# MERGING ACTIVE-SPACE AND RENORMALIZED COUPLED-CLUSTER METHODS VIA THE CC(P;Q) FORMALISM, WITH APPLICATIONS TO CHEMICAL REACTION PROFILES AND SINGLET–TRIPLET GAPS

By

Nicholas P. Bauman

#### A DISSERTATION

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

Chemistry - Doctor of Philosophy

2016

#### ABSTRACT

#### MERGING ACTIVE-SPACE AND RENORMALIZED COUPLED-CLUSTER METHODS VIA THE CC(P;Q) FORMALISM, WITH APPLICATIONS TO CHEMICAL REACTION PROFILES AND SINGLET-TRIPLET GAPS

#### By

#### Nicholas P. Bauman

The development of accurate and computationally efficient wave function methods that can capture and balance dynamical and non-dynamical many-electron correlation effects to describe multi-reference problems, such as potential energy surfaces involving bond breaking, biradicals, and excited states characterized by dominant many-electron excitations, is one of the main goals of quantum chemistry. Among the promising approaches in this endeavor are the completely renormalized and active-space coupled-cluster (CC) and equation-of-motion (EOM) CC methods. While the completely renormalized and active-space CC and EOMCC approaches have been very successful in many applications, there are some cases where they do not capture the dynamical or non-dynamical many-electron correlation effects in a satisfactory manner. In this dissertation, we introduce the CC(P;Q) formalism, which alleviates this concern by combining the completely renormalized and active-space together. The CC(P;Q) scheme provides a systematic approach to correcting energies obtained in the active-space CC and EOMCC calculations that recover much of the non-dynamical and some dynamical many-electron correlation effects for the remaining, mostly dynamical, correlation effects missing in the active-space CC and EOMCC considerations. We discuss the development of the CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) methods, which use the CC(P,Q) formalism to correct energies obtained with the CC and EOMCC approaches with singles, doubles, and active-space triples (CCSDt/EOMCCSDt) for missing triple excitations (CC(t;3)), or to correct energies obtained with the CC and EOMCC approaches with singles, doubles, and active-space triples and quadruples (CCSDtq/EOMCCSDtq) for missing triples (CC(t,q;3)) or missing triples and quadruples (CC(t,q;3,4)), or even to correct energies obtained with the CC and EOMCC approaches with singles, doubles, triples, and active-space quadruples (CCSDTq/EOMCCSDTq) for correlation effects due to the missing quadruple excitations (CC(q;4)). By examining the double dissociation of water, the Be + H<sub>2</sub>  $\rightarrow$  HBeH insertion, and the singlet-triplet gaps in the strongly biradical (HFH)<sup>-</sup> system and the BN molecule, we demonstrate that the CC(t;3), CC(t,q;3), and CC(t,q;3,4) methods reproduce the total and relative energies obtained with the parent full CC/EOMCC approaches with singles, doubles, and triples or singles, doubles, triples, and quadruples to within fractions of a millihartree at the tiny fraction of the computer cost, even when the electronic quasi-degeneracies become substantial.

The CC(P,Q) formulation prompted the development of efficient CCSDt, CCSDtq, and CCSDTq programs. In this dissertation, we describe the technique of spin-integration for both closed and open shells, and how the resulting equations for CCSDTQ were automatically derived and implemented in a factorized form. We also discuss how the efficiency of the code was improved by removing unnecessary operations through, in particular, the reorganization of the relevant loops. Finally, we explain how the CCSDTQ code was transformed to obtain the active-space CCSDtq and CCSDTq approaches, which are the most essential parts of the CC(t,q;3), CC(t,q;3,4), and CC(q;4) calculations. Copyright by NICHOLAS P. BAUMAN 2016 This is dedicated to my lovely wife Jennifer.

#### ACKNOWLEDGMENTS

I would like to thank my PhD. advisor, Professor Piotr Piecuch, for his patience and guidance throughout my graduate career. Since my first day in his group he has driven me to be a better scientist and shown me my potential. Through many late nights and long discussions he has provided me with a wealth of knowledge that I never anticipated. I am forever grateful for his continuous support and everything he has done for me.

I would like to thank the other members of my committee, Professor Benjamin G. Levine, Professor Robert I. Cukier, and Professor Marcos Dantus for all their support, advice, and patience in overseeing my graduate study.

I would also like to thank Dr. Jun Shen from our group who I owe a great deal of gratitude too. He has not only made the development work presented in this dissertation possible, but has provided valuable guidance and has always been patient and willing to answer any questions I may have had. Last, but not least, I would like to extend my thanks to graduate students from the Piecuch research group I have had an opportunity to interact with, especially, Dr. Jared A. Hansen, Mr. Adeayo O. Ajala, Mr. Jorge Emiliano Deustua, and Mr. Ilias Magoulas, for the many thought provoking discussions and all the help they have provided me over the years.

## TABLE OF CONTENTS

LIST (	OF T.	ABLES	viii
LIST (	OF F	IGURES	x
Chapte	er 1	Introduction	1
Chapte	er 2	Project Objectives	18
Chapte	er 3	Merging Active-Space and Completely Renormalized Coupled-	
		Cluster Methods	19
3.1	Theo	ory	19
	3.1.1	Completely Renormalized Coupled-Cluster Approaches	20
	3.1.2	Active-Space Coupled-Cluster Methods	31
	3.1.3	The $CC(P;Q)$ Formalism	42
	3.1.4	CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) Hierarchy	50
3.2	Num	erical Results	61
	3.2.1	The $C_{2v}$ -Symmetric Double Dissociation of Water $\ldots \ldots \ldots$	61
	3.2.2	$Be + H_2 \rightarrow HBeH C_{2v}$ Insertion Pathway	73
	3.2.3	Singlet–Triplet Gap in HFH <sup>-</sup> Along the $D_{\infty h}$ -Symmetric Double-Dissoc	iation
		Pathway	81
	3.2.4	Singlet–Triplet Gap in BN	100
Chapte	er 4	Algorithmic Advances: Efficient Automated Implementation of	
		Active-Space CCSDtq and CCSDTq Methods, and Their Full	
		CCSDTQ Counterpart	111
4.1	Spin	Integration for Closed and Open Shells	112
4.2	4.2 Automated Approach to Derivation and Computer Implementation of Couple		
Cluster Methods in Factorized Form		ter Methods in Factorized Form	125
4.3	Improvements in Efficiency via Loop Reorganization		137
4.4	Transformation from CCSDTQ to the CCSDtq and CCSDTq Methods 141		
Chapte	er 5	Conclusions and Future Outlook	145
BIBLI	OGR	APHY	148

### LIST OF TABLES

Dependence of the CPU steps on $n_o$ , $n_u$ , $N_o$ , and $N_u$ for the most expensive terms for various CC approximations. <sup><i>a</i></sup>	5
A comparison of various CC ground-state energies for the reactant and transition-state species defining the automerization of cyclobu- tadiene, based on the geometries optimized in the MR-AQCC calcu- lations in Ref. [254] as well as purely electronic barrier heights (in kcal/mol). <sup><math>a,b</math></sup>	9
The complete set of projections entering the CCSDTQ amplitude equations organized according to their active and inactive character. Upper-case bold letters represent active labels and lower-case bold letters represent inactive labels (cf. the text and Fig. 3.1 for details).	36
A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts and the corresponding full CI data for the equilibrium and two displaced geometries of the $\rm H_2O$ molecule, as described by the DZ basis set [279].	69
A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts and the corresponding full CI data for the equilibrium and four displaced geometries of the H <sub>2</sub> O molecule, as described by the spherical cc-pVDZ basis set [237].	71
Coordinates of points along the sampling path $C_{2v}$ -symmetric path describing insertion of Be into H <sub>2</sub> , introduced in Ref. [283]. Structure A is the linear HBeH product, while structure I represents the Be + H <sub>2</sub> reactant. Structure E is the transition state.	77
A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the geometries A–I defining the $C_{2v}$ -symmetric insertion pathway of Be into H <sub>2</sub> , as described by the [3s1p/2s] basis set, introduced in Ref. [283]. <sup><i>a</i></sup>	78
	Dependence of the CPU steps on $n_o$ , $n_u$ , $N_o$ , and $N_u$ for the most expensive terms for various CC approximations. <sup><i>a</i></sup>

Table 3.6:	A comparison of the energies resulting from various CC calculations including up to triple and up to quadruple excitations with their par- ent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the $X  {}^{1}\Sigma_{g}^{+}$ state of the linear, $D_{\infty h}$ -symmetric, (HFH) <sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at a few values of the H–F distance $R_{\text{H-F}}$ (in Å). <sup><i>a</i></sup>	91
Table 3.7:	A comparison of the energies resulting from various CC calculations including up to triple and up to quadruple excitations with their par- ent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the $A^{3}\Sigma_{u}^{+}$ state of the linear, $D_{\infty h}$ -symmetric, (HFH) <sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at a few values of the H–F distance $R_{\text{H-F}}$ (in Å). <sup><i>a</i></sup>	94
Table 3.8:	A comparison of the $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$ gap values resulting from various CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the linear, $D_{\infty h}$ -symmetric, (HFH) <sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at a few values of the H–F distance $R_{\text{H-F}}$ (in Å). <sup><i>a</i></sup>	97
Table 3.9:	Equilibrium bond lengths (in Å) for the lowest triplet and singlet states of BN, the corresponding adiabatic singlet-triplet splittings $T_e$ (in cm <sup>-1</sup> ), and energies relative to CCSDT and CCSDTQ (in cm <sup>-1</sup> ), as obtained with the cc-pVDZ basis set. <sup><i>a</i></sup>	109
Table 3.10:	Equilibrium bond lengths (in Å) for the lowest triplet and singlet states of BN, the corresponding adiabatic singlet-triplet splittings $T_e$ (in cm <sup>-1</sup> ), and energies relative to CCSDT and CCSDTQ (in cm <sup>-1</sup> ), as obtained with the cc-pVTZ basis set. <sup><i>a</i></sup>	110
Table 4.1:	Matrix elements for the standard Hamiltonian and amplitudes for the cluster operators that appear in CCSDTQ, along with their cor- responding cases for spin-integrated equations, and the respective antisymmetrizers for the spin cases	121
Table 4.2:	Excited determinants entering into Eq. (3.11) for spin-orbital imple- mentations of CCSDTQ and the corresponding spin variants for the spin-integrated implementations.	123
Table 4.3:	Average iteration time, in seconds, for calculations of the $X  {}^{1}\Sigma_{g}^{+}$ and $A  {}^{3}\Sigma_{u}^{+}$ states of (HFH) <sup>-</sup> , computed using our original implementation of CCSDTQ, the improved version of CCSDTQ, and NWCHEM for comparison. <sup><i>a</i></sup>	140

#### LIST OF FIGURES

- Figure 1.1: Automerization of cyclobutadiene. The leftmost and rightmost structures correspond to the degenerate reactant/product minima, whereas the structure in the middle represents the transition state. . . . . .
- Figure 3.1: The orbital classification used in the active-space SRCC methods, such as CCSDt, CCSDtq, and CCSDTq. Core, active, and virtual orbitals are represented by solid, dashed, and dotted lines, respectively. Full and open circles represent core and active electrons of the reference determinant  $|\Phi\rangle$  (a closed-shell reference  $|\Phi\rangle$  is assumed).
- Figure 3.2: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations, along with their parent full CCSDT and CCSDTQ counterparts, for the equilibrium and two displaced geometries of the H<sub>2</sub>O molecule, as described by the DZ basis set [279]. The numerical values of the errors are found in Table 3.2. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.
- Figure 3.3: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the equilibrium and four displaced geometries of the H<sub>2</sub>O molecule, as described by the spherical cc-pVDZ basis set [237]. The numerical values of the errors are found in Table 3.3. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results.

70

72

10

- Figure 3.4: A comparison of the energies resulting from various all-electron CC calculations, including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the geometries A–I defining the  $C_{2v}$ -symmetric insertion pathway of Be into H<sub>2</sub>, as described by the [3s1p/2s] basis set, introduced in Ref. [283]. The numerical values of the errors are found in Table 3.5. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results. In presenting the results for the transition state structure E, we used the same reference  $|\Phi_2\rangle$  as that used in the F–I region. . . .

80

Figure 3.7:	A comparison of the $A  {}^{3}\Sigma_{u}^{+} - X  {}^{1}\Sigma_{g}^{+}$ gap values resulting from various CC calculations including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the linear, $D_{\infty h}$ -symmetric, (HFH) <sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at several values of the H–F distance $R_{\text{H-F}}$ . The numerical values of the errors are found in Table 3.8: Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ	
	results	99
Figure 4.1:	The one-body vertex representing the one-body component $F_N = f_p^q N[a^p a_q]$ of $H_N$	115
Figure 4.2:	The two-body vertex representing the two-body component $V_N = \frac{1}{4} v_{pq}^{rs} N[a^p a^q a_s a_r]$ of $H_N$ .	115
Figure 4.3:	The one-, two-, three-, and four-body vertices representing the $T_1$ , $T_2$ , $T_3$ , and $T_4$ cluster operators	116
Figure 4.4:	Diagram representing one of the $\frac{1}{2}v_n t_2^2$ terms that appear in the projection of the CC equations on to doubly excited determinants	117
Figure 4.5:	Sample output of the Latex files corresponding to the $H_N T_{3B}$ term generated by the automatic derivation code. See text for a description of the notation.	131
Figure 4.6:	The general structure of each binary tensor product as produced by the automatic derivation and implementation program	133
Figure 4.7:	Example of binary tensor product between a two-body Hamiltonian operator and a $T_{2A}$ cluster operator that produces and intermediate labeled S14, which is summed with other intermediates and Hamiltonian matrix elements in a collective intermediate labeled X2	134
Figure 4.8:	An example of one of the "REORDER" routines. The number in the title of the routine reflects the order that the indices are permuted to.	135
Figure 4.9:	An example of one of the "SUM" routines. The number in the title of the routine reflects the order that the indices are permuted to	136

Figure 4.10:	An example of the product between an intermediate labeled U41 and the cluster operator $T_{4B}$ , which is projected onto the excited deter-		
	minant $\langle \Phi_{ijk\tilde{l}}^{abcd}  $ . The array corresponding to the product formed is		
	labeled $Z140$ and is added to the residual named V2B	139	
Figure 4.11:	An example of the product between an intermediate labeled X2 and the cluster operator $T_{2A}$ , which is projected onto the excited determinant $\langle \Phi_{ijk}^{abc}   \dots $	142	
Figure 4.12:	An example of the product between an intermediate labeled S5 and the cluster operator $T_{4A}$ , which is projected onto the excited determinant $\langle \Phi_{ijkl}^{abcd}   \dots $	143	

## Chapter 1

## Introduction

One of the main challenges in quantum chemistry is the accurate description of quasidegenerate electronic states in instances involving biradicals, bond breaking, and excited states characterized by dominant two-electron excitations. These situations are a challenge due to the strong non-dynamical many-electron correlation effects that have to be properly balanced with dynamical correlations. The framework of coupled-cluster (CC) theory [1–6] is ideal for handling these types of situations as it offers the best balance between accuracy [fast convergence to the exact, full configuration interaction (CI), limit] and computational costs.

Within the CC framework, two typical routes taken to describe situations involving strong non-dynamical correlation effects are: 1) single-reference (SR) CC methods, and 2) multi-reference (MR) CC approaches. Situations involving strong non-dynamical correlations, such as those described above, are often described as MR problems, so it would seem most natural to turn to MR methods to solve them. In these approaches, one introduces a multi-dimensional model space consisting of a number of reference determinants, which are obtained by distributing active electrons among active orbitals in various ways and which are chosen such as to provide a reasonable zeroth-order description of the quasi-degenerate electronic state(s) of interest, as in the complete-active-space self-consistent-field (CASSCF) calculations. One can then, for example, use the Jeziorski-Monkhorst ansatz [7] to capture the remaining dynamical correlation effects through excitations from each reference determinant. Within this framework, one can formulate the state-universal (SU) (see, e.g., Refs. [7–47]) or the state-specific (SS) (see Refs. [41,42,46–66] for representative examples and recent advances) MRCC methods. However, this formulation is just one possible direction in the MR framework, and unfortunately, unlike SRCC methods, there is no unambiguous way of writing the exponential wave function ansatz for MR approaches. Despite more than three decades of active development, MRCC methods continue to face various formal and practical challenges. For example, the genuine SUMRCC approaches suffer from the intruder state or intruder solution problem, singularities, and multiple unphysical states [12–15, 18, 67]. While SSMRCC approximations may not be afflicted by these issues, at least in principle, they still suffer from convergence problems, particularly in excited-state considerations and, especially, when one or more coefficients at the reference determinants become small [41,66]. In the end, none of the existing MRCC methods based on the Jeziorski-Monkhorst ansatz are characterized by an ease of use and application of approaches as those of the SRCC ansatz, and the same is true for a plethora of other MRCC methods that use other forms of the wave function (see, e.g., Refs [22] and [68] for reviews). For these reasons and others, this work focuses on the approaches of the SRCC type, with the objective of recovering strong non-dynamical correlation effects dynamically through excitations from a single reference determinant.

The basic SRCC approximations, such as CC with singles and doubles (CCSD) [69–72], can be applied to systems containing dozens of non-hydrogen atoms or hundreds of correlated electrons, and hundreds or even thousands of basis functions, in part due to the relatively inexpensive CPU steps that scale as  $n_o^2 n_u^4$ , where  $n_o$  ( $n_u$ ) is the number of occupied (unoccupied) orbitals, or as  $\mathcal{N}^6$ , with the system size  $\mathcal{N}$ , which can be further reduced to linear scaling steps via local correlation approaches (see, e.g., Refs. [73–80] and numerous

references therein). A comparison of CPU time scalings of canonical CCSD with other CC methods mentioned in this dissertation is shown in Table 1.1. While CCSD is generally more accurate than its CI counterpart (i.e., CISD), especially in larger systems where the lack of size-extensivity of CISD becomes a major problem, it has serious difficulties with capturing non-dynamical correlation effects characterizing chemical reaction profiles involving, for instance, bond breaking and biradicals, while missing important dynamical correlations, especially those due to connected triples needed to achieve a quantitative description. The excited-state equation-of-motion (EOM) analog of CCSD, EOMCCSD [81–83], and the corresponding symmetry-adapted-cluster (SAC) CI [84–87] and linear-response CC [88–93] counterparts are capable of describing excited states dominated by one-electron transitions, but are often not accurate enough to obtain a quantitative description of such states, especially when larger polyatomic species are examined (cf., e.g., Refs. [94–97] for selected examples; for a thorough evaluation of a number of EOMCC methods, including EOM-CCSD, illustrating this statement, see Refs. [98–104]). Furthermore, EOMCCSD and its SAC-CI and linear-response counterparts fail at characterizing excited states having significant two- or other many-electron contributions [104–119]. While the higher-order dynamical and stronger non-dynamical correlation effects cannot be captured using the above low-order CC/EOMCC methods, they can be recovered through the explicit and complete inclusion of higher-rank components of the cluster operator T, such as the connected triply and quadruply excited clusters,  $T_3$  and  $T_4$ , respectively, in the ground-state SRCC wave function ansatz  $|\Psi_0\rangle = e^T |\Phi\rangle$  and, in the case of excited states, through the inclusion of the analogous higher-order components of the linear excitation operator  $R_{\mu}$ , i.e.,  $R_{\mu,3}$  and  $R_{\mu,4}$ , in the EOMCC wave function ansatz  $|\Psi_{\mu}\rangle = R_{\mu}e^{T}|\Phi\rangle$ , where  $\mu = 0$  designates the ground state,  $\mu > 0$  labels excited states, and  $|\Phi\rangle$  is the reference determinant [in this document, a restricted Hartree-Fock (RHF or ROHF) configuration]. Unfortunately, the full incorporation of higher-order components of T and  $R_{\mu}$ , as in the CC approach with singles, doubles, and triples (CCSDT) [120, 121], the CC method with singles, doubles, triples, and quadruples (CCSDTQ) [122–125], and their excited-state EOMCC counterparts abbreviated as EOM-CCSDT [107, 108, 126–128] and EOMCCSDTQ [126, 127, 129, 130], leads to large, often prohibitive, computational costs. For example, CCSDT has iterative CPU operations that scale as  $n_o^3 n_u^5$  ( $\mathcal{N}^8$ ) and CCSDTQ, with its iterative  $n_o^4 n_u^6$  ( $\mathcal{N}^{10}$ ) steps, is even more expensive, limiting the use of such schemes to systems of only a dozen or so correlated electrons (see Table 1.1). This restriction has led to the development of various methods that approximate the effects of  $T_3$ ,  $T_4$ ,  $R_{\mu,3}$ , and  $R_{\mu,4}$  components in order to combat these steep CPU time scalings.

