^{i\alpha n\beta},
\end{aligned} \tag{5.12}$$

$$\begin{aligned}
\chi_{a\alpha}^{e\alpha} = & -v_{a\alpha n\alpha}^{f\alpha e\alpha} r_{f\alpha}^{n\alpha} + v_{a\alpha n\beta}^{e\alpha f\beta} r_{f\beta}^{n\beta} + v_{n\alpha m\alpha}^{e\alpha f\alpha} r_{f\alpha}^{n\alpha} t_{a\alpha}^{m\alpha} \\
& - v_{m\alpha n\beta}^{e\alpha f\beta} r_{f\beta}^{n\beta} t_{a\alpha}^{m\alpha} - \frac{1}{2} v_{m\alpha n\alpha}^{e\alpha f\alpha} r_{a\alpha f\alpha}^{m\alpha n\alpha} - v_{m\alpha n\beta}^{e\alpha f\beta} r_{a\alpha f\beta}^{m\alpha n\beta},
\end{aligned} \tag{5.13}$$

$$\begin{aligned}
\chi_{m\beta}^{i\beta} = & -v_{n\beta m\beta}^{i\beta f\beta} r_{f\beta}^{n\beta} + v_{m\beta n\alpha}^{i\beta f\alpha} r_{f\alpha}^{n\alpha} - v_{n\beta m\beta}^{e\beta f\beta} r_{f\beta}^{n\beta} t_{e\beta}^{i\beta} \\
& + v_{m\beta n\alpha}^{e\beta f\alpha} r_{f\alpha}^{n\alpha} t_{e\beta}^{i\beta} + \frac{1}{2} v_{m\beta n\beta}^{e\beta f\beta} r_{e\beta f\beta}^{i\beta n\beta} + v_{m\beta n\alpha}^{e\beta f\alpha} r_{e\beta f\alpha}^{i\beta n\alpha},
\end{aligned} \tag{5.14}$$

and

$$\begin{aligned}
\chi_{a\beta}^{e\beta} = & -v_{a\beta n\beta}^{f\beta e\beta} r_{f\beta}^{n\beta} + v_{a\beta n\alpha}^{e\beta f\alpha} r_{f\alpha}^{n\alpha} + v_{n\beta m\beta}^{e\beta f\beta} r_{f\beta}^{n\beta} t_{a\beta}^{m\beta} \\
& - v_{m\beta n\alpha}^{e\beta f\alpha} r_{f\alpha}^{n\alpha} t_{a\beta}^{m\beta} - \frac{1}{2} v_{m\beta n\beta}^{e\beta f\beta} r_{a\beta f\beta}^{m\beta n\beta} - v_{m\beta n\alpha}^{e\beta f\alpha} r_{a\beta f\alpha}^{m\beta n\alpha},
\end{aligned} \tag{5.15}$$

which is the final form of the open-shell EOMCCSD equations used in the efficient vectorized GAMESS code. Once the singly and doubly excited amplitudes defining the EOMCCSD excitation operator, $r_{\mu,a}^i$ and $r_{\mu,ab}^{ij}$, respectively, and the vertical excitation energy $\omega_{\mu}^{(\text{CCSD})}$ have been determined by solving Eqs. (3.24) (3.25), $r_{\mu,0}$ is calculated *a posteriori* from the following expression:

$$r_{\mu,0} = \langle \Phi | [\bar{H}_{N,\text{open}}^{(\text{CCSD})} (R_{\mu,1}^{(\text{CCSD})} + R_{\mu,2}^{(\text{CCSD})})]_C | \Phi \rangle / \omega_{\mu}^{(\text{CCSD})}. \tag{5.16}$$

5.1.2 Electron-Attached and Ionized Equation-of-Motion Coupled-Cluster Theories

The key difference between the open-shell EOMCC theory and EA- or IP-EOMCC is the sector of the Fock space the similarity-transformed Hamiltonian of CCSD is diagonalized within. As an example, in the IP-EOMCC approaches we diagonalize the similarity-transformed Hamiltonian obtained in calculations for an N -electron reference system in the sector of the Fock space corresponding to $(N - 1)$ electrons. Thus, the left-hand sides of the IP-

EOMCCSD($2h-1p$) eigenvalue problem,

$$(\bar{H}_{N,\text{open}}^{(\text{CCSD})} R_{\mu}^{(2h-1p)})_C |\Phi\rangle = \omega_{\mu}^{(2h-1p)} R_{\mu}^{(2h-1p)} |\Phi\rangle \quad (5.17)$$

are obtained by projecting Eq. (5.17) onto all $\langle\Phi_{i\beta}|$, $\langle\Phi_{i\beta j\beta}^{b\beta}|$, and $\langle\Phi_{i\beta j\alpha}^{b\alpha}|$ determinants. In this way the following equations are obtained:

$$\begin{aligned} \langle\Phi_{i\beta}|(\bar{H}_{N,\text{open}}^{(\text{CCSD})} R_{N,\text{open}}^{(2h-1p)})_C |\Phi\rangle &= -\bar{h}_{m\beta}^{i\beta} r^{m\beta} + \bar{h}_{m\alpha}^{e\alpha} r_{e\alpha}^{i\beta m\alpha} + \bar{h}_{m\beta}^{e\beta} r_{e\beta}^{i\beta m\beta} \\ &\quad -\bar{h}_{m\beta n\alpha}^{i\beta f\alpha} r_{f\alpha}^{m\beta n\alpha} - \frac{1}{2} \bar{h}_{m\beta n\beta}^{i\beta f\beta} r_{f\beta}^{m\beta n\beta}, \end{aligned} \quad (5.18)$$