Traditional ways of estimating the effects due to higher-than-doubly excited components of the cluster operator T and the EOMCC excitation operator  $R_{\mu}$  rely on manybody perturbation theory (MBPT). Included in this group are the iterative CCSDT-n[131–136] and CCSDTQ-n [137], or non-iterative CCSD[T] [135, 138, 139], CCSD(T) [140], CCSD(TQ<sub>f</sub>) [141], and similar approaches [135, 136, 142–146], and their perturbative excitedstate EOMCC [147–149] and linear response CC [150–153] extensions. While reducing the computational costs of the full CCSDT and CCSDTQ approximations and being useful in practice, these approaches still have serious difficulties in capturing non-dynamical correlation effects characterizing chemical reaction profiles involving bond breaking, biradicals, and excited states having significant two- and other many-electron contributions, which comprise many of the problems we are interested in, especially in areas such as reaction mechanisms and dynamics and photochemistry. Therefore, more robust, yet computationally feasible, approaches must be considered.

	CPU Timing Scalings		
Method	Iterative	Non-iterative	
CCSD	$n_o^2 n_u^4$		
CCSDT	$n_o^3 n_u^5$		
$\mathbf{CCSDTQ}$	$n_o^4 n_u^6$		
$\operatorname{CCSD}(T)$	$n_o^2 n_u^4$	$n_o^3 n_u^4$	
$\operatorname{CCSD}(\operatorname{TQ}_f)$	$n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^2 n_u^5$	
$CR-CC(2,3)_A$	$n_o^2 n_u^4$	$n_o^3 n_u^4$	
$CR-CC(2,3)_D$	$n_o^2 n_u^4$	$n_o^3 n_u^4$	
$CR-CC(2,4)_{AA}$	$n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^2 n_u^5$	
$CR-CC(2,4)_{DA}$	$n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^2 n_u^5$	
$CR-CC(2,4)_{DD}$	$n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^4 n_u^5$	
$\operatorname{CR-CC}(3,4)_A$	$n_o^3 n_u^5$	$n_o^2 n_u^5$	
$\operatorname{CR-CC}(3,4)_{D}$	$n_o^3 n_u^5$	$n_o^4 n_u^5$	
CCSDt	$N_o N_u n_o^2 n_u^4$		
CCSDtq	$N_o^2 N_u^2 n_o^2 n_u^4$		
CCSDTq	$N_o N_u n_o^3 n_u^5$		
$\mathrm{CC}(\mathrm{t};3)_{\mathrm{A}}$	$N_o N_u n_o^2 n_u^4$	$n_o^3 n_u^4$	
$\mathrm{CC}(\mathrm{t};3)_{\mathrm{D}}$	$N_o N_u n_o^2 n_u^4$	$n_o^3 n_u^4$	
$\mathrm{CC}(\mathrm{t},\mathrm{q};3)_{\mathrm{A}}$	$N_o^2 N_u^2 n_o^2 n_u^4$	$n_o^3 n_u^4$	
$\mathrm{CC}(\mathrm{t},\mathrm{q};3)_{\mathrm{D}}$	$N_o^2 N_u^2 n_o^2 n_u^4$	$n_o^3 n_u^4$	
$\mathrm{CC}(\mathrm{t},\mathrm{q};\!3,\!4)_{\mathrm{AA}}$	$N_o^2 N_u^2 n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^2 n_u^5$	
$\mathrm{CC}(\mathrm{t},\mathrm{q};3,4)_{\mathrm{DA}}$	$N_o^2 N_u^2 n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^2 n_u^5$	
$CC(t,q;3,4)_{DD}$	$N_o^2 N_u^2 n_o^2 n_u^4$	$n_o^3 n_u^4 + n_o^4 n_u^5$	
$\rm CC(q;4)_A$	$N_o N_u n_o^3 n_u^5$	$n_o^2 n_u^5$	
$CC(q;4)_D$	$N_o N_u n_o^3 n_u^5$	$n_o^4 n_u^5$	

Table 1.1: Dependence of the CPU steps on  $n_o$ ,  $n_u$ ,  $N_o$ , and  $N_u$  for the most expensive terms for various CC approximations.<sup>*a*</sup>

 $\frac{\text{CC}(q;4)_{\text{D}}}{a} N_o (< n_o \text{ or } \ll n_o) \text{ designates the number of active occupied orbitals and } N_u (\ll n_u) \text{ designates the number of active unoccupied orbitals.}$ 

The completely renormalized (CR) CC and CR-EOMCC schemes and other approximations resulting from the more general theoretical framework termed the method of mo-

ments of CC equations (MMCC) [105, 106, 111–117, 154–165] comprise one of the classes of robust and promising SRCC methods that can be used to solve problems with larger non-dynamical correlations, such as those described above. These methods are based on the idea of adding an *a posteriori*, non-iterative, and state specific correction,  $\delta_{\mu}^{(A)}$ , due to higher-order excitations neglected in the conventional CC/EOMCC calculations defined by some truncation level  $m_A$ , to the corresponding CC/EOMCC energies  $E^{(A)}_{\mu}$  (A refers to the conventional CC/EOMCC approximation we want to correct, say CCSD or EOMCCSD). For example, CR-CCSD(T) [105, 106, 154–156] and CR-CC(2,3) [116, 160–162] add a correction to the CCSD ( $m_A = 2$ ) energy to account for the effects of connected triple excitations, while CR-CCSD(TQ) [105, 106, 154–156] and CR-CC(2,4) [160, 161] add a correction that approximately accounts for the effects of triple and quadruple excitations in a non-iterative manner. These methods and their CR-EOMCC analogs [105,106,113,114,116,117] eliminate or considerably reduce the errors of traditional perturbative CC schemes, such as CCSD[T], CCSD(T),  $CCSD(TQ_f)$ , etc., in the regions of potential energy surfaces involving bond breaking, biradicals, and electronic states characterized by two-electron excitations. The CR-CC, CR-EOMCC, and other MMCC methods are very effective at capturing dynamical correlations while retaining the ease of use of "black-box" approaches.

While the CR-CC and CR-EOMCC approaches have demonstrated considerable success (cf., e.g., Refs. [94–96,104–106,113,114,116–118,154–162,166–215]), even in systems as large as methylcobalamin [175] or oxygen migration on silicon surface [216], none of them are applicable to all MR situations. A fundamental problem is the fact that the CR-CC/CR-EOMCC approaches, like all non-iterative CC/EOMCC methods, rely on a specific lower-order *a priori* CC or EOMCC calculation (e.g., CCSD or EOMCCSD), and thus there is no natural mechanism allowing for the lower-order cluster components, such as  $T_1$  and  $T_2$ , or

their excited-state  $R_{\mu,1}$  and  $R_{\mu,2}$  analogs, to relax in the presence of higher-order  $T_n$  and  $R_{\mu,n}$ components, with  $n \geq 3$ , which is important in cases of stronger non-dynamical correlations. This is discussed in more detail later in this thesis, so at this time we only emphasize that this decoupling can have serious consequences, especially when examining chemical reaction profiles involving biradical transition states, such as, for example, the transition state characterizing the challenging, and frequently studied, automerization of cyclobutadiene [31,65,164,217–232], shown in Figure 1.1 and Table 1.2, or the equally challenging transition state defining the disrotatory pathway for the isomerization of bicyclo [1.1.0] butane to transbuta-1,3-diene, which is one of the two lowest-energy pathways that have been studied in recent years [164, 171, 233–236]. Examining Table 1.2, we can see that for the automerization of cyclobutadiene CR-CC(2,3) can describe the reactant/product minima with errors of less than one millihartree relative to full CCSDT, but has trouble describing the transition state, with errors of 14–15 millihartree relative to CCSDT when the cc-pVDZ and cc-pVTZ basis sets [237] are used. This leads to an overestimation of the barrier height by about 9 kcal/mol, which can be attributed to the neglect of coupling between the lower-order  $T_1$  and  $T_2$  components and  $T_3$ . If one wants to be able to relax  $T_1$  and  $T_2$  or  $R_{\mu,1}$  and  $R_{\mu,2}$  clusters in the presence of higher-order cluster operators, such as  $T_3$  and  $T_4$ , then alternative approaches, other than the non-iterative correction methods discussed so far, must be considered.

Another useful class of robust and promising SRCC-like methods that may handle MR situations is that of the active-space CC and EOMCC approaches [107–109,119,125,163–165, 238–253], which retain the relatively inexpensive computational costs similar to the basic, yet inadequate, CCSD and EOMCCSD approaches, while allowing  $T_1$  and  $T_2$  or  $R_{\mu,1}$  and  $R_{\mu,2}$  to relax in the presence of higher-than-two-body components of T and  $R_{\mu}$ , respectively, through selection of the dominant  $T_n$  and  $R_{\mu,n}$  excitation amplitudes using active orbitals. This results in a hierarchy of methods, such as CC approaches with singles, doubles, and active-space triples (CCSDt/EOMCCSDt), often termed "little t", and singles, doubles, and active-space triples and quadruples (CCSDtq/EOMCCSDtq), usually termed "little tq". There may also be instances involving strong non-dynamical correlation effects that require the full treatment of triples and selection of higher-than-three-body components of T and  $R_{\mu}$ , as in the case of the CC singles, doubles, triples and active-space quadruples (CCS-DTq/EOMCCSDTq) approximation, sometimes termed "little q". These methods naturally approach their parent approaches when all molecular orbitals in the basis set become active. So, for example, CCSDt and CCSDtq become full CCSDT and CCSDTQ, respectively, when all orbitals are active. While active-space CC methods are capable of capturing the vast majority of non-dynamical correlation effects, they are not as efficient in describing dynamical correlations as are the CCSD(T), CR-CC(2,3),  $CCSD(TQ_f)$ , CR-CC(2,4), or similar noniterative approaches. As an example, Table 1.2 for the automerization of cyclobutadiene shows that CCSDt provides a balanced description of the reactant and transition states, producing a barrier height that is only 0.3–1.1 kcal/mol from the value given by CCSDT, but, due to the missing, mainly dynamical, correlation effects, total electronic energies resulting from CCSDt calculations at the reactant/product and transition-state geometries have errors of 20–30 millihartree relative to CCSDT. As shown in [164], larger errors of this magnitude may sometimes provide a few kcal/mol errors in characterizing stationary point energetics of chemical reaction pathways. It would be best to eliminate such large errors to avoid problems of this kind. The active-space CC/EOMCC approximations are no longer pure "black-box" methods since one has to select occupied and unoccupied active orbitals prior to calculation, but this is not the situation of CASSCF-based MR approaches, since the CPU times of the active-space CC/EOMCC calculations scale linearly or quadratically with the numbers of active occupied and active unoccupied orbitals, as opposed to exponential scalings of CASSCF-based methods, i.e., it is trivial to increase the number of active orbitals by a significant factor without running into prohibitive costs. Nevertheless, while the CR-CC/CR-EOMCC and active-space CC/EOMCC approaches have their respective strengths, it is clear that both types of methodologies have weaknesses which need to be addressed.

Table 1.2: A comparison of various CC ground-state energies for the reactant and transitionstate species defining the automerization of cyclobutadiene, based on the geometries optimized in the MR-AQCC calculations in Ref. [254] as well as purely electronic barrier heights (in kcal/mol).<sup>a,b</sup>

Method	Reactant	Transition State	Barrier Height
		cc-pVTZ	
CCSD	26.837	47.979	20.9
$\operatorname{CCSD}(T)$	1.123	14.198	15.8
$\operatorname{CR-CC}(2,3)_{\mathrm{D}}$	0.848	14.636	16.3
CCSDt	20.786	20.274	7.3
$CC(t;3)_D$	-0.137	0.071	7.8
CCSDT	-154.244157	-154.232002	7.6
		cc- $pVTZ$	
CCSD	36.106	55.205	22.6
$\operatorname{CCSD}(T)$	0.278	12.291	18.1
$\operatorname{CR-CC}(2,3)_{\mathrm{D}}$	0.941	13.793	18.6
CCSDt	30.007	28.259	9.5
$CC(t;3)_D$	-0.141	-1.038	10.0
CCSDT	-154.390763	-154.373902	10.6

 $^{a}$  All results were taken from Ref. [164].

 $^{b}$  The CCSDT values are total energies, in hartree. The remaining energies, excluding the barrier heights, represent errors relative to CCSDT, in millihartree. The active space used in the CCSDt, CCSD(T)-h, and CC(t;3) calculations consisted of the one highest-energy occupied and one lowest-energy unoccupied orbitals

One way to improve upon the deficiencies of the active-space CC/EOMCC and CR-CC/CR-EOMCC methods, while still retaining their respective advantages, is by combining



Figure 1.1: Automerization of cyclobutadiene. The leftmost and rightmost structures correspond to the degenerate reactant/product minima, whereas the structure in the middle represents the transition state.

the two types of approaches into a single formalism. This is accomplished by generalizing the existing biorthogonal MMCC theory, such that one can correct the CC/EOMCC energies obtained with the arbitrary, i.e., conventional (e.g., CCSD/EOMCCSD or CCS-DT/EOMCCSDT) as well as unconventional (e.g., CCSDt/EOMCCSDt or CCSDtq/EOM-CCSDtq) truncations in T and  $R_{\mu}$  for essentially any subset of the missing many-electron correlation effects of interest. The resulting moment expansions, defining the CC(P;Q)formalism [163–165, 251–253] developed in our group, enable one to contemplate a variety of new schemes addressing the above concerns with the CR-CC/EOMCC and active-space CC/EOMCC approaches. This work is focused on the ensuing hierarchy based on correcting energies obtained in active-space CC and EOMCC calculations for missing, primarily dynamical, higher-order correlations. For example, the basic CC(P;Q) method, abbreviated as CC(t;3), corrects the energy obtained at the CCSDt level for the remaining correlation effects due to connected triple excitations missing in CCSDt, but present in the MMCCbased CR-CC(2,3) non-iterative correction. It has been shown [163-165, 251-253] that the CC(t;3) method improves upon both the CCSDt and CR-CC(2,3) results, while providing potential energy surfaces along bond breaking coordinates in close agreement with the results obtained using the full treatment of triples, CCSDT, to within a small fraction of a millihartree and at a fraction of the computational cost. Other methods in the CC(P;Q) hierarchy include CC(t,q;3), CC(t,q;3,4), and CC(q;4), as well as their excited-state analogs, in which energies obtained using active-space CCSDtq/EOMCCSDtq are corrected for missing higher-order correlations, such as triples (CC(t,q;3)) or triples and quadruples (CC(t,q;3,4)), or in which energies obtained using active-space CCSDTq/EOMCCSDTq are corrected for missing higher-order correlations, such as quadruples (CC(q;4)), using the non-iterative energy corrections similar to those defining the previously discussed CR-CC approaches. In the same fashion as CC(t;3), the CC(t,q;3), CC(t,q;3,4), and CC(q;4) approaches have been shown [251–253] to improve upon both of their standard active-space and CR-CC components, reproducing results obtained using the full treatment of quadruples, CCSDTQ, once again to within a small fraction of a millihartree, and at a fraction of the computational cost. This thesis focuses on the development of the CC(t,q;3), CC(t,q;3,4), and CC(q;4)approaches and the corresponding computer codes used in Refs. [251–253].

Our interest in developing the CC(t;3), CC(t,q;3), CC(t,q;3,4), CC(q;4), etc. hierarchy and the underlying CC(P; Q) methodology has been inspired, to some extent, by the earlier work by Li and Paldus [255, 256], who decided to correct the reduced MRCCSD (RMR-CCSD) [257] energies, which contain some, but not all triples, for the triples correlation effects missing in RMRCCSD using the (T) corrections of CCSD(T), resulting in the RMR-CCSD(T) approach, and the analogous effort by Li *et al.* [258–261], who proposed to do the same within the active-space CCSDt framework, once again using the standard CCSD(T) expression to correct the CCSDt energies for the subset of triples missing in CCSDt. In particular, as explained in Refs. [163–165], the ground-state CC(t;3) approach can be regarded as an improvement over the CCSD(T)-h method of Li *et al.* [258–261], in which the perturbative (T) correction of CCSD(T) exploited in CCSD(T)-h to correct the CCSDt energies for the subset of triples missing in CCSDt is replaced by the much more robust CR-CC(2,3)-type expression resulting from the CC(P;Q) considerations. Similarly, the more recent CCSD(T)q-h approach developed in Ref. [262], in which one uses the (T) correction of CCSD(T) to correct the CCSDtq energies for the subset of triples missing in CCSDtq, is an approximation to the ground-state CC(t,q;3) method, which we formally proposed in Ref. [163] and which we implement and test, along with its CC(t,q;3,4) extension, in the present thesis research.

While recognizing the similarities between the CC(t;3) and CC(t;3) approaches and their CCSD(T)-h and CCSD(T)q-h counterparts, as summarized above, we should also point out the differences between the former and the latter schemes. First and foremost is the fact that the CC(t;3) and CC(t,q;3) approximations are part of a larger, systematically improvable CC(t;3), CC(t,q;3), CC(t,q;3,4), CC(q;4), etc. hierarchy introduced in Refs. [163, 164] and further developed here, which results from the rigorous moment energy expansions defining the CC(P;Q) framework. The CCSD(T)-h and CCSD(T)q-h methods of Refs. [258–262] and their RMRCCSD(T) predecessor proposed in Refs. [255, 256] are based on intuitive, ad hoc arguments, which are aimed at correcting the respective CCSDt, CCSDtq, and RMRCCSD energies for the missing triple excitation effects, without a proof that the (T) correction of CCSD(T) represents an appropriate mathematical expression to do it. Indeed, one must keep in mind that the (T) correction of CCSD(T) was originally derived [140] and subsequently rederived [160, 161, 263] as a quasi-perturbative correction to CCSD, not RMRCCSD, CCSDt, or CCSDtq. It is also unclear how to extend the CCSD(T)h, CCSD(T)q-h, and RMRCCSD(T) approaches to higher-order (i.e., higher-than-triple or higher-than-quadruple) excitations and, what is especially important in this work, what formulas to use to correct the CCSD(T)q-h and RMRCCSD(T) energies for the subsets of quadruples still missing in these schemes. None of this is a problem when the CC(P;Q)based CC(t;3), CC(t,q;3), CC(t,q;3,4), CC(q;4), etc. hierarchy is considered, since one can easily define the relevant higher-order schemes converging to the exact, full CI limit and having well-defined relationships with the parent schemes from a traditional CCSDT, CCS-DTQ, etc. CC sequence via suitable choices of the P and Q subspaces of the many-electron Hilbert space entering the CC(P; Q) expressions reported in Refs. [163, 164]. Furthermore, as shown in Refs. [163–165], it is safer to use the CC(P;Q) ideas to design the non-iterative corrections to the active-space CC methods, such as CCSDt, since the CC(P;Q)-inspired corrections, such as that defining CC(t;3), always bring the underlying CCSDt energies to a closer agreement with the parent full CCSDT results, whereas the CCSD(T)-inspired correction to CCSDt defining the aforementioned CCSD(T)-h approach may change the CCSDt energies in an opposite direction, i.e., away from CCSDT, which is an undesired behavior. Finally, although this study focuses on extending the earlier CC(t;3) work [163–165] to the higher-level CC(t,q;3) and CC(t,q;3,4) schemes, as applied to the ground electronic states, it should also be pointed out that unlike RMRCCSD(T), CCSD(T)-h, and CCSD(T)q-h, which apply to the ground-state problem only, the CC(t;3), CC(t,q;3), CC(t,q;3,4), CC(q;4), etc. hierarchy, and the underlying CC(P; Q) formalism have natural extensions to excited states within the EOMCC framework [163]. In this thesis, we focus on the ground electronic states and CC(t,q;3), CC(t,q;3,4), and CC(q;4) approaches that describe the combined effects of triples and quadruples.