$$\begin{aligned} \langle\Phi_{i\beta j\beta}^{b\beta}|(\bar{H}_{N,\text{open}}^{(\text{CCSD})} R_{N,\text{open}}^{(2h-1p)})_C |\Phi\rangle &= -\mathcal{A}_{i\beta j\beta} \bar{h}_{n\beta}^{j\beta} r_{b\beta}^{i\beta n\beta} + \bar{h}_{b\beta}^{f\beta} r_{f\beta}^{i\beta j\beta} \\ &\quad -\bar{h}_{m\beta b\beta}^{i\beta j\beta} r^{m\beta} + \mathcal{A}_{i\beta j\beta} \bar{h}_{m\beta b\alpha}^{e\beta j\alpha} r_{e\alpha}^{i\beta m\alpha} \\ &\quad -\mathcal{A}_{i\beta j\beta} \bar{h}_{b\beta m\beta}^{e\beta j\beta} r_{e\beta}^{i\beta m\beta} + \frac{1}{2} \bar{h}_{m\beta n\beta}^{i\beta j\beta} r_{e\beta}^{m\beta n\beta} \\ &\quad -\frac{1}{2} v_{m\beta n\beta}^{e\beta f\beta} t_{e\beta b\beta}^{i\beta j\beta} r_{f\beta}^{m\beta n\beta} \\ &\quad -v_{m\beta n\alpha}^{e\beta f\alpha} t_{e\beta b\beta}^{i\beta j\beta} r_{f\alpha}^{m\beta n\alpha}, \end{aligned} \quad (5.19)$$

and

$$\begin{aligned}
\langle \Phi_{i\beta j\alpha}^{b\alpha} | (\bar{H}_{N,\text{open}}^{(\text{CCSD})} R_{N,\text{open}}^{(2h-1p)})_C | \Phi \rangle = & -\bar{h}_{n\alpha}^{j\alpha} r_{b\alpha}^{i\beta n\alpha} + \bar{h}_{b\alpha}^{f\alpha} r_{f\alpha}^{i\beta j\alpha} \\
& -\bar{h}_{m\beta}^{i\beta} r_{b\alpha}^{m\beta j\alpha} - \bar{h}_{m\beta b\alpha}^{i\beta j\alpha} r^{m\beta} \\
& -\bar{h}_{m\beta j\alpha}^{e\beta b\alpha} r_{e\beta}^{i\beta m\beta} - \bar{h}_{b\alpha m\alpha}^{e\alpha j\alpha} r_{e\alpha}^{i\beta m\alpha} \\
& -\bar{h}_{m\beta b\alpha}^{i\beta e\alpha} r_{e\alpha}^{m\beta j\alpha} + \bar{h}_{m\beta n\alpha}^{i\beta j\alpha} r_{b\alpha}^{m\beta n\alpha} \\
& -\frac{1}{2} v_{m\beta n\beta}^{e\beta f\beta} t_{e\beta b\alpha}^{i\beta j\alpha} r_{f\beta}^{m\beta n\beta} \\
& -v_{m\beta n\alpha}^{e\beta f\alpha} t_{e\beta b\alpha}^{i\beta j\alpha} r_{f\alpha}^{m\beta n\alpha}.
\end{aligned} \tag{5.20}$$

As in the case of EOMCCSD, the IP-EOMCCSD(2h-1p) equations are solved using the Hirao-Nakatsuji algorithm [303].

Chapter 6

Summary and Concluding Remarks

In this dissertation, we addressed the problem of generating highly accurate potential energy surfaces (PESs) for reactive processes by introducing and demonstrating the performance of electronic structure methodologies that can provide a balanced description of chemical species with varying levels of electronic degeneracy, but are also practical enough to be applied to a wide range of chemical problems, as well as extrapolation techniques which facilitate the generation of PESs corresponding to high-level electronic structure calculations in a much more efficient manner than that offered by conventional and laborious point-wise computations. In particular, we examined the performance of two classes of coupled-cluster (CC) methods which are capable of accounting for the diverse electron correlation effects encountered in the majority of ground- and excited-state PES considerations. The first class of methods consisted of the size-extensive completely renormalized (CR) CC approaches for ground states and their equation-of-motion (EOM) CC extensions for excited states, in which noniterative corrections due to higher-order correlation effects are added to the energies obtained with the standard CC and EOMCC approximations, such as CCSD or EOMCCSD, respectively.

We showed that the left-eigenstate CR-CC(2,3) and CR-EOMCC(2,3) methods that belong to this category offer excellent performance for a diverse range of applications, including a benchmark database of barrier heights for thermochemical kinetics, a pair of bimolecular association mechanisms involving the ozone molecule, competing intramolecular reaction mechanisms describing the isomerization of bicyclobutane to butadiene, and the ground- and excited-state PES cuts for the water molecule. When necessary, corrections for quadruple excitations were also included via the CR-CC(2,3)+Q method which usually improved the performance of the CR-CC(2,3) methods from chemical to sub-chemical accuracies for many of the studied systems. A new variant of the CR-EOMCC(2,3) method was also presented and discussed, namely, the δ -CR-EOMCC(2,3) approach that can provide a size-intensive treatment of excitation energies. This method was applied to describe excitation energies and hydrogen-bonding-induced spectral shifts in complexes of 7-Hydroxyquinoline with considerable success, helping to explain problems with time-dependent density functional theory. The second class of methods considered here were the active-space variants of the electron attached (EA) and ionized (IP) EOMCC theories. The EA- and IP-EOMCC approaches were shown to be an excellent alternative to open-shell CC and EOMCC methods and their perturbative extensions for describing open-shell molecular systems, providing spin-adapted results while their active-space variants proved to be extremely efficient, significantly reducing the costs of the high-level parent EA- and IP-EOMCC approximations without sacrificing accuracy. We also developed a general strategy for reducing the cost of generating PESs with correlated electronic structure methods via the concept of correlation energy scaling. In order to demonstrate typical accuracies one may expect when using the two types of PES extrapolation schemes presented here, namely, the single-level and dual-level schemes, a number

of benchmark applications were presented, such as the previously mentioned bicyclobutane isomerization and single-bond breaking potential energy curves for the H_2O , HCl , and F_2 molecules. The single-level extrapolation schemes were shown to reproduce PESs obtained in laborious high-level point-by-point computations to within fractions of a millihartree in most cases, even when used to extrapolate the PES to the CBS-limit. Meanwhile, the dual-level PES extrapolation schemes were shown to be capable of producing similar accuracies at a tiny fraction of the computational cost of their single-level analogs. The insensitivity of the results to the choice of pivot geometry and improvements in accuracy available when a higher-order base wave function is chosen were also demonstrated. Finally, the development of new open-shell EOMCCSD and IP-EOMCCSD($2h-1p$) computer codes for the GAMESS software package, along with the corresponding programmable equations, was discussed.

REFERENCES

REFERENCES

- [1] P. A. M. Dirac, P. R. Soc. London-ContA. **123**, 714 (1929).
- [2] M. Born, R. Oppenheimer, Annalen der Physik **389**, 457 (1927).
- [3] H. M. James, A. S. Coolidge, J. Chem. Phys. **1**, 825 (1933).
- [4] D. R. Hartree, Proc. Cam. Phil. Soc. **24**, 89 (1928).
- [5] D. R. Hartree, Proc. Cam. Phil. Soc. **24**, 111 (1928).
- [6] D. R. Hartree, Proc. Cam. Phil. Soc. **24**, 426 (1928).
- [7] V. Fock, Z. Physik. **61**, 126 (1930).
- [8] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
- [9] A. C. Wahl, J. Chem. Phys. **41**, 2600 (1964).
- [10] S. F. Boys, Proc. Roy. Soc. **A201**, 125 (1950).
- [11] P. O. Löwdin, Phys. Rev. **97**, 1474 (1955).
- [12] P. O. Löwdin, Phys. Rev. **97**, 1490 (1955).
- [13] P. O. Löwdin, Phys. Rev. **97**, 1509 (1955).
- [14] J. A. Pople, J. S. Binkley, R. Seeger, Int. J. Quantum Chem. Symp. **10**, 1, (1976).

- [15] C. Møller, M. S. Plesset, Phys. Rev. **46**, 618 (1934).
- [16] K. A. Brueckner, Phys. Rev. **97**, 1353 (1955).
- [17] K. A. Brueckner, Phys. Rev. **100**, 36 (1955).
- [18] J. Goldstone, Proc. Roy. Soc. **A239**, 267 (1957).
- [19] J. Hubbard, Proc. Roy. Soc. **A240**, 539 (1957).
- [20] J. Hubbard, Proc. Roy. Soc. **A243**, 336 (1958).
- [21] J. Hubbard, Proc. Roy. Soc. **A244**, 199 (1958).
- [22] N. M. Hugenholtz, Physica **23**, 481 (1957).
- [23] F. Coester, Nucl. Phys. **7**, 421 (1958).
- [24] F. Coester, H. Kümmel, Nucl. Phys. **17**, 477 (1960).
- [25] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
- [26] J. Čížek, Adv. Chem. Phys. **14**, 35 (1969).
- [27] J. Čížek, J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).
- [28] S. R. Langhoff, E. R. Davidson, Int. J. Quantum Chem. **8**, 61 (1974).
- [29] E. R. Davidson, In *The World of Quantum Chemistry*, edited by R. Daudel, B. Pullmsn, pages 17–30 (Reidel, Dordrecht, 1974).
- [30] K. Emrich, Nucl. Phys. A **351**, 379 (1981).
- [31] J. Geertsen, M. Rittby, and R. J. Bartlett, Chem. Phys. Lett. **164**, 57 (1989).
- [32] D. C. Comeau and R. J. Bartlett, Chem. Phys. Lett. **207**, 414 (1993).
- [33] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).

- [34] P. Piecuch and R. J. Bartlett, *Adv. Quantum Chem.* **34**, 295 (1999).
- [35] H. Nakatsuji and K. Hirao, *Chem. Phys. Lett.* **47**, 569 (1977).
- [36] H. Nakatsuji and K. Hirao, *J. Chem. Phys.* **68**, 4279 (1978).
- [37] H. Nakatsuji, *Chem. Phys. Lett.* **59**, 362 (1978).
- [38] H. Nakatsuji, In *Computational Chemistry: Reviews of Current Trends*, edited by J. Leszczyński, vol. 2, pages 62–124 (World Scientific, Singapore, 1997), and references therein.
- [39] H. Nakatsuji, *Bull. Chem. Soc. Jpn.* **78**, 1705 (2005), and references therein.
- [40] H. Monkhorst, *Int. J. Quantum Chem. Symp.* **11**, 421 (1977).
- [41] E. Dalgaard and H. Monkhorst, *Phys. Rev. A* **28**, 1217 (1983).
- [42] M. Takahashi and J. Paldus, *J. Chem. Phys.* **85**, 1486 (1986).
- [43] H. Koch and P. Jørgensen, *J. Chem. Phys.* **93**, 3333 (1990).
- [44] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, *J. Chem. Phys.* **93**, 3345 (1990).
- [45] G. D. Purvis III, R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- [46] J. M. Cullen and M. C. Zerner, *J. Chem. Phys.* **77**, 4088 (1982).
- [47] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer III, *J. Chem. Phys.* **86**, 2991 (1987).
- [48] P. Piecuch and J. Paldus, *Int. J. Quantum Chem.* **36**, 429 (1989).
- [49] J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987); **89**, 3401 (1988) [Erratum].
- [50] G. E. Scuseria and H. F. Schaefer, III, *Chem. Phys. Lett.* **152**, 382 (1988).
- [51] N. Oliphant and L. Adamowicz, *J. Chem. Phys.* **95**, 6645 (1991).