The key feature of the CC(P;Q) formalism, which we would like to emphasize here, is as follows. The non-iterative corrections to the CCSDt, CCSDtq, and other active-space SRCC energies defining the CC(t;3), CC(t,q;3), CC(t,q;3,4), and similar approaches provide us with a conceptually straightforward and relatively inexpensive mechanism for relaxing the singly and doubly excited components of the cluster operator T,  $T_1$  and  $T_2$ , respectively, in the presence of higher-than-doubly excited clusters, such as  $T_3$  and  $T_4$ , without having to turn to the prohibitively expensive full CCSDT or CCSDTQ methods. In analogy to the conventional CCSD(T) and  $CCSD(TQ_f)$  schemes, the previous generations of the noniterative corrections exploiting the MMCC formalism, such as those defining the CR-CC(2,3)and CR-CC(2,4) methods, rely on the values of  $T_1$  and  $T_2$  determined by solving the standard CCSD equations. As already pointed out, the  $T_1$  and  $T_2$  clusters obtained with CCSD, where one assumes that they can be decoupled from the  $T_n$  components of T with  $n \ge 3$ , and the  $T_1$  and  $T_2$  clusters determined in the presence of  $T_n$ 's with  $n \ge 3$  can be quite different, particularly when the connected triples or the connected triples and quadruples become larger. Indeed, as demonstrated in Refs. [163–165], and as illustrated in Table 1.2, there are classes of chemical reactions involving biradical transition states, examples of singlet-triplet gaps in biradical systems, and cases of bond breaking, where the neglect of the coupling between the lower-order  $T_1$  and  $T_2$  components and  $T_3$  leads to rather large errors in the results of the CR-CC(2,3) and other similar non-iterative-based calculations, such as  $\Lambda$ -CCSD(T) [264,265] and CCSD(2)<sub>T</sub> [266], which are comparable, in absolute value, to those obtained with the standard (and failing) CCSD(T) approach. As demonstrated in Refs. [163–165] and as illustrated in Table 1.2, the CC(P;Q)-based CC(t;3) method, in which the  $T_1$  and  $T_2$  clusters that are used to determine the correction due to triples are taken from the CCSDt calculations, where one solves for  $T_1$  and  $T_2$  in the presence of the dominant  $T_3$  contributions captured with the help of active orbitals, addresses this issue, offering substantial improvements in the CR-CC(2,3) and  $CCSD(2)_T$  (also CCSD(T)) results in the bond breaking and biradical situations where these methods perform poorly, while providing the total as well as relative energies that often are within small fractions of a millihartree

relative to the corresponding full CCSDT data without making the calculations significantly more expensive than in the CR-CC(2,3),  $CCSD(2)_T$ ,  $\Lambda$ -CCSD(T), or CCSD(T) case. One of the main objectives of this work is to show that one observes similar improvements in the electronic energies along bond stretching coordinates, when the CR-CC(2,4) corrections due to triples and quadruples, which rely on the  $T_1$  and  $T_2$  clusters obtained with CCSD, are replaced by the CC(t,q;3) and CC(t,q;3,4) corrections to the CCSDtq energies that utilize the  $T_1$  and  $T_2$  amplitudes originating from the CCSDtq calculations. It is true that the active-space SRCC methods, such as CCSDt examined in Refs. [163–165] or CCSDtq examined in this work, which capture the coupling of  $T_1$  and  $T_2$  with the higher-order  $T_n$  components with  $n \geq 3$  selected with the help of active orbitals, provide high-quality relative energetics in many biradical and bond breaking situations too, but, as shown in Refs. [163–165], and as further demonstrated in this thesis, they often fail to provide accurate total energies relative to the parent SRCC approximations due to the neglect of the higher-order  $T_n$ contributions of the primarily dynamical character that do not involve active orbitals and this may negatively affect the resulting relative energetics of chemical reaction pathways (see, e.g., [164]). As shown in Refs. [163-165] using CC(t;3) and as demonstrated in this dissertation using CC(t,q;3) and CC(t,q;3,4), the CC(P,Q) methodology addresses all these issues, since it provides us with a transparent mechanism how to correct the CCSDt and CCSDtq energies for the subsets of triple (CCSDt) or triple and quadruple (CCSDtq) excitations neglected in these active-space approaches. Other ways of accounting for the relaxation of the  $T_1$ and  $T_2$  clusters in the presence of  $T_3$  or  $T_3$  and  $T_4$  within the SRCC framework, which can be viewed as alternatives to the CC(P,Q) schemes considered in this dissertation and the earlier [163–165] work, and its extensions in Refs. [251–253], have been proposed in Refs. [267, 268], where the authors combined full CCSDT calculations in an extended active space with the tailored CC framework [220,269,270] and in Refs. [271–273], where the authors exploited the Lagrangian SRCC and EOMCC frameworks and the appropriate perturbative analysis of the resulting equations. Our focus here is on the CC(P,Q) ideas originating from our group [163–165, 251–253].

In this dissertation, we lay down the theory of the CC(P;Q) methodology and describe the ensuing hierarchy of methods, namely CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4), with a focus on the latter three, which have been implemented in my Ph.D. research. We demonstrate and establish, particularly for approaches involving connected quadruple excitations, the robustness, utility, and accuracy of the CC(t,q;3), CC(t,q;3,4), and CC(q;4) approximations by examining the double dissociation of water, the Be + H\_2  $\rightarrow$  HBeH insertion, and the singlet-triplet gap in the strongly biradical  $(HFH)^{-}$  system [251], where quadruples are important to achieve a quantitative description. After examining these test sets, we apply these approaches to a small, but surprisingly difficult BN molecule, which serves as the final "torture" case for examining the CC(P;Q) methods in this dissertation and which required the development of highly efficient codes for the full and active-space treatment of quadruple excitations. As a result, we describe the technique of spin-integration for both closed and open shells, and how the resulting equations for full CCSDTQ and its active-space CCSDtq and CCSDTq counterparts are automatically derived and implemented in a highly efficient, fully vectorizable, factorized form. We also discuss how the efficiency of the resulting codes was further improved through the reorganization of the underlying loop structure. Finally, we explain how the CCSDTQ code, which can compete with the best codes of this kind around, was transformed to obtain the active-space CCSDtq and CCSDTq approaches. In summary, the focus of this dissertation is the extension of the CC(P;Q) hierarchy to methods involving connected quadruple excitations, as well as demonstrating and establishing that these approaches are capable of reproducing results obtained using the full treatment of quadruples, CCSDTQ, typically to within a fraction of a millihartree, at a tiny fraction of the computational effort of CCSDTQ, even when electronic quasi-degeneracies become substantial, as is the case when bonds are broken or challenging cases of singlet-triplet gaps are examined.

## Chapter 2

# **Project Objectives**

The main objectives of this thesis work are

- A. Introduce the CC(P;Q) theory and describe the ensuing hierarchy of methods, namely the CC(t;3), CC(t,q;3), CC(t,q;3,4) and CC(q;4) approaches.
- B. Investigate the performance of the CC(P;Q) methods by examining the interesting benchmark problems of the double dissociation of water, the Be + H<sub>2</sub>  $\rightarrow$  HBeH insertion, and the singlet-triplet gap in the strongly biradical (HFH)<sup>-</sup> system and by applying these methods to the enormously difficult BN molecule.
- C. Describe the procedure of spin integration for closed and open shell systems and how the spin-integrated CCSDTQ equations were derived, factorized, and translated into FORTRAN code using a program, developed in this work as well, that carries out these procedures automatically.
- D. Discuss how the performance of the spin-integrated CCSDTQ code was further improved by reorganizing the corresponding loop structure, and then describe how the code was modified to obtain the CCSDtq and CCSDTq routines used in the CC(t,q;3), CC(t,q;3,4) and CC(q;4) computations.

## Chapter 3

# Merging Active-Space and Completely Renormalized Coupled-Cluster Methods

### 3.1 Theory

This dissertation is concerned with the development and application of CC and EOMCC approaches that have emerged from the recently proposed CC(P;Q) methodology, with a focus on ground-state methods that efficiently account for the connected quadruple excitations. We begin by reviewing the conventional CC and EOMCC theories, the CR-CC and CR-EOMCC approaches, and the active-space CC and EOMCC methods. Then, we discuss how to improve upon the deficiencies of the CR-CC/CR-EOMCC and active-space CC/EOMCC approximations by combining them together via the CC(P;Q) formalism. We end this section by presenting the detailed equations of the resulting hierarchy of methods implemented to date, namely, the CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) approaches.

#### 3.1.1 Completely Renormalized Coupled-Cluster Approaches

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

$$|\Psi_0\rangle = e^T |\Phi\rangle,\tag{3.1}$$

where T is the cluster operator (the connected particle-hole excitation operator) and  $|\Phi\rangle$  is the reference (usually Hartree–Fock) determinant. Typically, we truncate the many-body expansion of T at a conveniently chosen excitation rank  $m_A \leq N$ , to obtain an approximate T, i.e.,  $T \simeq T^{(A)}$ . The truncated cluster operator  $T^{(A)}$  defining the approximate CC method A is given by

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

with

$$T_{n} = \sum_{\substack{i_{1} < \dots < i_{n} \\ a_{1} < \dots < a_{n}}} t_{a_{1}\dots a_{n}}^{i_{1}\dots i_{n}} E_{i_{1}\dots i_{n}}^{a_{1}\dots a_{n}},$$
(3.3)

where  $T_n$  is the *n*-body component of  $T^{(A)}$ ,  $t_{a_1...a_n}^{i_1...i_n}$  are the cluster amplitudes, and

$$E_{i_1\dots i_n}^{a_1\dots a_n} = \prod_{\kappa=1}^n a^{a_\kappa} a_{i_\kappa} \tag{3.4}$$

is the usual *n*-body excitation operator with  $a^p$  and  $a_p$  designating the creation and annihilation operators associated with the spin-orbital basis  $\{|p\rangle\}$ . Here and elsewhere in this thesis, we use the standard notation in which the indices  $i_1, i_2, \ldots$  or  $i, j, \ldots (a_1, a_2, \ldots$  or  $a, b, \ldots$ ) refer to the occupied (unoccupied) spin-orbitals in the reference determinant  $|\Phi\rangle$ . If  $m_A = N$ , we obtain the full CC wave function, which is equivalent to the full CI ground

state corresponding to the exact solution of the electronic Schrödinger equation in the basis set. When the cluster operator is truncated such that  $m_A < N$ , we obtain the well-known hierarchy of standard CC approximations ( $m_A = 2$  for CCSD,  $m_A = 3$  for CCSDT,  $m_A = 4$ for CCSDTQ, etc.). In the case of excited states, we obtain the excited-state wave function  $|\Psi_{\mu}\rangle$  by applying a linear particle-hole excitation operator  $R_{\mu}$  to ground-state SRCC wave function, i.e.,

$$|\Psi_{\mu}\rangle = R_{\mu}|\Psi_{0}\rangle,\tag{3.5}$$

where  $|\Psi_0\rangle$  is defined by Eq. (3.1). This is referred to as the EOMCC formalism. Usually, the  $R_{\mu}$  linear excitation operator is truncated at the same excitation level  $m_A$  as that used in the underlying ground-state CC method A, so that  $R_{\mu}$  is approximated by its truncated counterpart defining the EOMCC method A,

$$R_{\mu}^{(A)} = R_{\mu,0}^{(A)} + R_{\mu,\text{open}}^{(A)} = r_{\mu,0}\mathbf{1} + \sum_{n=1}^{m_A} R_{\mu,n}, \qquad (3.6)$$

in which **1** is the unit operator and

$$R_{\mu,n} = \sum_{\substack{i_1 < \dots < i_n \\ a_1 < \dots < a_n}} r_{\mu,a_1\dots a_n}^{i_1\dots i_n} E_{i_1\dots i_n}^{a_1\dots a_n},$$
(3.7)

where  $R_{\mu,n}$  is the *n*-body component of  $R_{\mu}^{(A)}$  and  $r_{\mu,a_1...a_n}^{i_1...i_n}$  are the excitation amplitudes for the  $\mu$ -th excited state. By varying the truncation level  $m_A$  defining  $R_{\mu}^{(A)}$  and the corresponding  $T^{(A)}$  operator, we obtain the conventional hierarchy of standard EOMCC approximations ( $m_A = 2$  for EOMCCSD,  $m_A = 3$  for EOMCCSDT,  $m_A = 4$  for EOMCCSDTQ, etc.). For consistency of our notation, where  $\mu = 0$  corresponds to the ground state and  $\mu > 0$  represents excited states, we define  $R_{\mu=0}^{(A)} = 1$ , so that  $r_{\mu=0,0}=1$  and  $R_{\mu,n} = 0$  for n > 0. In the ground-state considerations, the cluster amplitudes  $t_{a_1...a_n}^{i_1...i_n}$  defining  $T^{(A)}$ , Eq. (3.2), are obtained by solving the conventional SRCC equations. We arrive at these equations by first inserting the CC wave function  $|\Psi_0\rangle$ , Eq. (3.1), into the electronic Schrödinger equation,

$$H|\Psi_0\rangle = E_0|\Psi_0\rangle,\tag{3.8}$$

and multiplying both sides of Eq. (3.8) on the left by  $e^{-T(A)}$  to obtain the connected cluster form of the Schrödinger equation,

$$\bar{H}^{(A)}|\Phi\rangle = E_0|\Phi\rangle, \qquad (3.9)$$

where

$$\bar{H}^{(A)} = e^{-T^{(A)}} H e^{T^{(A)}} = (H e^{T^{(A)}})_C$$
(3.10)

is the similarity-transformed Hamiltonian and the subscript C denotes the connected part of the corresponding operator expression. Then, we project Eq. (3.9) onto the excited determinants  $|\Phi_{i_1...i_n}^{a_1...a_n}\rangle \equiv E_{i_1...i_n}^{a_1...a_n}|\Phi\rangle$  to obtain the following set of equations:

$$\langle \Phi^{a_1...a_n}_{i_1...i_n} | \bar{H}^{(A)} | \Phi \rangle = 0, \quad i_1 < \ldots < i_n, \quad a_1 < \ldots < a_n,$$
 (3.11)

where  $n = 1, ..., m_A$ . The excited determinants entering Eq. (3.11) correspond to the manybody components  $T_n$  included in  $T^{(A)}$ . For example, in the CCSD calculations  $(m_A = 2)$ , we project Eq. (3.9) on all singly and doubly excited determinants,  $|\Phi_i^a\rangle$  and  $|\Phi_{ij}^{ab}\rangle$ , respectively, while for CCSDTQ  $(m_A = 4)$  we project Eq. (3.9) on all singly, doubly, triply, and quadruply excited determinants,  $|\Phi_i^a\rangle$ ,  $|\Phi_{ijk}^{ab}\rangle$ ,  $|\Phi_{ijk}^{abc}\rangle$ , and  $|\Phi_{ijkl}^{abcd}\rangle$ . Once the system of equations, Eq. (3.11), is solved for the cluster amplitudes  $t_{a_1...a_n}^{i_1...i_n}$ , the ground-state energy of the standard CC method A is calculated using

$$E_0^{(A)} = \langle \Phi | \bar{H}^{(A)} | \Phi \rangle \equiv \langle \Phi | \bar{H}^{(A)}_{\text{closed}} | \Phi \rangle, \qquad (3.12)$$

where  $\bar{H}_{\text{closed}}^{(A)}$  is the 'closed' part of  $\bar{H}^{(A)}$ , represented by the connected diagrams that have no external Fermion lines.

Moving onto the excited states within the EOMCC framework, the excitation amplitudes  $r_{\mu,a_1...a_n}^{i_1...i_n}$  defining the *n*-body components of  $R_{\mu}^{(A)}$  with  $n = 1, ..., m_A$  are obtained by solving the non-hermitian eigenvalue problem involving the similarity-transformed Hamiltonian  $\bar{H}^{(A)}$ , Eq. (3.10), in the space spanned by the excited determinants  $|\Phi_{i_1...i_n}^{a_1...a_n}\rangle$  that correspond to the truncation level of  $R_{\mu}^{(A)}$ ,

$$\langle \Phi_{i_1\dots i_n}^{a_1\dots a_n} | (\bar{H}_{\text{open}}^{(A)} R_{\mu,\text{open}}^{(A)})_C | \Phi \rangle = \omega_{\mu}^{(A)} r_{\mu,a_1\dots a_n}^{i_1\dots i_n}, \quad i_1 < \dots < i_n, \quad a_1 < \dots < a_n, \quad (3.13)$$

where

$$\bar{H}_{\text{open}}^{(A)} = \bar{H}^{(A)} - \bar{H}_{\text{closed}}^{(A)} = \bar{H}^{(A)} - E_0^{(A)} \mathbf{1}$$
(3.14)

is the 'open' part of  $\bar{H}^{(A)}$ , represented by diagrams having external Fermion lines and

$$\omega_{\mu}^{(A)} = E_{\mu}^{(A)} - E_0^{(A)} \tag{3.15}$$

is the vertical excitation energy obtained with the EOMCC method A. Once the excitation amplitudes  $r_{\mu,a_1...a_n}^{i_1...i_n}$  with  $n = 1, ..., m_A$  defining  $R_{\mu,\text{open}}^{(A)}$  ( $\mu > 0$ ) are known, the coefficient  $r_{\mu,0}$  defining the zero-body component  $R_{\mu,0}^{(A)}$  is determined a posteriori using the following
equation:

$$r_{\mu,0} = \frac{\langle \Phi | (\bar{H}_{\text{open}}^{(A)} R_{\mu,\text{open}}^{(A)})_C | \Phi \rangle}{\omega_{\mu}^{(A)}}.$$
 (3.16)

Let us recall that the basic CCSD/EOMCCSD method has iterative CPU steps that scale as  $n_o^2 n_u^4$ , which are practical enough for many situations. Unfortunately, the higher-level CCS-DT/EOMCCSDT and CCSDTQ/EOMCCSDTQ approaches have CPU steps that scale as  $n_o^3 n_u^5$  and  $n_o^4 n_u^6$  respectively, which are prohibitively expensive for systems with more than a dozen or so correlated electrons. Thus, in order to incorporate the physics associated with the  $T_3$ ,  $T_4$ ,  $R_{\mu,3}$ , and  $R_{\mu,4}$  operators, which are needed to obtain an accurate description of dynamical and non-dynamical correlations in MR situations, such as bond breaking, biradicals, and two electron transitions, we must resort to the approximate treatments of these operators that reduce the iterative  $n_o^3 n_u^5$  and  $n_o^4 n_u^6$  steps to a more manageable level. As explained in the Introduction, one of the best approaches to this problem is the MMCC formalism, which allows one to come up with the relatively inexpensive corrections to the energies obtained in the low-order CC/EOMCC calculations, such as CCSD or EOMCCSD, defining the CR-CC and CR-EOMCC approaches, that are more robust in MR situations than the traditional perturbative methods of the CCSD(T) type.

The SR formulation of the MMCC theory relevant to this thesis research is based on the idea of adding the *a posteriori*, non-iterative, and state-specific corrections  $\delta_{\mu}^{(A)}$ , due to higher-order many-body excitations neglected in the conventional CC/EOMCC method A, to the corresponding CC/EOMCC energies  $E_{\mu}^{(A)}$ . The MMCC corrections  $\delta_{\mu}^{(A)}$  can be derived using one of the forms [105, 106, 111, 112, 116, 116, 154–156, 158, 160, 161, 163] of the expansion describing the differences between the exact full CI and CC/EOMCC method A energies, i.e.,

$$\delta_{\mu}^{(A)} \equiv E_{\mu} - E_{\mu}^{(A)}.$$
(3.17)

The name "MMCC" originates from the fact that the  $\delta_{\mu}^{(A)}$  corrections are expressed in terms of the generalized moments of the CC/EOMCC equations, designated as  $\mathfrak{M}_{\mu,a_1\dots a_n}^{i_1\dots i_n}(m_A)$ , characterizing the truncated CC/EOMCC method A we want to correct. These moments are defined as projections of the CC/EOMCC equations written for T approximated by  $T^{(A)}$ and  $R_{\mu}$  approximated by  $R_{\mu}^{(A)}$  on the excited determinants  $|\Phi_{i_1\dots i_n}^{a_1\dots a_n}\rangle$  with  $n > m_A$  that are normally disregarded in the CC/EOMCC calculations, truncated at  $m_A$ -fold excitations, i.e.,

$$\mathfrak{M}_{0,a_1...a_n}^{i_1...i_n}(m_A) = \langle \Phi_{i_1...i_n}^{a_1...a_n} | (\bar{H}^{(A)}) | \Phi \rangle$$
(3.18)

for the ground state, and

$$\mathfrak{M}^{i_1\dots i_n}_{\mu,a_1\dots a_n}(m_A) = \langle \Phi^{a_1\dots a_n}_{i_1\dots i_n} | (\bar{H}^{(A)} R^{(A)}_{\mu}) | \Phi \rangle$$
(3.19)

for excited states.