- [52] S. A. Kucharski and R. J. Bartlett, *Theor. Chim. Acta* **80**, 387 (1991).
- [53] S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992).
- [54] P. Piecuch and L. Adamowicz, *J. Chem. Phys.* **100**, 5792 (1994).
- [55] K. Kowalski and P. Piecuch, *J. Chem. Phys.* **113**, 8490 (2000).
- [56] K. Kowalski and P. Piecuch, *J. Chem. Phys.* **115**, 643 (2001).
- [57] K. Kowalski and P. Piecuch, *Chem. Phys. Lett.* **347**, 237 (2001).
- [58] S. A. Kucharski, M. Włoch, M. Musiał, and R. J. Bartlett, *J. Chem. Phys.* **115**, 8263 (2001).
- [59] S. Hirata, *J. Chem. Phys.* **121**, 51 (2004).
- [60] J. Paldus, In *Methods in computational molecular physics*, edited by S. Wilson and G. H. F. Dierksen, NATO ASI series B: physics, vol 293, page 99, (Plenum, New York, 1992).
- [61] R. J. Bartlett, In *Advanced series in physical chemistry. Modern electronic structure theory, part I*, edited by D. R. Yarkony, page 1047, (World Scientific, Singapore, 1995).
- [62] J. Paldus and X. Li, *Adv. Chem. Phys.* **110**, 1 (1999).
- [63] T. D. Crawford and H. F. Schaefer III, *Rev. Comp. Chem.* **14**, 33 (2000).
- [64] J. Gauss In *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III, and P. R. Schreiner, vol. 1, page 615 (Wiley, Chichester, 1998).
- [65] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, *Int. Rev. Phys. Chem.* **21** 527 (2002).
- [66] M. Musiał and R. J. Bartlett, *Rev. Mod. Phys.* **79**, 291 (2007).
- [67] A. P. Rendell, T. J. Lee, and A. Komornicki, *Chem. Phys. Lett.* **178**, 462 (1991).

- [68] A. P. Rendell, T. J. Lee, and R. Lindh, Chem. Phys. Lett. **194**, 84 (1992).
- [69] A. P. Rendell, M. F. Guest, and R. A. Kendall, J. Comput. Chem. **14**, 1429 (1993).
- [70] R. Kobayashi and A. P. Rendell, Chem. Phys. Lett. **265**, 1 (1997).
- [71] P. Piecuch and J. I. Landman, Parallel Comp. **26**, 913 (2000).
- [72] S. Hirata, J. Phys. Chem. A **107**, 9887 (2003).
- [73] G. Baumgartner, A. Auer, D. E. Bernholdt, A. Bibireata, V. Choppella, D. Cociorva, X. Gao, R. J. Harrison, S. Hirata, S. Krishnamoorthy, S. Krishnan, C. Lam, Q. Lu, M. Nooijen, R. M. Pitzer, J. Ramanujam, P. Sadayappan, and A. Sibiryakov, Proc. IEEE **93**, 276 (2005).
- [74] P. Piecuch, S. Hirata, K. Kowalski, P.-D. Fan, and T. L. Windus, Int. J. Quantum Chem. **106**, 79 (2006).
- [75] T. Janowski, A. R. Ford, and P. Pulay, J. Chem. Theory Comput. **3**, 1368 (2007).
- [76] R. M. Olson, L. J. Bentz, R. A. Kendall, M. W. Schmidt, and M. S. Gordon, J. Chem. Theory Comput. **3**, 1312 (2007).
- [77] M. E. Harding, T. Metzroth, J. Gauss, and A. A. Auer, J. Chem. Theory Comput. **4**, 64 (2008).
- [78] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
- [79] M. Schütz, J. Chem. Phys. **114**, 113 9986 (2000).
- [80] M. Schütz and H.-J. Werner, Chem. Phys. Lett. **318**, 370 (2000).
- [81] M. Schütz, J. Chem. Phys. **116**, 8772 (2002).
- [82] J. E. Subotnik and M. Head-Gordon, J. Chem. Phys. **123**, 064108 (2005).
- [83] J. E. Subotnik, A. Sodt, and M. Head-Gordon, J. Chem. Phys. **125** 074116 (2006).
- [84] W. Li, J.R. Gour, P. Piecuch, and S. Li, J. Chem. Phys. **131**, 114109 (2009).

- [85] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 6721 (2010).
- [86] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).
- [87] H. J. Nakano, Chem. Phys. **99**, 7983 (1993).
- [88] H. J. Nakano, Chem. Phys. Lett. **207**, 372 (1993).
- [89] K. Andersson, P. A. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).
- [90] H.-J. Werner and P. J. Knowles, J. Chem. Phys. **89**, 5803 (1988).
- [91] P. J. Knowles and H.-J. Werner, Chem. Phys. Lett. **145**, 514 (1988).
- [92] I. Lindgren and D. Mukherjee, Phys. Rep. **151**, 93 (1987).
- [93] D. Mukherjee and S. Pal, Adv. Quantum Chem. **20**, 291 (1989).
- [94] B. Jeziorski and H.J. Monkhorst, Phys. Rev. A **24**, 1668 (1981).
- [95] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A **47**, 2738 (1993).
- [96] P. Piecuch, R. Toboła, and J. Paldus, Chem. Phys. Lett. **210**, 243 (1993).
- [97] P. Piecuch and J. Paldus, Phys. Rev. A **49**, 3479 (1994).
- [98] P. Piecuch and K. Kowalski, Int. J. Mol. Sci. **3**, 676 (2002).
- [99] X. Li and J. Paldus, Int. J. Quantum Chem. **110**, 2734 (2010).
- [100] X. Li and J. Paldus, J. Phys. Chem. A **114**, 8591 (2010).
- [101] P. Piecuch, Mol. Phys. **108**, 2987 (2010).
- [102] Y.S. Lee and R.J. Bartlett, J. Chem. Phys, **80**, 4371 (1984).
- [103] Y.S. Lee, S.A. Kucharski, and R.J. Bartlett, J. Chem. Phys., **81**, 5906 (1984); **82**, 5761 (1985) [Erratum].

- [104] P. Piecuch and J. Paldus, *Theor. Chim. Acta.*, **78**, 65 (1990).
- [105] S.A. Kucharski and R.J. Bartlett, *Chem. Phys. Lett.*, **158**, 550 (1989).
- [106] M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, *J. Chem. Phys.* **83**, 4041 (1985).
- [107] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- [108] S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.*, **108**, 9221 (1998).
- [109] J. D. Watts and R. J. Bartlett, *Chem. Phys. Lett.* **233**, 81 (1995).
- [110] J. D. Watts and R. J. Bartlett, *Chem. Phys. Lett.* **258**, 581 (1996).
- [111] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **244**, 75 (1995).
- [112] O. Christiansen, H. Koch, and P. Jørgensen, *J. Chem. Phys.* **103**, 7429 (1995).
- [113] O. Christiansen, H. Koch, and P. Jørgensen, *J. Chem. Phys.* **105**, 1451 (1996).
- [114] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **256**, 185 (1996).
- [115] K. Kowalski and P. Piecuch, *J. Chem. Phys.* **120**, 1715 (2004).
- [116] X. Li and J. Paldus, *J. Chem. Phys.* **107**, 6257 (1997).
- [117] X. Li and J. Paldus, *J. Chem. Phys.* **108**, 637 (1998).
- [118] X. Li and J. Paldus, *J. Chem. Phys.* **113**, 9966 (2000).
- [119] X. Li and J. Paldus, *Mol. Phys.* **98**, 1185 (2000).
- [120] X. Li and J. Paldus, *Int. J. Quant. Chem.* **80**, 743 (2000).
- [121] X. Li and J. Paldus, *J. Chem. Phys.* **128**, 144118 (2008).

- [122] X. Li and J. Paldus, J. Chem. Phys. **128**, 144119 (2009).
- [123] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P. D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kus, and M. Musiał, Theor. Chem. Acc. **112**, 349 (2004).
- [124] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).
- [125] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 5644 (2000).
- [126] M. Włoch, J. R. Gour, K. Kowalski, and P. Piecuch J. Chem. Phys. **122**, 214107 (2005).
- [127] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).
- [128] P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, Chem. Phys. Lett. **418**, 463 (2005).
- [129] M. Włoch, M. D. Lodriguito, P. Piecuch, and J. R. Gour, Mol. Phys. **104**, 2149 (2006).
- [130] M. Włoch, J. R. Gour, and P. Piecuch, J. Phys. Chem. A **111**, 11359 (2007).
- [131] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. **108**, 2128 (2008).
- [132] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. **109**, 3268 (2009).
- [133] N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).
- [134] N. Oliphant and L. Adamowicz, J. Chem. Phys. **96**, 3739 (1992).
- [135] N. Oliphant and L. Adamowicz, Int. Rev. Phys. Chem. **12**, 339 (1993).
- [136] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. **99**, 1875 (1993).
- [137] P. Piecuch and L. Adamowicz, Chem. Phys. Lett. **221**, 121 (1993).
- [138] P. Piecuch and L. Adamowicz, J. Chem. Phys. **102**, 898 (1995).
- [139] K. B. Ghose, P. Piecuch, and L. Adamowicz, J. Chem. Phys. **103**, 9331 (1995).
- [140] K. B. Ghose and L. Adamowicz, J. Chem. Phys. **103**, 9324 (1995).

- [141] V. Alexandrov, P. Piecuch, and L. Adamowicz, J. Chem. Phys. **102**, 3301 (1995).
- [142] K. B. Ghose, P. Piecuch, S. Pal, and L. Adamowicz, J. Chem. Phys. **104**, 6582 (1996).
- [143] L. Adamowicz, P. Piecuch, and K. B. Ghose, Mol. Phys. **94**, 225 (1998).
- [144] P. Piecuch, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **110**, 6103 (1999).
- [145] P. Piecuch, S. A. Kucharski, and V. Špirko, J. Chem. Phys. **111**, 6679 (1999).
- [146] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **344**, 165 (2001).
- [147] J. R. Gour, P. Piecuch, and M. Włoch, J. Chem. Phys. **123**, 134113 (2005).
- [148] J. R. Gour, P. Piecuch, and M. Włoch, Int. J. Quantum Chem. **106**, 2854 (2006).
- [149] J. R. Gour and P. Piecuch, J. Chem. Phys. **125**, 234107 (2006).
- [150] P. Piecuch, R. Toboła, and J. Paldus, Phys. Rev. A **54**, 1210 (1996).
- [151] J. Paldus, J. Čížek, and M. Takahashi, Phys. Rev. A **30**, 2193 (1984).
- [152] J. Paldus and J. Planelles, Theor. Chim. Acta **89**, 13 (1994).
- [153] J. Paldus , and X. Li, Theor. Chim. Acta **89**, 33 (1994).
- [154] J. Paldus , and X. Li, Theor. Chim. Acta **89**, 59 (1994).
- [155] L. Stolarczyk, Chem. Phys. Lett. **217**, 1 (1994).
- [156] G. Peris, J. Planelles, F. Rajadall, and J. Paldus, J. Chem. Phys. **62**, 137 (1997).
- [157] X. Li, G. Peris, J. Planelles, F. Rajadall, and J. Paldus, J. Chem. Phys. **107**, 90 (1997).
- [158] X. Li and J. Paldus, J. Chem. Phys. **119**, 5334 (2003)
- [159] J. Paldus and X. Li, J. Chem. Phys. **118**, 6769 (2003).