In order to derive the MMCC corrections  $\delta^{(A)}_{\mu}$ , one typically begins with the asymmetric energy expression

$$E_{\mu} = \langle \Psi_{\mu} | H R_{\mu}^{(A)} e^{T^{(A)}} | \Phi \rangle / \langle \Psi_{\mu} | R_{\mu}^{(A)} e^{T^{(A)}} | \Phi \rangle, \qquad (3.20)$$

where  $\langle \Psi_{\mu} |$  is the full CI bra wave function for the ground  $(\mu = 0)$  or excited  $(\mu > 0)$  state, which gives the exact ground- or excited-state energy,  $E_{\mu}$ , independent of the truncation level  $m_A$  defining  $T^{(A)}$  and  $R^{(A)}_{\mu}$ . We recall that  $R^{(A)}_{\mu}$  is the unit operator in the groundstate  $(\mu = 0)$  case. In the biorthogonal MMCC theory, which interests us here most and which leads to methods mentioned in the Introduction, such as CR-CC(2,3), CR-CC(2,4), CR-EOMCC(2,3), and CR-EOMCC(2,4), we represent the exact bra state  $\langle \Psi_{\mu} |$  in Eq. (3.20) in the following manner:

$$\langle \Psi_{\mu} | = \langle \Phi | \mathscr{L}_{\mu} e^{-T^{(A)}}, \qquad (3.21)$$

where the hole-particle linear deexcitation operator  $\mathscr{L}_{\mu}$  is given by

$$\mathscr{L}_{\mu} = \mathscr{L}_{\mu}^{(A)} + \delta \mathscr{L}_{\mu}^{(A)} \equiv \sum_{n=0}^{m_{A}} \mathscr{L}_{\mu,n} + \sum_{n=m_{A}+1}^{N} \mathscr{L}_{\mu,n}, \qquad (3.22)$$

where  $\mathscr{L}_{\mu,0} = \delta_{\mu,0} \mathbf{1}$  and

$$\mathscr{L}_{\mu,n} = \sum_{\substack{i_1 < \dots < i_n \\ a_1 < \dots < a_n}} \ell^{a_1 \dots a_n}_{\mu, i_1 \dots i_n} E^{i_1 \dots i_n}_{a_1 \dots a_n},$$
(3.23)

with  $E_{a_1...a_n}^{i_1...i_n}$  defined by

$$E_{a_1...a_n}^{i_1...i_n} = \left(E_{i_1...i_n}^{a_1...a_n}\right)^{\dagger} = \prod_{\kappa=1}^n a^{i_\kappa} a_{a_\kappa}.$$
 (3.24)

By inserting Eq. (3.21) into Eq. (3.20) and imposing the normalization condition in which  $\langle \Phi | \mathscr{L}_{\mu}^{(A)} R_{\mu}^{(A)} | \Phi \rangle = \mathbf{1}$ , we immediately obtain

$$E_{\mu} = \langle \Phi | \mathscr{L}_{\mu} \bar{H}^{(A)} R_{\mu}^{(A)} | \Phi \rangle.$$
(3.25)

Inserting the resolution of identity in the N-electron Hilbert space,  $|\Phi\rangle\langle\Phi| + P + Q = \mathbf{1}$ , where  $P = \sum_{n=1}^{m_A} P_n$ ,  $Q = \sum_{n=m_A+1}^{N} P_n$ , and  $P_n = \sum_{\substack{i_1 < \cdots < i_n \\ a_1 < \cdots < a_n}} |\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle\langle\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}|$ , between

 $\mathscr{L}_{\mu}$  and  $\bar{H}^{(A)}$ , we arrive, after some manipulations, at the final formula for the exact energy

 $E_{\mu}$  that defines the biorthogonal MMCC theory,

$$E_{\mu} = E_{\mu}^{(A)} + \delta_{\mu}^{(A)}, \qquad (3.26)$$

where  $E_{\mu}^{(A)}$  is the energy of the  $\mu$ -th electronic state obtained with the CC/EOMCC method A and  $\delta_{\mu}^{(A)}$  is the correction toward full CI which takes on the following generic form:

$$\delta_{\mu}^{(A)} = \sum_{n=m_A+1}^{N_{\mu,A}} \sum_{\substack{i_1 < \dots < i_n \\ a_1 < \dots < a_n}} \ell_{\mu,i_1\dots i_n}^{a_1\dots a_n} \mathfrak{M}_{\mu,a_1\dots a_n}^{i_1\dots i_n}(m_A), \qquad (3.27)$$

where  $N_{\mu,A}$  is the highest value of *n* for which moments  $\mathfrak{M}_{\mu,a_1...a_n}^{i_1...i_n}(m_A)$ , Eq. (3.18) for  $\mu = 0$  or Eq. (3.19) for  $\mu > 0$ , are nonzero.

In practical calculations based on the above equations, the sum over n in Eq. (3.27) is truncated at some excitation level  $m_B$ , where  $m_A < m_B \leq N_{\mu,A}$ . For example, in the CR-CC(2,3) and CR-EOMCC(2,3) approaches introduced in Refs. [116, 117, 160–162] and further developed in Ref. [96], the CCSD and EOMCCSD energies,  $E_{\mu}^{(\text{CCSD})}$ , are corrected for the leading correlation effects due to triple excitations, so that  $m_A = 2$  and  $m_B = 3$ . As a result, the CR-CC(2,3) and CR-EOMCC(2,3) energy expressions are

$$E_{\mu}(2,3) = E_{\mu}^{(\text{CCSD})} + \delta_{\mu}(2,3), \qquad (3.28)$$

where

$$\delta_{\mu}(2,3) = \sum_{\substack{i < j < k \\ a < b < c}} \ell_{\mu,ijk}^{abc}(2) \mathfrak{M}_{\mu,abc}^{ijk}(2), \qquad (3.29)$$

with  $\mathfrak{M}_{\mu,abc}^{ijk}(2)$  representing the triply-excited moments of the CCSD ( $\mu = 0$ ) or EOMCCSD ( $\mu > 0$ ) equations. In order to make Eqs. (3.28) and (3.29) computationally manageable, one

has to come up with the approximate form of the  $\ell^{abc}_{\mu,ijk}(2)$  amplitudes that originate from the three-body component of the  $\mathscr{L}_{\mu}$  operator, Eq. (3.22) defining the exact bra state  $\langle \Psi_{\mu}|$  in Eq. (3.21), which in this particular case, where  $m_A = 2$ , becomes  $\langle \Psi_{\mu}| = \langle \Phi | \mathscr{L}_{\mu} e^{-T^{(\text{CCSD})}}$ , with  $T^{(\text{CCSD})} = T_1 + T_2$  representing the cluster operator obtained in CCSD calculations. This can be done in several ways. In the specific case of CR-CC(2,3) and CR-EOMCC(2,3) methods of Refs. [116, 117, 160, 161], the  $\ell^{abc}_{\mu,ijk}(2)$  that enter the corrections  $\delta_{\mu}(2,3)$ , Eq. (3.29), are calculated using the expression

$$\ell^{abc}_{\mu,ijk}(2) = \langle \Phi | L^{(\text{CCSD})}_{\mu} \bar{H}^{(\text{CCSD})} | \Phi^{abc}_{ijk} \rangle / D^{ijk}_{\mu,abc}(2)$$
(3.30)

where

$$D_{\mu,abc}^{ijk}(2) = E_{\mu}^{(\text{CCSD})} - \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} | \Phi_{ijk}^{abc} \rangle, \qquad (3.31)$$

which is obtained by performing an approximate quasi-perturbative analysis of the bra Schrödinger equation eigenvalue problem  $\langle \Psi_{\mu} | H = E_{\mu} \langle \Psi_{\mu} |$ , with  $\langle \Psi_{\mu} |$  given by  $\langle \Phi | \mathscr{L}_{\mu} e^{-T^{(\text{CCSD})}}$ , constrained to the subspace of the Hilbert space up to triple excitations (see Refs. [116, 117, 160, 161] for the details). The  $L_{\mu}^{(\text{CCSD})}$  operator in the above expression is the deexcitation operator defining the bra CCSD/EOMCCSD state  $\langle \tilde{\Psi}_{\mu}^{(\text{CCSD})} | = \langle \Phi | L_{\mu}^{(\text{CCSD})} e^{-T^{(\text{CCSD})}}$ , which matches the CCSD/EOMCCSD ket state  $|\Psi_{\mu}^{(\text{CCSD})}\rangle = R_{\mu}^{(\text{CCSD})}e^{T^{(\text{CCSD})}}|\Phi\rangle$  and which is defined as  $L_{\mu}^{(\text{CCSD})} = \delta_{\mu,0}\mathbf{1} + L_{\mu,1} + L_{\mu,2}$ , where  $L_{\mu,1}$  and  $L_{\mu,2}$  are the relevant oneand two-body components.  $\bar{H}^{(\text{CCSD})}$  in Eqs. (3.30) and (3.31) is the similarity-transformed Hamiltonian of CCSD,  $\bar{H}^{(\text{CCSD})} = e^{-T^{(\text{CCSD})}}He^{T^{(\text{CCSD})}} = (He^{T^{(\text{CCSD})}})_c$ . One obtains these components, or the deexcitation amplitudes  $l_{\mu,i}^a$  and  $l_{\mu,ij}^{ab}$  that define them, by solving the left-eigenstate CCSD/EOMCCSD equations. The general left-eigenstate CC/EOMCC system corresponding to truncation A at  $m_A$ -fold excitations has the form of a linear system

$$\delta_{\mu,0} \langle \Phi | \bar{H}_{\text{open}}^{(A)} | \Phi_{i_1 \dots i_n}^{a_1 \dots a_n} \rangle + \langle \Phi | L_{\mu,\text{open}}^{(A)} \bar{H}_{\text{open}}^{(A)} | \Phi_{i_1 \dots i_n}^{a_1 \dots a_n} \rangle$$
  
=  $\omega_{\mu}^{(A)} l_{\mu,i_1 \dots i_n}^{a_1 \dots i_n} (m_A) \quad i_1 < \dots < i_n, \quad a_1 < \dots < a_n \qquad n = 1, \dots, m_A \quad (3.32)$ 

where  $L_{\mu,\text{open}}^{(A)} = \sum_{n=1}^{m_a} L_{\mu,n}$  is the deexcitation operator defining the bra CC/EOMCC state  $\langle \tilde{\Psi}_{\mu}^{(A)} | = \langle \Phi | (\delta_{\mu,0} \mathbf{1} + L_{\mu,\text{open}}^{(A)}) e^{-T^{(A)}}$  matching the ket state  $|\Psi_{\mu}^{(A)}\rangle = R_{\mu}^{(A)} e^{T^{(A)}} |\Phi\rangle$ . The left-eigenstate CCSD/EOMCCSD equations are obtained by setting  $m_A$  in Eq. (3.32) at 2.

In performing the CR-CC(2,3) and CR-EOMCC(2,3) calculations, we often distinguish between variants A and D. Variant D is the full form of CR-CC(2,3)/CR-EOMCC(2,3), where we use the  $D_{\mu,abc}^{ijk}(2)$  denominator in its most complete Epstein–Nesbet form given by Eq. (3.31). Variant A is obtained by replacing Eq. (3.31) by the Møller–Plesset formula  $\omega_{\mu}^{(\text{CCSD})} - (\epsilon_i + \epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b - \epsilon_c)$ , where  $\omega_{\mu}^{(\text{CCSD})}$  is the EOMCCSD excitation energy (0 in the  $\mu = 0$  ground-state case) and  $\epsilon_p$ 's are the usual spin-orbital energies (diagonal elements of the Fock matrix). We will use similar variants in our CC(*P*;*Q*) considerations.

One can use similar ideas to develop other  $\operatorname{CR-CC}(m_A, m_B)/\operatorname{CR-EOMCC}(m_A, m_B)$  methods, such as  $\operatorname{CR-CC}(2,4)/\operatorname{CR-EOMCC}(2,4)$  and  $\operatorname{CR-CC}(3,4)/\operatorname{CR-EOMCC}(3,4)$ . In the  $\operatorname{CR-CC}(2,4)/\operatorname{CR-EOMCC}(2,4)$  case, where  $m_A = 2$  and  $m_B = 4$ , we correct the CCSD/EOM-CCSD energy for the combined effects of triples and quadruples using the formula

$$E_{\mu}(2,4) = E_{\mu}^{(\text{CCSD})} + \delta_{\mu}(2,4), \qquad (3.33)$$

where

$$\delta_{\mu}(2,4) = \sum_{\substack{i < j < k \\ a < b < c}} \ell^{abc}_{\mu,ijk}(2) \mathfrak{M}^{ijk}_{\mu,abc}(2) + \sum_{\substack{i < j < k < l \\ a < b < c < d}} \ell^{abcd}_{\mu,ijkl}(2) \mathfrak{M}^{ijkl}_{\mu,abcd}(2). \quad (3.34)$$

In the CR-CC(3,4)/CR-EOMCC(3,4) case, where  $m_A = 3$  and  $m_B = 4$ , we correct the CCSDT/EOMCCSDT energy for correlation effects of quadruples using

$$E_{\mu}(3,4) = E_{\mu}^{(\text{CCSDT})} + \delta_{\mu}(3,4), \qquad (3.35)$$

where

$$\delta_{\mu}(3,4) = \sum_{\substack{i < j < k < l \\ a < b < c < d}} \ell^{abcd}_{\mu,ijkl}(3) \mathfrak{M}^{ijkl}_{\mu,abcd}(3).$$
(3.36)

The  $\ell_{\mu,ijk}^{abc}(2)$ ,  $\ell_{\mu,ijkl}^{abcd}(2)$ , and  $\ell_{\mu,ijkl}^{abcd}(3)$  amplitudes entering the above expressions are determined in a similar way as in the CR-CC(2,3)/CR-EOMCC(2,3) methods using quasiperturbative expressions similar to Eqs. (3.30) and (3.31). Again, in each case, we typically distingush between variants A and D in handling the respective denominators analogous to  $D_{\mu,abc}^{ijk}(2)$  entering CR-CC(2,3) and CR-EOMCC(2,3), with variant D implying the full Epstein–Nesbet-like treatment and variant A implying the more approximate Møller–Plesset treatment.

All of the above CR-CC/CR-EOMCC approaches offer tremendous savings in the computer effort when compared to their iterative parents. For example, CR-CC(2,3) reduces the iterative  $n_o^3 n_u^5$  CPU steps of full CCSDT by the iterative  $n_o^2 n_u^4$  steps of CCSD and noniterative  $n_o^3 n_u^4$  steps associated with the determination of the  $\delta_{\mu}(2,3)$  corrections. Similarly, CR-CC(2,4) replaces the iterative  $n_o^4 n_u^6$  steps of CCSDTQ by the iterative  $n_o^2 n_u^4$  steps of CCSD and non-iterative  $n_o^3 n_u^4$  and  $n_o^2 n_u^5$  (variant A) or  $n_o^4 n_u^5$  (variant D) steps. CR-CC(3,4) replaces the iterative  $n_o^4 n_u^6$  steps of CCSDTQ by the iterative  $n_o^3 n_u^5$  steps of CCSDT and non-iterative  $n_o^2 n_u^5$  (variant A) or  $n_o^4 n_u^5$  (variant D) steps. All of these cost reductions offered by the various CR-CC( $m_A, m_B$ )/CR-EOMCC( $m_A, m_B$ ) methods with  $m_A = 2$  or 3 and  $m_B = 3$  or 4 are summarized in Table 1.1.

The non-iterative CR-CC( $m_a, m_b$ ) approaches and their EOMCC extensions are often as accurate as the parent iterative methods at the fraction of the cost. For example, the CR-CC(2,3) approach recovers the full CCSDT results for potential energy surfaces describing single bond breaking to within a millihartree or so, eliminating failures of perturbative methods such as CCSD(T). Unfortunately, as explained in the Introduction, there are cases where CR-CC(2,3) and other CR-CC( $m_A, m_B$ ) methods have problems. This happens when the relaxation of the cluster and excitation amplitudes corresponding to the underlying method A we are trying to correct becomes significant in the presence of higher many-body components of T and  $R_{\mu}$  operators. In all such cases it is useful to turn to the active-space CC and EOMCC approaches, which provide better cluster and excitation amplitudes and which we discuss next.

## 3.1.2 Active-Space Coupled-Cluster Methods

As discussed in the Introduction, if we are to handle systems involving stronger non-dynamical correlations in a robust manner by recovering such correlations through the explicit inclusion of higher-than-doubly excited operators defining T and  $R_{\mu}$  in the SRCC framework, the biggest issue we face is that of the prohibitive computer costs of the CCS-DT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, and other high-level SRCC approaches. The CR-CC/CR-EOMCC corrections described in Section 3.1.1 are one way to address it, but one cannot solve every MR problem in this way, so it is worth exploring alternatives, such as the active-space CC and EOMCC methods.

The active-space CC and EOMCC approximations are a class of robust methods that seek to capture stronger non-dynamical correlation effects through suitable selection of the dominant  $T_3$ ,  $T_4$ ,  $R_{\mu,3}$ ,  $R_{\mu,4}$ , and other higher-than-doubly excited components of T and  $R_{\mu}$ , while retaining relatively inexpensive computational costs. This is accomplished with the help of active orbitals that are used to identify dominant  $T_3$ ,  $T_4$ ,  $R_{\mu,3}$ ,  $R_{\mu,4}$ , etc. amplitudes. Each active-space CC /EOMCC calculation begins by partitioning the one-electron basis of occupied and unoccupied spin-orbitals used in the conventional SRCC considerations into four disjoint groups of core or inactive occupied spin-orbitals, labeled by lower-case bold letters  $\mathbf{i}_1, \mathbf{i}_2, \ldots$ , or  $\mathbf{i}, \mathbf{j}, \ldots$ , active spin-orbitals occupied in the reference determinant  $|\Phi\rangle$ , labeled by upper-case bold letters  $I_1, I_2, \ldots$ , or  $I, J, \ldots$ , active spin-orbitals unoccupied in  $|\Phi\rangle$ , labeled by upper-case bold letters  $A_1, A_2, \ldots$ , or  $A, B, \ldots$ , and virtual or inactive unoccupied spin-orbitals, labeled by lower-case bold letters  $\mathbf{a}_1, \mathbf{a}_2, \ldots$ , or  $\mathbf{a}, \mathbf{b}, \ldots$  (see Fig. 3.1). In the following, we continue to use symbols set in italics for occupied spinorbitals  $(i_1, i_2, \ldots, \text{ or } i, j, \ldots)$  and unoccupied spin-orbitals  $(a_1, a_2, \ldots, \text{ or } a, b, \ldots)$ if the active/inactive character of a given spin-orbital is not specified. Typically, in the active-space CC/EOMCC methods, the low-order cluster and excitation operators, such as  $T_1$ ,  $T_2$ ,  $R_{\mu,1}$ , and  $R_{\mu,2}$ , are treated exactly, i.e., all available spin-orbitals are used to define them, as in Eqs. (3.2) and (3.6), but the higher-order operators, such as  $T_3$ ,  $T_4$ ,  $R_{\mu,3}$ , and  $R_{\mu,4}$ , are treated in an approximate manner, by selecting their dominant parts through active spin-orbitals, which represent a relatively small subset of all spin-orbitals, thus significantly reducing the computer costs. This way of thinking can be extended to any SRCC or EOMCC [107–109, 119, 125, 163–165, 238–253] (even, MRCC [64, 66]) approach. In general, if  $T_{\sigma}$  and  $R_{\mu,\sigma}$  are the highest-order many-body components of T and  $R_{\mu}$  considered, one can treat the  $T_n$  and  $R_{\mu,n}$  components with  $n \leq \rho \ (\rho < \sigma)$  exactly, as in Eqs. (3.2) and



Figure 3.1: The orbital classification used in the active-space SRCC methods, such as CCSDt, CCSDtq, and CCSDTq. Core, active, and virtual orbitals are represented by solid, dashed, and dotted lines, respectively. Full and open circles represent core and active electrons of the reference determinant  $|\Phi\rangle$  (a closed-shell reference  $|\Phi\rangle$  is assumed).

(3.6), while using active spin-orbitals to define the remaining higher-order  $T_n$  and  $R_{\mu,n}$ many-body components with  $\rho < n \leq \sigma$ . The freedom to choose the highest-order operators  $T_{\sigma}$  and  $R_{\mu,\sigma}$  and which operators to treat exactly and which approximately using active orbitals leads to a hierarchy of the active-space CC/EOMCC methods.

The simplest of these methods that explicitly incorporates higher-order cluster and excitation components beyond those defining CCSD and EOMCCSD are the CC approaches with singles, doubles, and active-space triples (CCSDt/EOMCCSDt), often termed "little t". In this case, we solve the SRCC amplitude equations, Eq. (3.11), in which  $T^{(A)}$  is replaced by

$$T^{(\text{CCSDt})} = T_1 + T_2 + t_3 \tag{3.37}$$

and  $\bar{H}^{(A)}$  by

$$\bar{H}^{(\text{CCSDt})} = e^{-T^{(\text{CCSDt})}} H e^{T^{(\text{CCSDt})}} = \left(H e^{T^{(\text{CCSDt})}}\right)_C,$$
(3.38)

where

$$T_1 = \sum_{i,a} t_a^i E_i^a, (3.39)$$

$$T_2 = \sum_{i>j,a>b} t_{ab}^{ij} E_{ij}^{ab},$$
(3.40)

and

$$t_3 = \sum_{\mathbf{I} > j > k, a > b > \mathbf{C}} t_{ab\mathbf{C}}^{\mathbf{I}jk} E_{\mathbf{I}jk}^{ab\mathbf{C}}.$$
(3.41)

In other words, we solve the CCSDT equations using all singles, all doubles, and a subset of triples defined by Eq. (3.41). The  $t_a^i$ ,  $t_{ab}^{ij}$ , and  $t_{ab}^{\mathbf{I}jk}$  amplitudes are obtained by solving the SRCC system, Eq. (3.11), in which  $T^{(A)} = T^{(\text{CCSDt})}$ , in a subspace of the many-electron Hilbert space,  $\mathscr{H}$ , spanned by the singly excited determinants  $|\Phi_i^a\rangle = E_i^a |\Phi\rangle$ , doubly excited determinants  $|\Phi_{ij}^{ab}\rangle = E_{ij}^{ab}|\Phi\rangle$ , and selected triply excited determinants  $|\Phi_{ijk}^{ab}\rangle = E_{ij}^{ab}|\Phi\rangle$ , and selected triply excited determinants  $|\Phi_{ijk}^{ab}\rangle = E_{ijk}^{ab}|\Phi\rangle$  that correspond to the excitations in  $T^{(\text{CCSDt})}$ , referred to as  $\mathscr{H}^{(t)}$ . In terms of core, active, and virtual orbitals described above and shown in Fig. 3.1, this is equivalent to projecting the connected cluster form of the Schrödinger equation with  $T = T^{(\text{CCSDt})}$  on all singly, all doubly, and selected triply excited determinants that belong to classes 1–9 on Table 3.1.

In the case of excited states, the corresponding EOMCCSDt analog of CCSDt is defined by replacing the  $R_{\mu}^{(A)}$  operator entering the excited-state wave function ansatz, Eq. (3.5), by

$$R_{\mu}^{(\text{CCSDt})} = r_{\mu,0}\mathbf{1} + R_{\mu,1} + R_{\mu,2} + r_{\mu,3}, \qquad (3.42)$$

where

$$R_{\mu,1} = \sum_{i,a} r^i_{\mu,a} E^a_i, \qquad (3.43)$$

$$R_{\mu,2} = \sum_{i>j,a>b} r^{ij}_{\mu,ab} E^{ab}_{ij}, \qquad (3.44)$$

and

$$r_{\mu,3} = \sum_{\mathbf{I}>j>k,a>b>\mathbf{C}} r_{\mu,ab\mathbf{C}}^{\mathbf{I}jk} E_{\mathbf{I}jk}^{ab\mathbf{C}}.$$
(3.45)

The  $r_{\mu,a}^{i}$ ,  $r_{\mu,ab}^{ij}$ , and  $r_{\mu,ab\mathbf{C}}^{\mathbf{I}jk}$  amplitudes and the corresponding vertical excitation energies  $\omega_{\mu}^{(\text{CCSDt})}$  are determined by diagonalizing the similarity-transformed Hamiltonian of CCSDt,  $\bar{H}^{(\text{CCSDt})}$ , defined by Eq. (3.38), in the space spanned by  $|\Phi_{i}^{a}\rangle$ ,  $|\Phi_{ij}^{ab}\rangle$ , and  $|\Phi_{\mathbf{I}jk}^{ab\mathbf{C}}\rangle$  determinants corresponding to classes 1–9 in Table 3.1.

The main benefit of restricting the higher-order  $T_3$  and  $R_{\mu,3}$  components to their activespace  $t_3$  and  $r_{\mu,3}$  variants is the significant saving in computational costs compared to the parent full CCSDT and EOMCCSDT approaches. If  $n_o$  and  $n_u$  are the numbers of occupied and unoccupied orbitals, respectively, used in the correlated calculation, and if  $N_o$  (<  $n_o$  or  $\ll n_o$  and  $N_u$  ( $\ll n_u$ ) designate the corresponding numbers of the active occupied and active unoccupied orbitals, the CCSDt and EOMCCSDt approaches replace the iterative  $n_o^3 n_u^5$  CPU steps of full CCSDT/EOMCCSDT by the much less expensive  $N_o N_u n_o^2 n_u^4$  steps (see Table 1.1). In other words, when properly implemented, the CCSDt and EOMCCSDt methods replace the computationally expensive  $\mathcal{N}^8$  iterative CPU steps of CCSDT/EOMCCSDT, by the relatively inexpensive steps equivalent to the costs of CCSD or EOMCCSD calculations multiplied by a small prefactor equal to the number of single excitations in the active space. The low-order polynomial growth of the CPU time represented by the  $N_o N_u$  prefactor also means that it is trivial to improve the results or check if they are converged by increasing the active space. This should be contrasted with the genuine MR methods, which are usually based on a model space defined by the CAS approach, which result in dimensionalities that scale factorially with the numbers of active orbitals and electrons, in addition to other costs of post-CASSCF calculations. Similar applies to memory requirements, which reduces

Table 3.1: The complete set of projections entering the CCSDTQ amplitude equations organized according to their active and inactive character. Upper-case bold letters represent active labels and lower-case bold letters represent inactive labels (cf. the text and Fig. 3.1 for details).

Projection	Single	Double	Triple	Quadruple
1	$\langle \Phi_{\mathbf{I}}^{\mathbf{A}}  $	$\langle \Phi_{{f I}{f J}}^{{f A}{f B}} $	$\langle \Phi^{\mathbf{ABC}}_{\mathbf{IJK}} $	$\langle \Phi^{\mathbf{ABCD}}_{\mathbf{IJKL}}  $
2	$\langle \Phi_{\mathbf{I}}^{\mathbf{a}}  $	$\langle \Phi_{\mathbf{IJ}}^{\mathbf{aB}} $	$\langle \Phi^{\mathbf{aBC}}_{\mathbf{IJK}}  $	$\langle \Phi^{\mathbf{aBCD}}_{\mathbf{IJKL}}  $
3	$\langle \Phi_{\mathbf{i}}^{\mathbf{A}}  $	$\langle \Phi_{\mathbf{I}\mathbf{i}}^{\mathbf{\overline{AB}}} $	$\langle \Phi_{\mathbf{IJk}}^{\mathbf{ABC}}  $	$\langle \Phi_{\mathbf{IJKl}}^{\mathbf{ABCD}}  $
4	$\langle \Phi_{\mathbf{i}}^{\mathbf{a}}  $	$\langle \Phi_{\mathbf{I}\mathbf{i}}^{\mathbf{a}\mathbf{B}} $	$\langle \Phi^{aBC}_{IJk}  $	$\langle \Phi^{\mathbf{aBCD}}_{\mathbf{IJKl}}  $
5		$\langle \Phi_{\mathbf{IJ}}^{\mathbf{ab}} $	$\langle \Phi_{\mathbf{IJK}}^{\mathbf{abC}}  $	$\langle \Phi^{abCD}_{IJKL}  $
6		$\langle \Phi_{\mathbf{I}\mathbf{i}}^{\mathbf{ab}} $	$\langle \Phi_{\mathbf{IJk}}^{\mathbf{abC}}  $	$\langle \Phi^{abCD}_{IJKl}  $
7		$\langle \Phi_{\mathbf{i}\mathbf{i}}^{\mathbf{\overline{A}B}}  $	$\langle \Phi_{\mathbf{Iik}}^{\mathbf{ABC}}  $	$\langle \Phi_{\mathbf{IJkl}}^{\mathbf{ABCD}}  $
8		$\langle \Phi_{\mathbf{i}\mathbf{j}}^{\mathbf{a}\mathbf{B}}  $	$\langle \Phi_{\mathbf{Iik}}^{\mathbf{aBC}}  $	$\langle \Phi^{\mathbf{aBCD}}_{\mathbf{IJkl}}  $
9		$\langle \Phi_{\mathbf{i}\mathbf{i}}^{\mathbf{a}\mathbf{b}} $	$\langle \Phi^{f abC}_{Iik}  $	$\langle \Phi_{\mathbf{I}\mathbf{J}\mathbf{k}\mathbf{l}}^{\mathbf{ab}\mathbf{C}\mathbf{D}} $
10		-0	$\langle \Phi^{abc}_{IJK}  $	$\langle \Phi^{\mathbf{abcD}}_{\mathbf{IJKL}}  $
11			$\langle \Phi^{abc}_{IJk}  $	$\langle \Phi^{\mathbf{abcD}}_{\mathbf{IJKl}}  $
12			$\langle \Phi^{\mathbf{abc}}_{\mathbf{Iik}}  $	$\langle \Phi^{\mathbf{abcD}}_{\mathbf{IJkl}}  $
13			$\langle \Phi^{f ABC}_{f ijk} $	$\langle \Phi_{\mathbf{Iikl}}^{\mathbf{ABCD}}  $
14			$\langle \Phi_{\mathbf{ijk}}^{\mathbf{aBC}}  $	$\langle \Phi_{\mathbf{Ijkl}}^{\mathbf{aBCD}}  $
15			$\langle \Phi^{\mathbf{abC}}_{\mathbf{ijk}}  $	$\langle \Phi_{\mathbf{Ijkl}}^{\mathbf{abCD}}  $
16			$\langle \Phi^{\mathbf{abc}}_{\mathbf{iik}}  $	$\langle \Phi_{f Iikl}^{f abc D}  $
17			5	$\langle \Phi^{f abcd}_{f IJKL}  $
18				$\langle \Phi^{\mathbf{abcd}}_{\mathbf{IJKl}}  $
19				$\langle \Phi^{\mathbf{abcd}}_{\mathbf{IJkl}}  $
20				$\langle \Phi^{\mathbf{abcd}}_{\mathbf{Iikl}}  $
21				$\langle \Phi_{\mathbf{i}\mathbf{i}\mathbf{k}\mathbf{l}}^{\mathbf{\overline{ABCD}}}  $
22				$\langle \Phi_{\mathbf{ijkl}}^{\mathbf{aBCD}}  $
23				$\langle \Phi_{f ijkl}^{f abCD}  $
24				$\langle \Phi_{\mathbf{iikl}}^{\mathbf{abcD}}  $
25				$\langle \Phi_{\mathbf{ijkl}}^{\mathbf{\check{a}bcd}}  $

from the  $\sim n_o^3 n_u^3$  requirements of CCSDT/EOMCCSDT, related to the need to store all triply excited amplitudes, to the much more manageable  $\sim N_o N_u n_o^2 n_u^2$  requirement of CCS-Dt/EOMCCSDt. While the  $t_3$  and  $r_{\mu,3}$  operators presented above require a minimum of one occupied and one unoccupied index to be restricted to the active orbitals, one can always consider alternative definitions, where more spin-orbital indices are from the active set. One can, for example, consider the so-called CCSDt(II) [118, 163, 275] and CCSDt(III) [118, 163, 275]

or CCSDt' [242, 243] approaches and their excited-state analogs (as a note, the CCSDt approximation presented above is alternatively referred to as CCSDt(I) [118, 163, 275]). The savings that such simplified active-space methods provide are even greater than those offered by the above CCSDt/EOMCCSDt model.

The flexibility in the choice of the level of truncation in T and  $R_{\mu}$  and which operators to treat exactly and which to treat approximately allows us to extend the active-space considerations beyond CCSDt/EOMCCSDt to include other high-order correlations, such as quadruples, which are of particular interest in this work. In the CC/EOMCC approaches with singles, doubles, and active-space triples and quadruples (CCSDtq/EOMCCSDtq), termed "little tq", the truncated cluster operator defining the ground-state wave function is defined as follows:

$$T^{(\text{CCSDtq})} = T_1 + T_2 + t_3 + t_4, \tag{3.46}$$

and  $\bar{H}^{(A)}$  entering the SRCC amplitude equations, Eq. (3.11), is replaced by

$$\bar{H}^{(\text{CCSDtq})} = e^{-T^{(\text{CCSDtq})}} H e^{T^{(\text{CCSDtq})}} = \left(H e^{T^{(\text{CCSDtq})}}\right)_C, \qquad (3.47)$$

where  $T_1$ ,  $T_2$ , and  $t_3$  are defined by Eqs. (3.39)–(3.41) and

$$t_4 = \sum_{\mathbf{I} > \mathbf{J} > k > l, a > b > \mathbf{C} > \mathbf{D}} t_{ab\mathbf{C}\mathbf{D}}^{\mathbf{I}\mathbf{J}kl} E_{\mathbf{I}\mathbf{J}kl}^{ab\mathbf{C}\mathbf{D}}.$$
(3.48)

The cluster amplitudes are determined by solving Eq. (3.11), in which  $T^{(A)}$  is replaced by  $T^{(\text{CCSDtq})}$  in the space spanned by the  $|\Phi_i^a\rangle$ ,  $|\Phi_{ij}^{ab}\rangle$ ,  $|\Phi_{\mathbf{I}jk}^{ab}\mathbf{C}\rangle$  and  $|\Phi_{\mathbf{I}Jkl}^{ab}\mathbf{CD}\rangle$  determinants, using the projections 1–9 in Table 3.1. In the case of the excited-state EOMCCSDtq analog of CCSDtq, we replace the  $R_{\mu}^{(A)}$  operator entering the EOMCC wave function ansatz, Eq. (3.5),

by

$$R_{\mu}^{(\text{CCSDtq})} = r_{\mu,0}\mathbf{1} + R_{\mu,1} + R_{\mu,2} + r_{\mu,3} + r_{\mu,4}, \qquad (3.49)$$

where  $R_{\mu,1}$ ,  $R_{\mu,2}$ , and  $r_{\mu,3}$  are defined by Eqs. (3.43)–(3.45) and

$$r_{\mu,4} = \sum_{\mathbf{I} > \mathbf{J} > k > l, a > b > \mathbf{C} > \mathbf{D}} r_{\mu,ab\mathbf{C}\mathbf{D}}^{\mathbf{I}\mathbf{J}kl} E_{\mathbf{I}\mathbf{J}kl}^{ab\mathbf{C}\mathbf{D}}.$$
(3.50)

The excitation amplitudes defining  $R_{\mu}^{(\text{CCSDtq})}$  and the corresponding vertical excitation energies  $\omega_{\mu}^{(\text{CCSDtq})}$  are determined by diagonalizing the similarity-transformed Hamiltonian of CCSDtq,  $\bar{H}^{(\text{CCSDtq})}$ , given by Eq. (3.47), in the space spanned by the  $|\Phi_{i}^{a}\rangle$ ,  $|\Phi_{ij}^{ab}\rangle$ ,  $|\Phi_{\mathbf{I}jk}^{ab}\rangle$ and  $|\Phi_{\mathbf{IJ}kl}^{ab}\mathcal{CD}\rangle$  determinants, using projections 1–9 in Table 3.1.

We can also treat the tri-excited components of T and  $R_{\mu}$  exactly, while approximating the quadruples components,  $T_4$  and  $R_{\mu,4}$ , approximately using active orbitals, which is how the CCSDTq and EOMCCSDTq approaches are defined. In the CCSDTq/EOMCCSDTq methodology, the truncated cluster operator defining the ground-state wave function is given by

$$T^{(\text{CCSDTq})} = T_1 + T_2 + T_3 + t_4, \tag{3.51}$$

where, once again,  $T_1$  and  $T_2$  are defined by Eqs. (3.39) and (3.40), and  $T_3$  is treated fully as well,

$$T_3 = \sum_{i>j>k, a>b>c} t^{ijk}_{abc} E^{abc}_{ijk}, \qquad (3.52)$$

but we use active orbitals to define the four-body cluster components,

$$t_4 = \sum_{\mathbf{I} > j > k > l, a > b > c > \mathbf{D}} t_{abc\mathbf{D}}^{\mathbf{I}jkl} E_{\mathbf{I}jkl}^{abc\mathbf{D}}.$$
(3.53)

We then solve the SRCC amplitude equations, Eq. (3.11), in which  $\bar{H}^{(A)}$  is replaced by

$$\bar{H}^{(\text{CCSDTq})} = e^{-T^{(\text{CCSDTq})}} H e^{T^{(\text{CCSDTq})}} = \left(H e^{T^{(\text{CCSDTq})}}\right)_C, \quad (3.54)$$

in the space spanned by the  $|\Phi_i^a\rangle$ ,  $|\Phi_{ij}^{ab}\rangle$ ,  $|\Phi_{ijk}^{abc}\rangle$  and  $|\Phi_{\mathbf{I}jkl}^{abc}\mathsf{D}\rangle$  determinants, using projections 1–16 in Table 3.1, and if we are interested in calculating excited states using EOMCCSDTq, we replace the  $R_{\mu}^{(A)}$  operator entering the EOMCC wave function ansatz, Eq. (3.5), by

$$R_{\mu}^{(\text{CCSDTq})} = r_{\mu,0}\mathbf{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3} + r_{\mu,4}, \qquad (3.55)$$

where  $R_{\mu,1}$  and  $R_{\mu,2}$  are defined by Eqs. (3.43) and (3.44),

$$R_{\mu,3} = \sum_{i>i>k,a>b>c} r_{\mu,abc}^{ijk} E_{ijk}^{abc},$$
(3.56)

and

$$r_{\mu,4} = \sum_{\mathbf{I} > j > k > l, a > b > c > \mathbf{D}} r_{\mu,abc\mathbf{D}}^{\mathbf{I}jkl} E_{\mathbf{I}jkl}^{abc\mathbf{D}}.$$
(3.57)

The excitation amplitudes of EOMCCSDTq and the corresponding vertical excitation energies  $\omega_{\mu}^{(\text{CCSDTq})}$  are determined by diagonalizing the similarity-transformed Hamiltonian of CCSDTq,  $\bar{H}^{(\text{CCSDTq})}$ , defined by Eq. (3.54), in the space spanned by the  $|\Phi_i^a\rangle$ ,  $|\Phi_{ijk}^{ab}\rangle$ and  $|\Phi_{\mathbf{I}jkl}^{abc\mathbf{D}}\rangle$  determinants, using projections 1–16 in Table 3.1.

As shown in Table 1.1, the computational savings offered by the active-space CCSDtq/ EOMCCSDtq and CCSDTq/EOMCCSDTq approaches can be huge compared to the parent CCSDTQ and EOMCCSDTQ methods. The most expensive CPU steps of the full CCSDtq approach, as described above, scale as  $N_o^2 N_u^2 n_o^2 n_u^4$ , i.e., as a small prefactor proportional to the number of double excitations within the active space times the  $n_o^2 n_u^4$  steps of CCSD/EOMCCSD. Since  $N_o < n_o$  (or  $\ll n_o$ ) and  $N_u \ll n_u$ , we obtain a significant saving in computer effort compared to the  $n_o^4 n_u^6$  steps of CCSDTQ/EOMCCSDTQ. Similar applies to the storage requirements for the triply and quadruply excited amplitudes, which are  $\sim n_o^3 n_u^3$  and  $\sim n_o^4 n_u^4$ , respectively, in the CCSDTQ/EOMCCSDTQ case, but only  $\sim N_o N_u n_o^2 n_u^2$  and  $\sim N_o^2 N_u^2 n_o^2 n_u^2$  in the case of CCSDtq/EOMCCSDtq. The CCSDTq and EOMCCSDTQ approaches presented above also reduce the computational effort compared to CCSDTQ/EOMCCSDTQ, with CPU steps that scale as  $N_o N_u n_o^3 n_u^5$ , i.e., the number of single excitations in active space times the  $n_o^3 n_u^5$  steps of the CCSDT type, and storage requirements for quadruples that are  $\sim N_o N_u n_o^3 n_u^3$ . In designing the best and most complete CCSDtq/EOMCCSDtq and CCSDTq/EOCCSDTq approaches, as described above, one follows the general recipe elaborated on below, but one can always reduce the costs of the CCSDtq/EOMCCSDtq and CCSDTq/EOMCCSDTq calculations further by restricting more occupied and unoccupied indices of the higher-order cluster and excitation operators to the active space, much like the case of the different variants of CCSDt. In fact, while the quadruple excitations corresponding to CCSDTq presented above are defined by a minimum of one active occupied and one active unoccupied indices, the initial implementations of CCSDTq [243, 276] employed a  $t_4$  operator defined by a minimum of two active occupied and two active unoccupied indices. Once again, and unlike the genuine MR methods, the low-order polynomial growth represented by the prefactors of CCSDtq and CCSDTq means it is rather easy to improve our description of a system of interest by increasing the active space.