- [160] X. Li and J. Paldus, J. Chem. Phys. **124**, 034112 (2006).
- [161] X. Li and J. Paldus, J. Chem. Phys. **125**, 064107 (2006).
- [162] X. Li and J. Paldus, Chem. Phys. Lett. **431**, 179 (2006).
- [163] J. Paldus and X. Li, Collect. Czech. Chem. Commun. **72**, 100 (2007).
- [164] X. Li and J. Paldus, J. Chem. Phys. **126**, 224304 (2007).
- [165] X. Li and J. Paldus, J. Phys. Chem. A **111**, 11189 (2007).
- [166] X. Li, J.R. Gour, J. Paldus, and P. Piecuch, Chem. Phys. Lett. **461**, 321 (2008).
- [167] X. Li and J. Paldus, J. Theor. Comp. Chem. **7**, 805 (2008).
- [168] X. Li and J. Paldus, J. Chem. Phys. **131**, 114103 (2009).
- [169] X. Li and J. Paldus, Int. J. Quantum Chem. **109**, 3305 (2009).
- [170] X. Li and J. Paldus, J. Chem. Phys. **132**, 114103 (2010).
- [171] X. Li and J. Paldus, J. Chem. Phys. **134**, 074301 (2011).
- [172] V. Spirko, X. Li, and J. Paldus, Collect. Czech. Chem. Commun. **76**, 327 (2011).
- [173] X. Li and J. Paldus, J. Chem. Phys. **119**, 5320 (2003).
- [174] X. Li and J. Paldus, J. Chem. Phys. **119**, 5346 (2003).
- [175] X. Li and J. Paldus, J. Chem. Phys. **120**, 5890 (2004).
- [176] X. Li and J. Paldus, Mol. Phys. **104**, 661 (2006).
- [177] X. Li, Collect. Czech. Chem. Commun. **70**, 755 (2005).
- [178] X. Li and J. Paldus, J. Chem. Phys. **124**, 034112 (2006).

- [179] X. Li and J. Paldus, J. Phys. Chem. A **114**, 8591 (2010).
- [180] X. Li and J. Paldus, Int. J. Quantum Chem. **110**, 2734 (2010).
- [181] X. Li and J. Paldus, J. Chem. Phys. **133**, 024102 (2010).
- [182] K. Kowalski and P. Piecuch, J. Chem. Phys., **115**, 2966 (2001).
- [183] K. Kowalski and P. Piecuch, J. Chem. Phys., **116**, 7411 (2002).
- [184] A. J. C. Varandas and P. Piecuch, Chem. Phys. Lett. **430**, 448 (2006).
- [185] Y. Ge, M. S. Gordon, and P. Piecuch, J. Chem. Phys. 2007, **127**, 174106 (2007).
- [186] P. Piecuch, M. Włoch, and A. J. C. Varandas, Theor. Chem. Acc. **120**, 59 (2008).
- [187] Y. Z. Song, A. Kinal, P. J. S. B. Caridade, A. J. C. Varandas, and P. Piecuch, J. Mol. Struct.: THEOCHEM. **859**, 22 (2008).
- [188] Y. Ge, M. S. Gordon, P. Piecuch, M. Włoch, and J. R. Gour, J. Phys. Chem. A **112**, 11873 (2008).
- [189] A. Kinal and P. Piecuch, J. Phys. Chem. A **111**, 734 (2007).
- [190] C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, J. Phys. Chem. A **110**, 1991 (2006); **111**, 4871 (2007) [Erratum].
- [191] C. J. Cramer, A. Kinal, M. Włoch, P. Piecuch, and L. Gagliardi, J. Phys. Chem. A **110**, 11557 (2006); **111**, 4871 (2007) [Erratum].
- [192] C. J. Cramer, J. R. Gour, A. Kinal, M. Włoch, P. Piecuch, A. R. M. Shahi, and Gagliardi, L. J. Phys. Chem. A **112**, 3754 (2008).
- [193] J. J. Lutz and P. Piecuch, J. Chem. Phys. **128**, 154116 (2008).
- [194] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).

- [195] M. S. Gordon and M. W. Schmidt, In: Theory and Applications of Computational Chemistry: The First Forty Years, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria, pages 1167–1190, (Elsevier, Amsterdam, 2005).
- [196] L. Meissner and R. J. Bartlett J. Chem. Phys. **102**, 7490 (1995).
- [197] E. A. Salter, G. W. Trucks, and R. J. Bartlett, J. Chem. Phys. **90**, 1752 (1989).
- [198] P. S. Epstein, Phys. Rev. **28**, 695 (1934).
- [199] R. D. Nesbet, Proc. Roy. Soc. (London) **A230**, 312 (1955).
- [200] T. Shiozaki, K. Hirao, S. Hirata, J. Chem. Phys. **126**, 244106 (2007).
- [201] G. Fradelos, J. J. Lutz, T. A. Wesolowski, P. Piecuch, M. Wloch, J. Chem. Theory Comput. **7**, 1647 (2011).
- [202] B. J. Lynch and D. G. Truhlar, J. Phys. Chem. A **107**, 3898 (2003).
- [203] Y. Zhao, N. Gonzalez-Garcia, and D. G. Truhlar, J. Phys. Chem. A **109**, 2012 (2005).
- [204] J. Zheng, Y. Zhao, and D. G. Truhlar, J. Chem. Theory Comput. **3**, 569 (2007).
- [205] J. M. L. Martin and G. de Oliverira, J. Chem. Phys. **111**, 1843 (1999).
- [206] J. Zheng, J. R. Gour, J. J. Lutz, M. Wloch, P. Piecuch, D. G. Truhlar, J. Chem. Phys. **128**, 044108 (2008).
- [207] B. J. Lynch, Y. Zhao, and D. G. Truhlar, J. Phys. Chem. A **107**, 1384 (2003).
- [208] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
- [209] D. E. Woon and T. H. Dunning, J. Chem. Phys. **98**, 1358 (1993).
- [210] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
- [211] T. H. Dunning, K. A. Peterson, and A. K. Wilson, J. Chem. Phys. **114**, 9244 (2001).
- [212] D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).