The success of active-space SRCC methods in applications involving quasi-degenerate situations, such as bond-breaking, is rationalized by recognizing that the active-space SRCC

framework mimics a genuine MRCC description. In general, as shown in Ref. [240], the active-space SRCC theory corresponds to the MRCC formalism in which we consider up to  $\rho$ -tuple excitations from the multi-dimensional model space

$$\mathcal{M}_{0} = \operatorname{span}\left\{ |\Phi\rangle, \left|\Phi_{\mathbf{I}_{1}}^{\mathbf{A}_{1}}\right\rangle, \left|\Phi_{\mathbf{I}_{1}\mathbf{I}_{2}}^{\mathbf{A}_{1}\mathbf{A}_{2}}\right\rangle, \dots, \left|\Phi_{\mathbf{I}_{1}\cdots\mathbf{I}_{\tau}}^{\mathbf{A}_{1}\cdots\mathbf{A}_{\tau}}\right\rangle \right\}_{\mathbf{I}_{1} > \mathbf{I}_{2} > \cdots > \mathbf{I}_{\tau}, \mathbf{A}_{1} > \mathbf{A}_{2} > \cdots > \mathbf{A}_{\tau}}$$
(3.58)

spanned by all possible at most  $\tau$ -tuple excitations from the reference determinant  $|\Phi\rangle$ . In that case the cluster operator T of the SRCC theory should have the form

$$T = T_1 + \dots + T_{\rho} + t_{\rho+1} + \dots + t_{\rho+\tau}, \qquad (3.59)$$

where  $T_1, \ldots, T_\rho$  are treated fully, to capture all dynamical correlations outside  $\mathcal{M}_0$ , and

$$t_{\rho+\kappa} = \sum_{\substack{\mathbf{I}_1 > \dots > \mathbf{I}_{\kappa} > i_1 > \dots > i_{\rho} \\ a_1 > \dots > a_{\rho} > \mathbf{A}_1 > \dots > \mathbf{A}_{\kappa}} t_{a_1 \dots a_{\rho} \mathbf{A}_1 \dots \mathbf{A}_{\kappa}}^{\mathbf{I}_1 \dots \mathbf{I}_{\kappa} i_1 \dots a_{\rho} \mathbf{A}_1 \dots \mathbf{A}_{\kappa}} E_{\mathbf{I}_1 \dots \mathbf{I}_{\kappa} i_1 \dots i_{\rho}}^{a_1 \dots a_{\rho} \mathbf{A}_1 \dots \mathbf{A}_{\kappa}}$$
(3.60)

for  $\kappa = 1, ..., \tau$ . In the full CCSDt method,  $\rho = 2$  and  $\tau = 1$ , meaning that CCSDt is analogous to MRCC methods in which we consider all single and double excitations from the model space spanned by the reference determinant  $|\Phi\rangle$  and the singly excited determinants in the active space (plus all their products). In the full CCSDtq approach,  $\rho = 2$  and  $\tau =$ 2, corresponding to all single and double excitations from the model space spanned by the reference determinant  $|\Phi\rangle$  and the singly and doubly excited determinants relative to  $|\Phi\rangle$ in the active space, while in the full CCSDTq approach presented above  $\rho = 3$  and  $\tau = 1$ , corresponding to all single, double, and triple excitations from the model space spanned by the reference determinant  $|\Phi\rangle$  and the singly excited determinants in the active space (plus the reference determinant  $|\Phi\rangle$  and the singly excited determinants in the active space (plus the reference determinant  $|\Phi\rangle$  and the singly excited determinants in the active space (plus their various products). While this is not a focus of this work, it is important to mention that the above generalized treatment of the cluster operator to mimic MRCC also allows us to consider active-space analogs of the SRCC methods with higher-than-quadruple excitations.

We have now reviewed the active-space CC/EOMCC approaches and the MMCC-based CR-CC/CR-EOMCC corrections, and we also know from the discussion in the Introduction that both of these methodologies face challenges in some MR situations, such as potential energy surfaces involving certain types of biradical transition states, so our next task is to show how to merge the CR-CC/CR-EOMCC and active-space CC/EOMCC methodologies into a single formalism which, as shown in this thesis, enables one to calculate potential energy surfaces along bond breaking coordinates and singlet-triplet gaps in challenging biradical species to within fractions of a millihartree relative to the high-level CC (e.g., CCSDTQ) or exact (full CI) data at the small fraction of the computer cost.

## 3.1.3 The CC(P;Q) Formalism

The overall purpose of the CC(P;Q) theory is to correct the results of CC/EOMCC calculations in a subspace of the N-electron Hilbert space, referred to as the P-space, for the electron correlation effects in the subspace called the Q-space. For example, if the P-space is spanned by all singly and doubly excited determinants, and the Q-space by the triply excited determinants, then we end up with the previously discussed CR-CC(2,3) approach. If the P-space is spanned by all singles, doubles, and triples, and the Q-space by all quadruples, we obtain CR-CC(3,4), etc. However, by being more flexible in the definitions of the P- and Q-spaces in the CC(P;Q) theory, we can correct the results of CC or EOMCC calculations using non-traditional truncations in the cluster and excitation operators for the electron correlation effects missing in such non-traditional calculations. An example of the the CC(P;Q) method relevant to this work is the CC(t;3) approach, which we discuss below, where the P-space is spanned by singles, doubles, and a subset of triples, as defined by the active-space CCSDt approach, and the Q-space is spanned by the remaining triples missing in CCSDt but present in CCSDT. Or, we could correct the active-space CCSDtq energy for the missing triples and quadruples, as in the CC(t,q;3,4) approach, also discussed below. In order to introduce these methods, we have to define the underlying mathematical concepts. We begin with the formal definition of the P-space CC/EOMCC calculations.

Let us designate a subspace of the N-electron Hilbert space of interest, referred to as the P-space, denoted  $\mathscr{H}^{(P)}$ , which is spanned by the excited determinants  $|\Phi_K\rangle = E_K |\Phi\rangle$ , where  $E_K$  is the elementary excitation operator which generates  $|\Phi_K\rangle$  from the reference determinant  $|\Phi\rangle$ . The corresponding truncated definitions of the T and  $R_{\mu}$  operators for the P-space CC/EOMCC calculations, designated by  $T^{(P)}$  and  $R^{(P)}_{\mu}$ , respectively, can be written as

$$T^{(P)} = \sum_{|\Phi_K\rangle \in \mathscr{H}^{(P)}} t_K E_K \tag{3.61}$$

and

$$R_{\mu}^{(P)} = R_{\mu,0}^{(P)} + R_{\mu,\text{open}}^{(P)} = r_{\mu,0} \mathbf{1} + \sum_{|\Phi_K\rangle \in \mathscr{H}^{(P)}} r_{\mu,K} E_K, \qquad (3.62)$$

where  $t_K$  and  $r_{\mu,K}$  are the cluster and excitation amplitudes obtained in the CC/EOMCC calculations in the  $\mathscr{H}^{(P)}$  subspace. In analogy to the conventional SRCC formalism, the  $t_K$  amplitudes defining  $T^{(P)}$  are determined by solving the system of non-linear, energyindependent equations,

$$\langle \Phi_K | \bar{H}_{\text{open}}^{(P)} | \Phi \rangle = 0, \ | \Phi_K \rangle \in \mathscr{H}^{(P)},$$
(3.63)

obtained by projecting the electronic Schrödinger equation on the determinants  $|\Phi_K\rangle$  belonging to the *P*-space  $\mathscr{H}^{(P)}$ . The corresponding ground-state energy is computed using the following expression:

$$E_0^{(P)} = \langle \Phi | \bar{H}^{(P)} | \Phi \rangle, \qquad (3.64)$$

where

$$\bar{H}^{(P)} = e^{-T^{(P)}} H e^{T^{(P)}} = (H e^{T^{(P)}})_C$$
(3.65)

is the relevant similarity-transformed Hamiltonian and  $\bar{H}_{open}^{(P)} = \bar{H}^{(P)} - \bar{H}_{closed}^{(P)} = \bar{H}^{(P)} - E_0^{(P)} \mathbf{1}$  is the 'open' part of  $\bar{H}^{(P)}$  defined using diagrams of  $\bar{H}^{(P)}$  that have external Fermion lines. We obtain the excited-state information, particularly, the amplitudes  $r_{\mu,K}$  defining  $R_{\mu}^{(P)}$  and the corresponding excited-state energies  $E_{\mu}^{(P)}$ , by diagonalizing the similaritytransformed Hamiltonian  $\bar{H}^{(P)}$  in  $\mathscr{H}^{(P)}$  and solving the corresponding *P*-space EOMCC equations:

$$\langle \Phi_K | (\bar{H}_{\text{open}}^{(P)} R_{\mu,\text{open}}^{(P)})_C | \Phi \rangle = \omega_\mu^{(P)} r_{\mu,K}, \quad | \Phi_K \rangle \in \mathscr{H}^{(P)}, \tag{3.66}$$

where

$$\omega_{\mu}^{(P)} = E_{\mu}^{(P)} - E_0^{(P)}.$$
(3.67)

The remaining  $r_{\mu,0}$  zero-body coefficient is calculated using the *P*-space analog of Eq. (3.16). The above set of equations encompasses various kinds of CC and EOMCC methods, including standard and nonstandard truncation schemes of the *T* and  $R_{\mu}$  operators.

If we stopped at this point, we would not know how important the correlation effects outside the *P*-space are. Thus, the only way to check the convergence would be to increase the *P*-space by adding more determinants to it, which can be very expensive. The CC(P;Q) theory offers a better and more economical solution. Thus, once the above *P*space CC/EOMCC equations have been solved and the information about the truncated forms of *T* and  $R_{\mu}$ , given by Eqs. (3.61) and (3.62), obtained, we correct the resulting energies,  $E_{\mu}^{(P)}$ , for the correlation effects involving excited determinants  $|\Phi_K\rangle$  from another subspace  $\mathscr{H}^{(Q)} \subseteq (\mathscr{H}^{(0)} \oplus \mathscr{H}^{(P)})^{\perp}$ , referred to as the *Q*-space ( $\mathscr{H}^{(0)}$  is the one-dimensional subspace spanned by the reference determinant  $|\Phi\rangle$ ). This is accomplished by generalizing the previously discussed biorthogonal MMCC expansion, given by Eq. (3.27), to obtain the *Q*-space-corrected energies  $E_{\mu}^{(P+Q)}$ , in the following manner [163, 164]:

$$E_{\mu}^{(P+Q)} \equiv E_{\mu}^{(P)} + \delta_{\mu}(P;Q), \qquad (3.68)$$

where the non-iterative correction  $\delta_{\mu}(P;Q)$  due to the Q-space contributions missing in the *P*-space CC/EOMCC calculations is defined as

$$\delta_{\mu}(P;Q) = \sum_{\substack{|\Phi_{K}\rangle \in \mathscr{H}^{(Q)}\\ \operatorname{rank}(|\Phi_{K}\rangle) \leq \min(N_{\mu}^{(P)},\Xi^{(Q)})}} \ell_{\mu,K}(P) \mathfrak{M}_{\mu,K}(P), \qquad (3.69)$$

with

$$\mathfrak{M}_{\mu,K}(P) = \langle \Phi_K | (\bar{H}^{(P)} R^{(P)}_{\mu}) | \Phi \rangle$$
(3.70)

representing the generalized moments of the CC/EOMCC equations corresponding to the calculations with  $T = T^{(P)}$  and  $R_{\mu} = R_{\mu}^{(P)}$ , associated with the projections of these equations on the  $|\Phi_K\rangle$  determinants from the Q-space  $\mathscr{H}^{(Q)}$ . As in the case of the considerations discussed in the previous sections, we formally define  $R_{\mu=0}^{(P)} = \mathbf{1}$ , so that Eq. (3.70) and the equations below cover the ground-state ( $\mu = 0$ ) case and excited ( $\mu > 0$ ) states. The symbol  $N_{\mu}^{(P)}$  in Eq. (3.69) is the highest possible many-body rank of the excited determinant  $|\Phi_K\rangle$  relative to the reference  $|\Phi\rangle$  for which the generalized moments  $\mathfrak{M}_{\mu,K}(P)$  of the P-space calculations, given by Eq. (3.70), are still still non-zero. The symbol  $\Xi^{(Q)}$  in Eq. (3.69) is the highest many-body rank of the excited determinants  $|\Phi_K\rangle$  included in  $\mathscr{H}^{(Q)}$ . For

example, if we want to use Eq. (3.69) to correct the ground-state CCSDt energies for the triples missing in CCSDt, as in CC(t;3),  $N_{\mu}^{(P)} = 10$ , but  $\Xi^{(Q)} = 3$ , so that, as expected,  $\min(N_{\mu}^{(P)}, \Xi^{(Q)}) = 3$ . If we want to use it to correct CCSD energies for triples and quadruples, as in CR-CC(2,4),  $N_{\mu}^{(P)} = 6$ , but  $\Xi^{(Q)} = 4$ , so  $\min(N_{\mu}^{(P)}, \Xi^{(Q)}) = 4$ , etc.

As shown in Ref. [163], the Q-space-corrected energy  $E_{\mu}^{(P+Q)}$  defined by Eq. (3.68) can be derived by approximating its exact analog, similar to Eqs. (3.26) and (3.27), which defines the difference  $\delta_{\mu}^{(P)}$  between the full CI energies  $E_{\mu}$  and the corresponding energies  $E_{\mu}^{(P)}$ obtained in the P-space CC/EOMCC calculations,

$$\delta_{\mu}^{(P)} \equiv E_{\mu} - E_{\mu}^{(P)} = \sum_{\substack{|\Phi_{K}\rangle \in (\mathscr{H}^{(0)} \oplus \mathscr{H}^{(P)})^{\perp} \\ \operatorname{rank}(|\Phi_{K}\rangle) \leq N_{\mu}^{(P)}}} \ell_{\mu,K}(P) \mathfrak{M}_{\mu,K}(P), \qquad (3.71)$$

which can be derived by considering the asymmetric energy expression

$$E_{\mu} = \langle \Psi_{\mu} | H R_{\mu}^{(P)} e^{T^{(P)}} | \Phi \rangle / \langle \Psi_{\mu} | R_{\mu}^{(P)} e^{T^{(P)}} | \Phi \rangle$$
(3.72)

and replacing the exact bra state  $\langle \Psi_{\mu} |$  by the *P*-space analog of Eq. (3.21):

$$\langle \Psi_{\mu} | = \langle \Phi | \mathscr{L}_{\mu} e^{-T^{(P)}}, \qquad (3.73)$$

where

$$\mathscr{L}_{\mu} = \mathscr{L}_{\mu}^{(P)} + \delta \mathscr{L}_{\mu}^{(P)}, \qquad (3.74)$$

with

$$\mathscr{L}^{(P)}_{\mu} = \delta_{\mu,0} \mathbf{1} + \sum_{|\Phi_K\rangle \in \mathscr{H}^{(P)}} \ell_{\mu,K} (E_K)^{\dagger}$$
(3.75)

and

$$\delta \mathscr{L}^{(P)}_{\mu} = \sum_{|\Phi_K\rangle \in (\mathscr{H}^{(0)} \oplus \mathscr{H}^{(P)})^{\perp}} \ell_{\mu,K}(E_K)^{\dagger}.$$
(3.76)

One obtains Eq. (3.68) by limiting the summation over the excited determinants  $|\Phi_K\rangle$  from the orthogonal complement of  $\mathscr{H}^{(0)} \oplus \mathscr{H}^{(P)}$  in Eq. (3.71) to the determinants that belong to  $\mathscr{H}^{(Q)}$  which is a subspace of  $(\mathscr{H}^{(0)} \oplus \mathscr{H}^{(P)})^{\perp}$ , while replacing the exact values of the  $\ell_{\mu,K}$ amplitudes entering Eq. (3.71), which one can obtain by solving the similarity-transformed form of the bra Schrödinger equation,

$$\langle \Phi | \mathscr{L}_{\mu} \bar{H}^{(P)} = E_{\mu} \langle \Phi | \mathscr{L}_{\mu}, \qquad (3.77)$$

in the entire many-electron Hilbert space  $\mathscr{H}$ , by their approximate  $\ell_{\mu,K}(P)$  values resulting from constraining Eq. (3.77) to the total subspace of interest, i.e.,  $\mathscr{H}^{(0)} \oplus \mathscr{H}^{(P)} \oplus \mathscr{H}^{(Q)}$ , based on the information obtained in the preceding *P*-space CC/EOMCC calculations. In analogy to the CR-CC/CR-EOMCC methods discussed in Section 3.1.1, in order to come up with practical schemes Eqs. (3.68)–(3.70), we have to propose a procedure which would enable one to determine the  $\ell_{\mu,K}(P)$  amplitudes that multiply moments  $\mathfrak{M}_{\mu,K}(P)$  in the definition of the correction  $\delta_{\mu}(P;Q)$  in a computationally manageable fashion.

One such procedure, proposed in Ref. [163] and adopted in this thesis, is based on the realization that Eqs. (3.73)–(3.76) are reminiscent of the expressions that define the CC/EOMCC bra states

$$\langle \tilde{\Psi}_{\mu}^{(P)} | = \langle \Phi | L_{\mu}^{(P)} e^{-T^{(P)}},$$
(3.78)

resulting from the P-space CC/EOMCC calculations and forming a biorthogonal basis with

the corresponding CC/EOMCC ket states

$$|\Psi_{\mu}^{(P)}\rangle = R_{\mu}^{(P)}|\Psi_{0}^{(P)}\rangle = R_{\mu}^{(P)}e^{T^{(P)}}|\Phi\rangle.$$
(3.79)

Thus, we can approximate the deexcitation operator  $\mathscr{L}_{\mu}$  that serves as a source of amplitudes  $\ell_{\mu,K}(P)$  by splitting it into the known, *a priori* determined, *P*-space component  $L_{\mu}^{(P)}$  and the unknown component  $\mathscr{L}_{\mu}^{(Q)}$  that provides information about the desired  $\ell_{\mu,K}(P)$  amplitudes corresponding to the determinants  $|\Phi_K\rangle \in \mathscr{H}^{(Q)}$  and entering the correction  $\delta_{\mu}(P;Q)$ , Eq. (3.69). The former component, given by

$$L_{\mu}^{(P)} = L_{\mu,0}^{(P)} + L_{\mu,\text{open}}^{(P)} \equiv \delta_{\mu,0} \mathbf{1} + \sum_{|\Phi_K\rangle \in \mathscr{H}^{(P)}} l_{\mu,K} (E_K)^{\dagger}, \qquad (3.80)$$

is obtained in the usual way by solving the truncated, P-space, left CC/EOMCC eigenvalue problem (cf. Eq. (3.77)),

$$\delta_{\mu,0} \langle \Phi | \bar{H}_{\text{open}}^{(P)} | \Phi_K \rangle + \langle \Phi | L_{\mu,\text{open}}^{(P)} \bar{H}_{\text{open}}^{(P)} | \Phi_K \rangle = \omega_\mu^{(P)} l_{\mu,K}, \quad |\Phi_K \rangle \in \mathscr{H}^{(P)}.$$
(3.81)

The unknown  $\mathscr{L}^{(Q)}_{\mu}$  contribution to the  $\mathscr{L}_{\mu}$  operator, which is now approximated in the following manner:

$$\mathscr{L}_{\mu} \approx L_{\mu}^{(P)} + \mathscr{L}_{\mu}^{(Q)}, \qquad (3.82)$$

where

$$\mathscr{L}^{(Q)}_{\mu} = \sum_{|\Phi_K\rangle \in \mathscr{H}^{(Q)}} \ell_{\mu,K}(P) (E_K)^{\dagger}, \qquad (3.83)$$

can be obtained by examining the projected form of the exact left eigenvalue problem given

by Eq. (3.77) within the  $\mathscr{H}^{(Q)}$  subspace and exploiting the information obtained in the preceding *P*-space CC/EOMCC calculations. Indeed, by replacing the exact  $\mathscr{L}_{\mu}$  operator Eq. (3.77) by its approximate form given by Eq. (3.82) and by right projecting the resulting equation on the  $|\Phi_K\rangle$  determinants from  $\mathscr{H}^{(Q)}$ , we can show that the  $\ell_{\mu,K}(P)$  amplitudes defining  $\mathscr{L}^{(Q)}_{\mu}$  satisfy the system of coupled linear equations that looks as follows:

$$\begin{split} \langle \Phi | L_{\mu}^{(P)} \bar{H}^{(P)} | \Phi_K \rangle &+ \sum_{|\Phi_{K'}\rangle \in \mathscr{H}^{(Q)}} \langle \Phi_{K'} | \bar{H}^{(P)} | \Phi_K \rangle \, \ell_{\mu,K'}(P) \\ &= E_{\mu} \, \ell_{\mu,K}(P), \ |\Phi_K \rangle \in \mathscr{H}^{(Q)}. \end{split}$$
(3.84)