- [213] K. A. Peterson and T. H. Dunning, J. Chem. Phys. **117**, 10548 (2002).
- [214] L. A. Curtiss, K. Raghavachari, C. Redfern, V. Rassolov, and J. A. Pople, J. Chem. Phys. **109**, 7764 (1998).
- [215] P. R. Taylor, In *Lecture Notes in Quantum Chemistry*, edited by B. O. Roos, page 406, (Springer-Verlag, Berlin, 1992).
- [216] A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, and J. Olsen, Chem. Phys. Lett. **302**, 437 (1999).
- [217] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).
- [218] W. D. Laidig and H. F. Schaefer, J. Chem. Phys. **74**, 3411 (1981).
- [219] P. Borowski, K. Andersson, P.A. Malmqvist and B. O. Roos, J. Chem. Phys. **97**, 5568 (1992).
- [220] M. L. Leininger and H. F. Schaefer, J. Chem. Phys. **107**, 9059 (1997).
- [221] X. Li and J. Paldus, J. Chem. Phys. **110**, 2884 (1999).
- [222] S. P. Walch and W. A. Goddard III, J. Am. Chem. Soc. **97**, 5319 (1975).
- [223] S. D. Kahn, W. J. Hehre, and J. A. Pople, J. Am. Chem. Soc. **109**, 1871 (1987).
- [224] P. C. Hiberty and C. J. Leforestier, J. Am. Chem. Soc. **100**, 2012 (1978).
- [225] S. E. Wheeler, D. G. Ess, and K. N. Houk, J. Phys. Chem. A **112**, 1798 (2008).
- [226] W. B. DeMore and C. L. Lin, J. Org. Chem. **38**, 985 (1973).
- [227] W. B. DeMore, Int. J. Chem. Kinet. **1**, 209 (1969).
- [228] D. J. Miller, T. E. Nemo, and L. A. Hull, J. Org. Chem. **40**, 2675 (1975).
- [229] S. Jackson and L. A. Hull, J. Org. Chem. **41**, 3340 (1976).
- [230] K. Griesbaum and Y. Dong, J. Prakt. Chem. **339**, 575 (1997).

- [231] O. Horie and G. K. Moortgat, *Acc. Chem. Res.* **31**, 381 (1998).
- [232] J. M. Anglada, R. Crehuet, and J. M. Bofill, *Chem. Eur. J.* **5**, 1809 (1999).
- [233] L. C. Li, P. Deng, M. H. Xu, and N. B. Wong, *Int. J. Quantum Chem.* **98**, 309 (2004).
- [234] W. T. Chan and I. P. Hamilton, *J. Chem. Phys.* **118**, 1688 (2003).
- [235] I. Ljubic and A. Sabljic *J. Phys. Chem. A* **106**, 4745 (2002).
- [236] J. Z. Gillies, C. W. Gillies, F. J. Lovas, K. Matsumura, R. D. Suenram, E. Kraka, and D. Cremer, *J. Am. Chem. Soc.* **113**, 6408 (1991).
- [237] M. Olzmann, E. Kraka, D. Cremer, R. Gutbrod, and S. Anderson, *J. Phys. Chem. A* **101**, 9421 (1997).
- [238] D. Cremer, R. Crehuet, J. Anglada, *J. Am. Chem. Soc.* **123**, 6127 (2001).
- [239] D. Cremer, E. Kraka, R. Crehuet, J. Anglada, and J. Grafenstein, *Chem. Phys. Lett.* **347**, 268 (2001).
- [240] J. Z. Gillies, C. W. Gillies, F. J. Lovas, K. Matsumura, R. D. Suenram, E. Kraka, and D. Cremer, *J. Am. Chem. Soc.* **113**, 2412 (1991).
- [241] W. T. Chan, C. Weng, and J. D. Goddard, *J. Phys. Chem. A* **111**, 4792 (2007).
- [242] Y. Zhao, O. Tishchenko, J. R. Gour, W. Li, J. J. Lutz, P. Piecuch, and D. G. Truhlar, *J. Phys. Chem. A* **113**, 5786 (2009).
- [243] Y. Zhao, N. E. Schultz, and D. G. Truhlar, *J. Chem. Theory Comput.* **2**, 364 (2006).
- [244] S. G. V. Ornum, R. M. Champeau, and R. Pariza, *Chem. Rev.* **106**, 2990 (2006).
- [245] K. B. Wiberg and J. M. Lavanish, *J. Am. Chem. Soc.* **88**, 5272 (1966).
- [246] G. L. Closs and P. E. Pfeffer, *J. Am. Chem. Soc.* **90**, 2452 (1968).
- [247] R. B. Woodward and R. Hoffman, *The Conservation of Orbital Symmetry* (Academic, New York, 1970).

- [248] K. A. Nguyen and M. S. Gordon, J. Am. Chem. Soc. **117**, 3835 (1995).
- [249] P. B. Shevlin and M. L. McKee, J. Am. Chem. Soc. **110**, 1666 (1988).
- [250] M. J. S. Dewar and S. Kirschner, J. Am. Chem. Soc. **97**, 2931 (1975).
- [251] R. Srinivasan, A. Levi, and I. Haller, J. Phys. Chem. **58**, 1775 (1965).
- [252] K. B. Wiberg and R. A. Fenoglio, J. Am. Chem. Soc. **90**, 3395 (1968).
- [253] D. A. Mazziotti, J. Phys. Chem. A **112**, 13684 (2008).
- [254] R. Berner and A. Luchow, J. Phys. Chem. A **114**, 13222 (2010).
- [255] J. Shen and P. Piecuch, Chem. Phys., *submitted* (2011).
- [256] W.L. Glab, J. Chem. Phys. **107**, 5979 (1997).
- [257] P. Farmanara, O. Steinkellner, M.T. Wick, M. Wittmann, G. Korn, V. Stert, and W. Radloff, J. Chem. Phys. **111**, 6264 (1999).
- [258] M.S. Child and W.L. Glab, J. Chem. Phys. **112**, 3754 (2000).
- [259] J.H. Fillion, R. van Harreveld, J. Ruiz, M. Castillejo, A.H. Zanganeh, J.-L. Lemaire, M.C. van Hemert, and F. Rostas, J. Phys. Chem. A **105**, 11414 (2001).
- [260] W.L. Glab, J. Chem. Phys. **117**, 9316 (2002).
- [261] F. Ederly and A. Kanaev, Eur. Phys. J. D **23**, 257 (2003).
- [262] B.-M. Cheng, C.-Y. Chung, M. Bahou, Y.-P. Lee, L.C. Lee, R. van Harreveld, and M.C. van Hemert, J. Chem. Phys. **120**, 224 (2004).
- [263] J.H. Fillion, J. Ruiz, X.-F. Yang, M. Castillejo, F. Rostas, and J.-L. Lemaire, J. Chem. Phys. **120**, 6531 (2004).
- [264] O. Steinkellner, F. Noack, H.-H. Ritze, W. Radloff, and V. Hertel, J. Chem. Phys. **121**, 1765 (2004).

- [265] D.M. Hirst and M.S. Child, *Molec. Phys.* **77**, 463 (1992).
- [266] M. von Dirke, B. Heumann, K. Kuhl, T. Schroder, and R. Schinke, *J. Chem. Phys.* **101**, 2051 (1994).
- [267] F. Schneider, F. Di Giacomo, and F.A. Gianturco, *J. Chem. Phys.* **104**, 5153 (1996).
- [268] R. van Harrelvelt and M.C. van Hemert, *J. Chem. Phys.* **112**, 5777 (2000).
- [269] R. van Harrelvelt and M.C. van Hemert, *J. Chem. Phys.* **114**, 9453 (2001).
- [270] V. Engel, V. Staemmler, R.L. Vander Wal *et al.*, *J. Phys. Chem.* **96**, 3201 (1992).
- [271] X. Li and J. Paldus, *J. Chem. Phys.* **133**, 024102 (2010).
- [272] C. Tanner, C. Manca, and S. Leutwyler, *Science* **302**, 1736 (2003).
- [273] D. Bruhwiler, G. Calzaferri, T. Torres, J. H. Ramm, N. Gartmann, L. Q. Dieu, I. Lopez-Duarte, and M. V. Martinez-Diaz, *J. Mater. Chem.* **19**, 8040 (2009).
- [274] F. E. Hernandez, S. Yu, M. Garcia, and A. D. Campiglia, *J. Phys. Chem. B* **109** 9499 (2005).
- [275] J. M. Goldberg, S. Batjargal, and E. J. Petersson, *J. Am. Chem. Soc.* **132**, 14719 (2010).
- [276] M. Thut, C. Tanner, A. Steinlin, and S. Leutwyler, *J. Phys. Chem. A* **112**, 5566 (2008).
- [277] T. A. Wesolowski and A. Warshel, *J. Phys. Chem.* **97**, 8050 (1993).
- [278] T. A. Wesolowski, In *Computational Chemistry: Reviews of Current Trends*; edited by J. Leszczyński, Vol. 10, pages 1–82, (World Scientific, Singapore, 2006).
- [279] T. A. Wesolowski, *Phys. Rev. A* **77**, 012504 (2008).
- [280] K. Pernal and T. A. Wesolowski, *Int. J. Quantum Chem.* **109**, 2520 (2009).
- [281] T. Wesolowski, *J. Am. Chem. Soc.* **126**, 11444 (2004).

- [282] W. J. Hehre, R. Ditchfield, and J. A. Pople, **56**, 2257 (1972).
- [283] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
- [284] T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. v. R. Schleyer, J. Comput. Chem. **4**, 294 (1983).
- [285] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, J. Chem. Phys. **72**, 650 (1980).
- [286] A. J. Sadlej Coll. Czech. Chem. Commun. **53**, 1995 (1988).
- [287] S. Coussan, Y. Ferro, A. Trivella, P. Roubin, R. Wieczorek, C. Manca, P. Piecuch, K. Kowalski, M Włoch, S. A. Kucharski, and M. Musiał, J. Phys. Chem. A **110**, 3920 (2006).
- [288] K. Kowalski, S. Krishnamoorthy, O. Villa, J. R. Hammond, N. Govind, J. Chem. Phys. **132**, 154103 (2010).
- [289] M. Ehara, J. R. Gour, and P. Piecuch, Mol. Phys. **107**, 871 (2009).
- [290] J. A. Hansen, P. Piecuch, J. J. Lutz, and J. R. Gour, Phys. Scr. **84**, 028110 (2011).
- [291] T. H. Dunning Jr., J. Chem. Phys. **53**, 2823 (1970).
- [292] T. Nakajima and H. Nakatsuji, Chem. Phys. Lett. **79**, 280 (1997).
- [293] M. Ishida, K. Toyota, M. Ehara, and H. Nakatsuji, **347** 493 (2001).
- [294] M. Ishida, K. Toyota, M. Ehara, H. Nakatsuji, and M. J. Frisch, J. Chem. Phys., **120** 2593 (2004).
- [295] G. Herzberg, *Molecular Spectra and Molecular Structure III: Electronic Spectra and Electronic Structure of Polyatomic Molecules* (Van Nostrand, London, 1967).
- [296] A. J. Merer and D. N. Travis, Can. J. Phys. **44**, 353 (1966).
- [297] A. J. Merer and D. N. Travis, Can. J. Phys. **43**, 1795 (1966).
- [298] Y. Ohtsuka, P. Piecuch, J. R. Gour, M. Ehara, H. Nakatsuji, J. Chem. Phys. **126**, 164111 (2007).

- [299] P. D. Rajendra and P. Chandra, **114**, 1589 (2001).
- [300] J. J. Lutz and P. Piecuch, In *Nuclei and Mesoscopic Physics, Workshop on Nuclei and Mesoscopic Physics, WNMP 2007*, edited by P. Danielewicz, P. Piecuch, and V. Zelevinsky AIP Conferenc Proceedings, Vol. 995, pages 62-71, (American Institute of Physics, Melville, NY, 2008).
- [301] A. J. C. Varandas, Chem. Phys. Lett. **443**, 398 (2007).
- [302] A. J. C. Varandas, J. Chem. Phys. **127**, 114316 (2007).
- [303] K. Hirao and H. Nakatsuji, J. Comput. Phys. **45**, 246 (1982).
- [304] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).