If we further approximate the exact energy  $E_{\mu}$  in Eq. (3.84) by the energy  $E_{\mu}^{(P)}$  resulting from the preceding *P*-space CC/EOMCC calculations and move the diagonal, K' = K, contributions to the summation on the left-hand side of Eq. (3.84) to the right-hand side, as in the Epstein–Nesbet partitioning, we obtain

$$\begin{split} \langle \Phi | L_{\mu}^{(P)} \bar{H}^{(P)} | \Phi_K \rangle &+ \sum_{|\Phi_{K'}\rangle \in \mathscr{H}^{(Q)}, K' \neq K} \langle \Phi_{K'} | \bar{H}^{(P)} | \Phi_K \rangle \, \ell_{\mu,K'}(P) \\ &= D_{\mu,K}(P) \, \ell_{\mu,K}(P), \ |\Phi_K \rangle \in \mathscr{H}^{(Q)}, \end{split}$$
(3.85)

where

$$D_{\mu,K}(P) = E_{\mu}^{(P)} - \langle \Phi_K | \bar{H}^{(P)} | \Phi_K \rangle$$
(3.86)

is the corresponding perturbative denominator, analogous to that used in the CR-CC(2,3) and other  $CC(m_a, m_b)$  approximations (see in Eq. (3.31) in Section 3.1.1.). The system defined by Eq. (3.85) contains couplings among the  $|\Phi_K\rangle$  determinants from  $\mathscr{H}^{(Q)}$ , defined by the off-diagonal matrix elements of  $\bar{H}^{(P)}$  in the  $\mathscr{H}^{(Q)}$  subspace, and we can attempt to solve it iteratively. This would constitute one of the possible ways of determining the desired  $\ell_{\mu,K}(P)$  amplitudes that enter the corrections  $\delta_{\mu}(P;Q)$ , Eq. (3.69). Alternatively, we could follow the philosophy of the CR-CC(2,3) and other CR-CC( $m_a, m_b$ ) schemes and ignore the off-diagonal elements of the matrix representing  $\bar{H}^{(P)}$  in the Q-space and replace the expression given by Eq. (3.85) by the simple non-iterative formula

$$\ell_{\mu,K}(P) = \langle \Phi | L_{\mu}^{(P)} \bar{H}^{(P)} | \Phi_K \rangle / D_{\mu,K}(P), \qquad (3.87)$$

and use Eq. (3.85), limited to small blocks of the  $\bar{H}^{(P)}$  matrix involving the degenerate Q-space determinants  $|\Phi_K\rangle$ , only if there are orbital degeneracies and only if we are interested in maintaining the strict invariance of the resulting energies  $E^{(P+Q)}_{\mu}$  with respect to rotations among degenerate orbitals relevant to non-Abelian symmetries only.

Eqs. (3.68)–(3.70), with  $\ell_{\mu,K}(P)$  defined by Eq. 3.87 are the basis expressions for the CC(P;Q) methods used in this thesis project. The CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) approaches which are based on these expression and which are exploited and further developed in the work, are discussed next.

## 3.1.4 CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) Hierarchy

While the general formalism described in the previous section was laid down for both ground and excited states, the detailed working equations describing the CC(P;Q)-based CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) methods presented below focus on ground-state considerations, since this thesis project deals with the development of methods for quasidegeneracies in the ground electronic states or lowest-energy states of a given multiplicity.

The ground-state CC(t;3) approach, initially implemented and applied by Dr. Jun Shen

from our group [163-165] and reimplemented here, starts by solving the CCSDt amplitude equations (see Eqs. (3.11) and (3.37)–(3.41)) and determining the CCSDt energy,

$$E_0^{(\text{CCSDt})} = \langle \Phi | \bar{H}^{(\text{CCSDt})} | \Phi \rangle.$$
(3.88)

Then, following the general CC(P;Q) recipe discussed in Section 3.1.2, we proceed to the calculation of the desired CC(t;3) ground-state energy  $E_0^{(CC(t;3))}$  defined as

$$E_0^{(\mathrm{CC}(\mathbf{t};3))} = E_0^{(\mathrm{CCSDt})} + \delta_0(\mathbf{t};3), \qquad (3.89)$$

where, the non-iterative correction  $\delta_0(t; 3)$  due to the triples missing in the CCSDt consideration is determined using the formula (cf. Eqs. (3.68)–(3.70))

$$\delta_{0}(\mathbf{t};3) = \sum_{\substack{|\Phi_{ijk}^{abc}\rangle \in \mathscr{H}^{(\mathrm{T})} \ominus \mathscr{H}^{(\mathrm{t})}}} \ell_{0,ijk}^{abc}(\mathrm{CCSDt}) \,\mathfrak{M}_{0,abc}^{ijk}(\mathrm{CCSDt}),$$
(3.90)

with the  $\mathfrak{M}^{ijk}_{0,abc}(\mathrm{CCSDt})$  moments defined as:

$$\mathfrak{M}_{0,abc}^{ijk}(\text{CCSDt}) = \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSDt})} | \Phi \rangle.$$
(3.91)

The  $\ell_{0,ijk}^{abc}(\text{CCSDt})$  amplitudes entering Eq. (3.90) are calculated as (cf. Eq. (3.87))

$$\ell_{0,ijk}^{abc}(\text{CCSDt}) = \langle \Phi | L_0^{(\text{CCSDt})} \bar{H}^{(\text{CCSDt})} | \Phi_{ijk}^{abc} \rangle / D_{0,abc}^{ijk}(\text{CCSDt}), \qquad (3.92)$$

where

$$D_{0,abc}^{ijk}(\text{CCSDt}) = E_0^{(\text{CCSDt})} + \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSDt})} | \Phi_{ijk}^{abc} \rangle.$$
(3.93)

The summation over triply excited determinants  $|\Phi_{ijk}^{abc}\rangle$  in Eq. (3.90) excludes those that are already included in the CCSDt calculations, which are listed on Table 3.1 as classes 10–16. The deexcitation operator  $L_0^{(\text{CCSDt})}$ , entering Eq. (3.92) and given by

$$L_0^{(\text{CCSDt})} = \mathbf{1} + L_{0,1} + L_{0,2} + l_{0,3}, \qquad (3.94)$$

where  $L_{0,1}$  and  $L_{0,2}$  are the regular one- and two-body components of  $L_0^{(\text{CCSDt})}$  defined using all correlated spin-orbitals and

$$l_{0,3} = \sum_{\mathbf{I}>j>k,a>b>\mathbf{C}} l_{0,\mathbf{I}jk}^{ab\mathbf{C}} E_{ab\mathbf{C}}^{\mathbf{I}jk}$$
(3.95)

is the approximate triply deexcited active-space component, which defines the CCSDt bra state

$$\langle \tilde{\Psi}_0^{(\text{CCSDt})} | = \langle \Phi | L_0^{(\text{CCSDt})} e^{-T^{(\text{CCSDt})}}, \qquad (3.96)$$

is obtained by solving the left-eigenstate CCSDt equations. In analogy to the CR-CC(2,3) approach discussed in Section 3.1.1, the Epstein–Nesbet-like  $D_{0,abc}^{ijk}$ (CCSDt) denominator entering Eq. (3.92) and given by Eq. (3.93) defines the most complete variant D of CC(t;3) abbreviated as CC(t;3)<sub>D</sub>. If we replace the  $D_{0,abc}^{ijk}$ (CCSDt) denominator given by Eq. (3.93) by the corresponding Møller–Plesset-like expression,

$$D_{0,abc}^{ijk}(\text{CCSDt}) = \epsilon_i + \epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b - \epsilon_c, \qquad (3.97)$$

where  $\epsilon_p$  is the single-particle energy associated with spin-orbital p (the diagonal elements of the Fock matrix) we end up with variant A of CC(t;3) labeled CC(t;3)<sub>A</sub>. As shown in Table 1.1, the CC(t;3)<sub>A</sub> and CC(t;3)<sub>D</sub> methods have a CPU time scalings of  $N_o N_u n_o^2 n_u^4$  in the iterative active-space CCSDt part and  $\sim n_o^3 n_u^4$  in the calculation of the corresponding triples correction  $\delta_0(t;3)$ .

In extending the CC(P;Q) hierarchy to include connected quadruple excitations, which is the main objective of this thesis project, we can start by solving the CCSDtq amplitude equations and determining the CCSDtq energy,

$$E_0^{(\text{CCSDtq})} = \langle \Phi | \bar{H}^{(\text{CCSDtq})} | \Phi \rangle, \qquad (3.98)$$

and then correct the  $E_0^{(\text{CCSDtq})}$  for those triples and quadruples that are missing in CCSDtq to obtain the CC(t,q;3,4) energy  $E_0^{(\text{CC}(t,q;3,4))}$  as follows:

$$E_0^{(\mathrm{CC}(\mathbf{t},\mathbf{q};3,4))} = E_0^{(\mathrm{CCSDtq})} + \delta_0(\mathbf{t},\mathbf{q};3,4),$$
(3.99)

where

$$\delta_{0}(\mathbf{t},\mathbf{q};3,4) = \sum_{\substack{|\Phi_{ijk}^{abc}\rangle \in \mathscr{H}^{(\mathrm{T})} \ominus \mathscr{H}^{(\mathrm{t})} \\ + \sum_{\substack{|\Phi_{ijkl}^{abcd}\rangle \in \mathscr{H}^{(\mathrm{Q})} \ominus \mathscr{H}^{(\mathrm{q})}}} \ell_{0,ijkl}^{abcd}(\mathrm{CCSDtq}) \mathfrak{M}_{0,abcd}^{ijkl}(\mathrm{CCSDtq}), \quad (3.100)$$

with the  $\mathfrak{M}^{ijk}_{0,abc}(\text{CCSDtq})$  and  $\mathfrak{M}^{ijkl}_{0,abcd}(\text{CCSDtq})$  moments given by

$$\mathfrak{M}_{0,abc}^{ijk}(\text{CCSDtq}) = \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSDtq})} | \Phi \rangle$$
(3.101)

and

$$\mathfrak{M}_{0,abcd}^{ijkl}(\text{CCSDtq}) = \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSDtq})} | \Phi \rangle, \qquad (3.102)$$

respectively. The  $\mathscr{H}^{(\mathrm{T})} \oplus \mathscr{H}^{(\mathrm{t})}$  and  $\mathscr{H}^{(\mathrm{Q})} \oplus \mathscr{H}^{(\mathrm{q})}$  subspaces entering Eq. (3.100) are spanned by the triply and quadruply excited determinants other than those that belong to the  $|\Phi_{ijk}^{abc}\rangle$  and  $|\Phi_{ijkl}^{abcd}\rangle$  categories included in the CCSDtq calculations. To be more precise, the determinants entering the  $\mathscr{H}^{(\mathrm{T})} \oplus \mathscr{H}^{(\mathrm{t})}$  and  $\mathscr{H}^{(\mathrm{Q})} \oplus \mathscr{H}^{(\mathrm{q})}$  subspaces are those that belong to classes 10–25 in Table 3.1. The  $\ell_{0,ijk}^{abc}$  (CCSDtq) and  $\ell_{0,ijkl}^{abcd}$  (CCSDtq) amplitudes entering Eq. (3.100) are calculated as

$$\ell_{0,ijk}^{abc}(\text{CCSDtq}) = \langle \Phi | L_0^{(\text{CCSDtq})} \bar{H}^{(\text{CCSDtq})} | \Phi_{ijk}^{abc} \rangle / D_{0,abc}^{ijk}(\text{CCSDtq})$$
(3.103)

and

$$\ell_{0,ijkl}^{abcd}(\text{CCSDtq}) = \langle \Phi | L_0^{(\text{CCSDtq})} \bar{H}^{(\text{CCSDtq})} | \Phi_{ijkl}^{abcd} \rangle / D_{0,abcd}^{ijkl}(\text{CCSDtq}), \qquad (3.104)$$

where the  $L_0^{(\text{CCSDtq})}$  deexcitation operator resulting from the left-eigenstate CCSDtq calculations is defined as

$$L_0^{(\text{CCSDtq})} = \mathbf{1} + L_{0,1} + L_{0,2} + l_{0,3} + l_{0,4}, \qquad (3.105)$$

with  $l_{0,3}$  given by Eq. (3.95) and

$$l_{0,4} = \sum_{\mathbf{I} > \mathbf{J} > k > l, a > b > \mathbf{C} > \mathbf{D}} l_{0,\mathbf{IJ}kl}^{ab\mathbf{CD}} E_{ab\mathbf{CD}}^{\mathbf{IJ}kl}.$$
(3.106)

In the formally most complete formulation of the CC(t,q;3,4) approach, abbreviated as  $CC(t,q;3,4)_{DD}$ , the denominators that enter Eqs. (3.103) and (3.104) are given by the

Epstein–Nesbet-like expressions,

$$D_{0,abc}^{ijk}(\text{CCSDtq}) = E_0^{(\text{CCSDtq})} - \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSDtq})} | \Phi_{ijk}^{abc} \rangle$$
(3.107)

and

$$D_{0,abcd}^{ijkl}(\text{CCSDtq}) = E_0^{(\text{CCSDtq})} - \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSDtq})} | \Phi_{ijkl}^{abcd} \rangle.$$
(3.108)

We can, however, contemplate simplified CC(t,q;3,4) schemes, where one or both denominators  $D_{0,abc}^{ijk}(CCSDtq)$  and  $D_{0,abcd}^{ijkl}(CCSDtq)$  is or are replaced by the Møller–Plesset-like expressions,

$$D_{0,abc}^{ijk}(\text{CCSDtq}) = \epsilon_i + \epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b - \epsilon_c$$
(3.109)

and

$$D_{0,abcd}^{ijkl}(\text{CCSDtq}) = \epsilon_i + \epsilon_j + \epsilon_k + \epsilon_l - \epsilon_a - \epsilon_b - \epsilon_c - \epsilon_d.$$
(3.110)

If we use the Møller–Plesset forms of  $D_{0,abc}^{ijk}$ (CCSDtq) and  $D_{0,abcd}^{ijkl}$ (CCSDtq), the resulting CC(t,q;3,4) method is abbreviated as CC(t,q;3,4)<sub>AA</sub>. If we use the the more complete Epstein–Nesbet form for the  $D_{0,abc}^{ijk}$ (CCSDtq) denominator entering the triples part of the correction  $\delta_0(t,q;3,4)$  and the Møller–Plesset form for  $D_{0,abcd}^{ijkl}$ (CCSDtq), defining the correction due to quadruples, we end up with CC(t,q;3,4)<sub>DA</sub>. It does not make much sense to consider the CC(t,q;3,4)<sub>AD</sub> scheme, where  $D_{0,abcd}^{ijkl}$ (CCSDtq) is defined by the less complete Møller–Plesset expression and  $D_{0,abcd}^{ijkl}$ (CCSDtq) by the more accurate Epstein–Nesbet form, but, as already mentioned, we can use the Epstein–Nesbet denominator in the triples and quadruples parts of  $\delta_0(t,q;3,4)$ , as in the aforementioned CC(t,q;3,4)<sub>DD</sub> approximation. There is, however, one issue that makes the CC(t,q;3,4)<sub>DD</sub> scheme less attractive, namely, the issue of computer costs. As shown in Table 1.1, the CC(t,q;3,4)<sub>AA</sub> and CC(t,q;3,4)<sub>DD</sub> approaches are characterized by the iterative CPU steps of CCSDtq that scale as  $N_o^2 N_u^2 n_o^2 n_u^4$ and the non-iterative steps of  $\delta_0(t, q; 3, 4)$  that scale as  $n_o^3 n_u^4$  in the triples part of it and  $n_o^2 n_u^5$  in the quadruples part of it. In other words, the AA and DA variants of CC(t,q;3,4) replace the iterative  $\mathcal{N}^{10}$  steps of full CCSDTQ by the iterative  $\mathcal{N}^6$ -like and non-iterative  $\mathcal{N}^7$ -type operations. The situation with CC(t,q;3,4)<sub>DD</sub> is, however, different (see Table 1.1). The CC(t,q;3,4)<sub>DD</sub> method has identical  $\mathcal{N}^6$ -like iterative steps and identical  $\mathcal{N}^7$ -like steps in the triples part of  $\delta_0(t,q;3,4)$  as in the case of the AA and DA approximations, but it is considerably more expensive in the quadruples part of  $\delta_0(t,q;3,4)$ , increasing the  $n_o^2 n_u^5$ scaling of the AA and DA variants of CC(t,q;3,4) to a  $n_o^4 n_u^5$  ( $\mathcal{N}^9$ -like) level. For this reason, our focus in this work is on the CC(t,q;3,4)<sub>DA</sub> approach, with CC(t,q;3,4)<sub>AA</sub> being a possible alternative.

As shown in this thesis, the CC(t,q;3,4) approach can be very accurate, producing the CCSDTQ-level results at the fraction of the cost, but one can improve the CC(t,q;3,4) description further by treating triples fully via full CCSDTq and correcting the CCSDTq calculations for the quadruples outside the "little q" set using the CC(P;Q)-style correction defining the CC(q,4) theory. In this case, we solve the CCSDTq amplitude equations and determine the CCSDTq energy,

$$E_0^{(\text{CCSDTq})} = \langle \Phi | \bar{H}^{(\text{CCSDTq})} | \Phi \rangle, \qquad (3.111)$$

and then correct  $E_0^{(\text{CCSDTq})}$  for the quadruples outside the "little q" set using the formula

$$E_0^{(\mathrm{CC}(\mathbf{q};4))} = E_0^{(\mathrm{CCSDTq})} + \delta_0(\mathbf{q};4), \qquad (3.112)$$

where

$$\delta_{0}(\mathbf{q};4) = \sum_{\substack{|\Phi_{ijkl}^{abcd}\rangle \in \mathscr{H}^{(\mathbf{Q})} \ominus \mathscr{H}^{(\mathbf{q})}}} \ell_{0,ijkl}^{abcd}(\mathbf{CCSDTq}) \,\mathfrak{M}_{0,abcd}^{ijkl}(\mathbf{CCSDTq}), \quad (3.113)$$

with

$$\mathfrak{M}_{0,abcd}^{ijkl}(\text{CCSDTq}) = \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSDTq})} | \Phi \rangle$$
(3.114)

representing the corresponding moments of CCSDTq representing the projections of the CCSDTq equations on the missing quadruply excited determinants. The  $\ell_{0,ijkl}^{abcd}$ (CCSDTq) amplitudes entering Eq. (3.114) are calculated as

$$\ell_{0,ijkl}^{abcd}(\text{CCSDTq}) = \langle \Phi | L_0^{(\text{CCSDTq})} \bar{H}^{(\text{CCSDTq})} | \Phi_{ijkl}^{abcd} \rangle / D_{0,abcd}^{ijkl} (\text{CCSDTq}), \qquad (3.115)$$

where the  $L_0^{(\text{CCSDTq})}$  deexcitation operator, obtained in the left-eigenstate CCSDTq calculations, is defined as

$$L_0^{(\text{CCSDTq})} = \mathbf{1} + L_{0,1} + L_{0,2} + L_{0,3} + l_{0,4}$$
(3.116)

and the denominator  $D_{0,abcd}^{ijkl}$  (CCSDTq) in the most complete formulation of CC(q;4) is given by

$$D_{0,abcd}^{ijkl}(\text{CCSDTq}) = E_0^{(\text{CCSDTq})} - \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSDTq})} | \Phi_{ijkl}^{abcd} \rangle.$$
(3.117)

Once again, the denominator  $D_{0,abcd}^{ijkl}$  (CCSDTq), Eq. (3.117), is the Epstein–Nesbet-like expression corresponding to variant D of CC(q;4), denoted CC(q;4)<sub>D</sub>, but we can also replace

Eq. (3.117) by the corresponding Møller-Plesset-like equation,

$$D_{0,abcd}^{ijkl}(\text{CCSDTq}) = \epsilon_i + \epsilon_j + \epsilon_k + \epsilon_l - \epsilon_a - \epsilon_b - \epsilon_c - \epsilon_d, \qquad (3.118)$$

resulting in the A variant of CC(q;4), denoted CC(q;4)<sub>A</sub>. The CPU time scaling associated with the CC(q;4) method is  $N_o N_u n_o^3 n_u^5$  in the active-space CCSDTq calculations and  $n_o^2 n_u^5$ , when CC(q;4)<sub>A</sub> is employed, or  $n_o^4 n_u^5$ , when CC(q;4)<sub>D</sub> is used, in the quadruples correction part. In analogy to CC(t,q;3,4)<sub>DA</sub> vs CC(t,q;3,4)<sub>DD</sub>, the CC(q;4)<sub>A</sub> model is preferred due to its less expensive  $n_o^2 n_u^5$  ( $\mathcal{N}^7$ -like) steps compared to  $n_o^4 n_u^5$  ( $\mathcal{N}^9$ -like) steps of CC(q;4)<sub>D</sub>.

We can also utilize the flexibility of the CC(P;Q) methodology and consider methods "in between". For example, we can assume that almost all dynamical and non-dynamical correlations are captured by the active-space approach with selected triple and quadruple excitations (CCSDtq) and we then correct the corresponding CCSDtq energy  $E_0^{(CCSDtq)}$  for the remaining triples, ignoring the quadruples outside the "little q" set missing in CCSDtq. The resulting CC(t,q;3) approach, in which the energy is calculated as

$$E_0^{(\mathrm{CC}(\mathbf{t},\mathbf{q};3))} = E_0^{(\mathrm{CCSDtq})} + \delta_0(\mathbf{t},\mathbf{q};3), \qquad (3.119)$$

where

$$\delta_{0}(\mathbf{t},\mathbf{q};3) = \sum_{\substack{|\Phi_{ijk}^{abc}\rangle \in \mathscr{H}^{(\mathrm{T})} \ominus \mathscr{H}^{(\mathrm{t})}}} \ell_{0,ijk}^{abc}(\mathrm{CCSDtq}) \,\mathfrak{M}_{0,abc}^{ijk}(\mathrm{CCSDtq}), \quad (3.120)$$

is studied in this work as well, being a reasonable compromise between the CC(t;3) and CC(t,q;3,4) methods, as discussed in the following section. The CPU time scaling in the

CC(t,q;3) approach is  $N_o^2 N_u^2 n_o^2 n_u^4$  in the active-space CCSDtq calculations and  $n_o^3 n_u^4$  associated with the non-iterative triples correction part.

The ground-state CC(t;3) method was first implemented and applied by Dr. Jun Shen in our group [163–165]. The development and application of the CC(t,q;3), CC(t,q;3,4), and CC(q;4) methods are the basis for this dissertation, although the CC(t;3) codes were produced in this project as well. In our present implementation, discussed further in Chapter 4, all of these methods are interfaced with the RHF, ROHF, and integral transformation routines available in the GAMESS package [277,278]. Since we have not yet developed the codes that could solve the left-eigenstate CCSDt, CCSDtq, or CCSDTq equations, which would normally be needed to determine the  $\ell_{0,i_1...i_n}^{a_1...a_n}(A)$  (A = CCSDt, CCSDtq, or CCSDTq), amplitudes for the ground-state corrections described above, we have introduced a few simplifications in the CC(P;Q) routines. Thus, we approximate the similarity-transformed Hamiltonians of the active-space CCSDt, CCSDtq, and CCSDTq methods, Eqs. (3.38), (3.47), and (3.54), respectively, which enter the ground-state moments,  $\mathfrak{M}_{0,abc}^{ijk}$  (CCSDtq), Eq. (3.101),  $\mathfrak{M}_{0,abcd}^{ijkl}$  (CCSDtq), Eq. (3.102), and  $\mathfrak{M}_{0,abcd}^{ijkl}$  (CCSDTq), Eq. (3.114), by the CCSD-like

$$\bar{H}^{(A)}(2) = e^{-T_1 - T_2} H e^{T_1 + T_2} = (H e^{T_1 + T_2})_C, \qquad (3.121)$$

in which the higher-than-two-body components of the relevant  $T^{(\text{CCSDt})}$ ,  $T^{(\text{CCSDtq})}$ , and  $T^{(\text{CCSDTq})}$  operators are neglected, although – and this needs to be strongly emphasized – the  $T_1$  and  $T_2$  amplitudes entering Eq. (3.121) originate from the true CCSDt, CCSDtq, or CCSDTq calculations, i.e., they are properly relaxed in the presence of  $t_3$  or  $T_3$  and  $t_4$ . Moreover, we replace the full definitions of  $L_0^{(\text{CCSDt})}$ ,  $L_0^{(\text{CCSDtq})}$ , and  $L_0^{(\text{CCSDTq})}$ , Eqs.
(3.94), (3.105), and (3.116), respectively, which enter the  $\ell_{0,ijk}^{abc}(\text{CCSDt})$ ,  $\ell_{0,ijk}^{abc}(\text{CCSDtq})$ ,  $\ell_{0,ijk}^{abcd}(\text{CCSDtq})$ , and  $\ell_{0,ijkl}^{abcd}(\text{CCSDTq})$  amplitudes in the definitions of the CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) corrections, by

$$L_0^{(A)}(2) = \mathbf{1} + L_{0,1} + L_{0,2}, \qquad (3.122)$$

where A = CCSDt, CCSDtq, and CCSDTq, in which the higher-than-two-body components of  $L_0^{(A)}$  are neglected and the one- and two-body components,  $L_{0,1}$  and  $L_{0,2}$ , respectively, are obtained by solving the left eigenvalue problem involving  $\bar{H}^{(A)}(2)$ , Eq. (3.121), in the space of single and double excitations. In this way, we account for the relaxation of the  $T_1$ ,  $T_2$ ,  $L_{0,1}$ , and  $L_{0,2}$  amplitudes in the presence of the higher-than-two-body components of the CCSDt, CCSDtq, and CCSDTq cluster operators, which becomes significant when the latter components become large, as is the case in the biradical and bond-breaking regions of the potential energy surface, but avoid complex computational steps related to the full use of  $\bar{H}^{(\text{CCSDt})}$ ,  $\bar{H}^{(\text{CCSDtq})}$ , and  $\bar{H}^{(\text{CCSDTq})}$  Hamiltonians. The initial studies involving CC(t;3) defined in this approximate manner, reported in Refs. [163–165], and calculations performed in this thesis, have revealed that the CC(t;3) results match the total and relative energetics of full CCSDT almost exactly, despite its truncated form, defined by Eqs. (3.121) and 3.122. In other words, replacing the true similarity-transformed Hamiltonian of CCSDt by  $\bar{H}^{(\text{CCSDt})}(2)$  and the true deexcitation operator  $L_0^{(\text{CCSDt})}$  by  $L_0^{(\text{CCSDt})}(2)$ , as in Eqs. (3.121) and (3.122), has no noticeable effect on the calculated energies. We can expect that the same applies to the CC(t,q;3), CC(t,q;3,4), and CC(q;4) calculations, and all of our tests to date indicate that using Eqs. (3.121) and (3.122) instead of the full forms of  $L_0^{(A)}$  and  $\bar{H}^{(A)}(2)$  is fine, but we will return to this topic in the future, once routines enabling the complete treatment of the similarity-transformed Hamiltonians and deexcitation operators of CCSDt, CCSDtq, and CCSDTq, entering the CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) corrections are developed. Additional technical details of our efficient implementation of the CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) methods, especially the latter three approaches and the underlying CCSDtq and CCSDTq and parent CCSDTQ codes, are given in Chapter 4. In the next section, we illustrate the accuracies the CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations can provide, especially when compared to the parent CCSDT and CCSDTQ data, and full CI. Our CC(q;4) codes have been used by us as well, but since they are more expensive, needing CCSDTq computations, and since they have been mostly used by Mr. Ilias Magoulas from our group, we focus here on the CC(t;3), CC(t,q;3), and CC(t,q;3,4) computations.

## **3.2** Numerical Results

## 3.2.1 The $C_{2v}$ -Symmetric Double Dissociation of Water

The performance of the CC(t;3), CC(t,q;3), and CC(t,q;3,4) approaches, particularly the latter two methods that describe quadruples effects, was examined and evaluated by applying them to a few molecular problems for which the exact, full CI, or nearly exact, CCSDTQ, solutions are known or not to difficult to generate. Our focus is on situations where connected quadruples are important, which is the case when double bonds are broken or when certain types of chemical reactions and biradicals are examined. One particularly common example used to test the performance of new quantum chemistry methods, especially CC methods with quadruples, is the double dissociation of water, in which we examine the ground-state potential energy surface as both O–H bonds are simultaneously stretched by up to two or even three times the equilibrium bond lengths.

The various CC calculations of the double-dissociation of  $H_2O$  were performed with the DZ [279] and cc-pVDZ [237] basis sets. For the smaller DZ basis set, the results are shown in Table 3.2 and Figure 3.2. In this case, the equilibrium geometry, abbreviated as  $R_{\text{O-H}} =$  $R_e$ , and corresponding full CI energy were taken from [280]. The geometries that represent a simultaneous stretching of both O–H bonds by a factor of 1.5  $(R_{\text{O-H}} = 1.5R_e)$  and 2  $(R_{\text{O-H}}$  $= 2R_e$ ) without changing the  $\angle$ (H–O–H) angle and the corresponding full CI energies were taken from Ref. [281]. The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t;3), and CC(t,q;3,4) calculations included the  $3a_1$ ,  $1b_2$ ,  $4a_1$ , and  $2b_2$  orbitals, which are the two highest occupied and the two lowest unoccupied orbitals at  $R_{\text{O-H}} = 2R_e$ . The CCSD(T) and  $CCSD(TQ_f)$  energies can be found, for example, in Ref. [141], the CCSDT energies in Ref. [120], the CCSDtq results in Ref. [241], and the CCSDTQ results in Ref. [124], although all of these values were recalculated by us. For the larger cc-pVDZ basis set, the results are shown in Table 3.3 and Figure 3.3. In this case, the corresponding geometries covering the  $R_{\text{O-H}} = R_e - 3R_e$  region and the full CI energies were taken from Ref. [282]. Again, the active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations included the valence  $3a_1$ ,  $1b_2$ ,  $4a_1$ , and  $2b_2$  orbitals. In all calculations, all electrons were correlated and for the cc-pVDZ basis set spherical components of the d orbitals were employed. In addition to the energies or, to be more precise, errors relative to full CCSDT, full CCSDTQ, and full CI at each geometry, Tables 3.2 and 3.3 also report the overall maximum unsigned error (MUE) and non-parallelity error (NPE) values characterizing the various methods.

Starting with the DZ basis set (see Table 3.2 and Figure 3.2), we can see that the noniterative CCSD(T) method performs well at equilibrium, reproducing the parent CCSDT value to a small fraction of a millihartree. However, this situation changes drastically as the O–H bonds are stretched. By the time we reach  $R_{\text{O-H}} = 2R_e$  the error in the CCSD(T) calculations is about 5.5 millihartree compared to CCSDT, diverging quickly from it. When CCSD(T) is compared to full CI, the error is even greater (7.699 millihartree). As for CR-CC(2,3), the A variant performs similar to CCSD(T) at the equilibrium, and also diverges from CCSDT as both O–H bonds are stretched, but the errors are not as severe as in the case of CCSD(T), with CR-CC(2,3)<sub>A</sub> deviating from CCSDT by about 2.7 millihartree at  $R_{O-H}$  $= 2R_e$ . The D variant of CR-CC(2,3) greatly improves all of the results, producing small errors relative to full CCSDT that do not exceed 0.3 millihartree. Unfortunately, the CCSDT method itself begins to break in the  $R_{\text{O-H}} = 2R_e$  region, i.e., we have to consider methods with a more robust treatment of quadruples to obtain further improvement. The activespace CCSDt method compared to CCSD(T),  $CR-CC(2,3)_A$ , and  $CR-CC(2,3)_D$ , provides a smaller NPE value of 0.056 millihartree relative to CCSDT, but this is a result of error cancellation at the various geometries, since errors in the CCSDt results relative to CCSDT, of 0.583–0.639 millihartree, are not small, indicating some missing triples correlations. The agreement of CCSDt with CCSDT can be methodically improved by increasing the active space in the former case, but in this work we choose the alternative, less expensive approach, based on correcting the CCSDt energies for the correlation effects due to the missing triples using the CC(t;3) approach.

When we correct the active-space CCSDt energies for the missing triples using the CC(t;3) method, we see a dramatic improvement over both the active-space and completely renormalized approaches. The  $CC(t;3)_A$  and  $CC(t;3)_D$  methods have maximum errors (MUE values) of 0.295 and 0.212 millihartree, respectively, when compared to CCSDT. That is a major improvement over the maximum errors of 5.489 and 2.717 millihartree relative to full CCSDT observed in the CCSD(T) and the CR-CC(2,3)<sub>A</sub> calculations and a small im-

provement over CR-CC(2,3)<sub>D</sub> which gives MUE = 0.293 millihartree. At the same time, the CC(t;3) corrections are capable of improving the CCSDt results, reducing the MUE value characterizing CCSDt, of 0.639 millihartree, to a 0.2-0.3 millihartree level. In addition to the lower maximum errors, both variants of CC(t;3) improve the results of CR-CC(2,3) and CCSDt calculations for all three geometries almost perfectly, reproducing the full CCSDT energies at each point on the potential energy surface. Unfortunately, as already pointed out above and as shown in Figure 3.2 (top panel), CCSDT itself is insufficient and begins to fail as we approach the  $R_{\rm O-H} = 2R_e$  region. So the fact that CC(t;3) agrees with it almost perfectly is not solving the problem. We need to incorporate connected quadruply excited clusters and do it in a robust manner to obtain an accurate description of double bond dissociation in water.

For the three geometries examined in the DZ case, we can see that full CCSDTQ is essentially exact, with a maximum error of only 0.141 millihartree relative to full CI (as opposed to 2.210 millihartree in the CCSDT case at  $R_{\text{O-H}} = 2R_e$ ). The CCSD(TQ<sub>f</sub>) approach works well at the equilibrium, but, in the same fashion as CCSD(T), it diverges from full CCSDTQ and full CI as the O–H bonds are stretched, producing errors of over 6 millihartree relative to CCSDTQ at  $R_{\text{O-H}} = 2R_e$ . The AA variant of CR-CC(2,4), which corresponds to using Møller–Plesset denominators for the triples and quadruples corrections, performs better than CCSD(TQ<sub>f</sub>), with a maximum error of just over 1 millihartree relative to CCSDTQ, and CR-CC(2,4)<sub>DA</sub> works in a similar way, but none of the CR-CC(2,4) approaches can compete with the active-space CCSDtq method when it comes to the NPE values. In analogy to the case of CCSDt vs CCSDT, the CCSDtq calculations are characterized by the very small NPE values relative to the parent CCSDTQ approach, of 0.259 millihartree, as opposed to NPEs of over 1 millihartree observed in the CR-CC(2,4) calculations. The active-space CCSDtq method provides consistent errors relative to CCSDTQ that range from 0.970 millihartree at  $R_{\text{O-H}} = R_e$  to 1.229 millihartree at  $R_{\text{O-H}} = 2R_e$ , but some triples and quadruples correlations are clearly missing. Instead of simply increasing the active-space to improve these results, we correct the CCSDtq energies for the missing higher-order correlation effects by employing the CC(P;Q) framework.

When we correct the CCSDtq results for just the missing triples via the CC(t,q;3) approach, we reduce the maximum error relative to full CCSDTQ of 1.229 millihartree to 0.888 and 0.805 millihartree when the  $CC(t,q;3)_A$  and  $CC(t,q;3)_D$  methods are employed. The CCSDtq-based  $CC(t,q;3)_A$  and  $CC(t,q;3)_D$  methods also have smaller maximum and non-parallelity errors than the CCSD-based CR-CC(2,4)<sub>AA</sub> and CR-CC(2,4)<sub>DA</sub> approximations, which shows that the use of better  $T_1$  and  $T_2$  amplitudes constructing the MMCC-style corrections helps the overall accuracy. When we correct the CCSDtq values for the missing triples and quadruples, we further improve our agreement with CCSDTQ. For the three geometries, the MUE values relative to full CCSDTQ characterizing the  $CC(t,q;3,4)_{AA}$  and  $CC(t,q;3,4)_{DA}$  calculations are only 0.412 and 0.329 millihartree, respectively, and the corresponding NPEs are equally small. We can, thus, conclude that for the DZ basis set, the CC(t,q;3) and CC(t,q;3,4) approaches reproduce the parent and virtually exact CCSDTQ results to a fraction of a millihartree, at a tiny fraction of the computational cost as we go from the iterative  $\mathcal{N}^{10}$  CPU steps of CCSDTQ to  $\mathcal{N}^6$ -like iterative steps of CCSDtq plus the relatively inexpensive  $\mathcal{N}^7$ -like steps associated with the non-iterative CC(t,q;3) and CC(t,q;3,4) corrections.

The larger cc-pVDZ basis set also allows us to examine how well the various CC methods perform as we further stretch the O-H bond lengths to  $2.5R_e$  and  $3R_e$  (see Table 3.3 and Figure Figure 3.3. Once again, as we break both bonds, the CCSD(T) method quickly diverges from full CCSDT, producing huge 50 millihartree error relative to CCSDT at  $R_{\text{O-H}} = 3R_e$ . The CR-CC(2,3)<sub>A</sub> and CR-CC(2,3)<sub>D</sub> methods, especially CR-CC(2,3)<sub>D</sub>, are much more well behaved, with maximum errors of 8.922 and 1.652 millihartree, respectively, relative to CCSDT, but this is not too helpful, since CCSDT itself completely fails when  $R_{\text{O-H}} > 2R_e$ , giving errors relative to full CI that exceed 40 millihartree in the  $R_{\text{O-H}} = 3R_e$  region. As in the case of the DZ basis set, the active-space CCSDt approach reduces the NPE values relative to full CCSDT compared to CR-CC(2,3) calculations, but this is a result of cancelling rather substantial errors that range from 0.860 to 2.216 millihartree when the entire  $R_{\text{O-H}} = R_e - 3R_e$  region is examined, indicating missing triples.

When we correct the active-space CCSDt results for the missing triple excitations that are excluded in CCSDt using the CC(t;3) corrections, the errors at all geometries relative to CCSDT reduce to small fractions of a millihartree on the order of 0.1–0.5 millihartrees, resulting in equally impressive NPE values, but we must keep in mind that CCSDT itself completely fails when both O-H bonds are stretched beyond  $2R_e$ , i.e., we need to incorporate  $T_4$  clusters and do it in a robust manner to obtain a more accurate description. As shown in Table 3.3, up to  $2R_e$ , CCSDTQ is virtually exact, while beginning to deviate from full CI in the  $R_{\text{O-H}} = 2.5R_e - 3R_e$  region, signaling the need for even high-order correlations in order to accurately describe the simultaneous O-H bond-breaking in water, but this deviation is not nearly as big as in the CCSDT case and the point of this investigation is to evaluate how well the CC(P;Q) methods reproduce their parent methods in comparison to previously established approximations, so CCSDTQ remains an important reference in judging other methods including quadruples. As can be seen in Table 3.3 and similarly to the DZ case, the  $CCSD(TQ_f)$  approach describes the equilibrium geometry well compared to CCSDTQ, but diverges as the bonds are stretched, with a maximum error of approximately 14 millihartree at  $R_{\text{O-H}} = 3R_e$  relative to CCSDTQ. From 2.5R<sub>e</sub> to  $3R_e$ , CCSD(TQ<sub>f</sub>) goes from overstabilizing the dissociating water molecule by over 11 millihartree to understabilizing it by nearly 14 millihartree, indicating the massive failure of  $CCSD(TQ_f)$  in the bond breaking region. This erratic behavior of  $CCSD(TQ_f)$  can also be seen in Figure 3.3. The  $CR-CC(2,4)_{AA}$  approach tends to have better agreement with CCSDTQ, but still has a maximum error of over 6 millihartree relative to CCSDTQ, while the  $CR-CC(2,4)_{DA}$  method has a larger maximum error of 13.621 millihartree relative to CCSDTQ, even though it seems to be stabilizing at the longer bond lengths with respect to full CCSDTQ. Unlike the above non-iterative approaches based on the CCSD amplitudes, the active-space CCSDtq results do not exhibit the divergent behavior, giving rather stable energies that are 1.497–2.656 millihartree higher in energy than their CCSDTQ counterparts, but it is quite clear from Table 3.3 that CCSDtq is still missing some dynamical correlations of the  $T_3$  and  $T_4$  type, which one would like to capture. We can always try to do it by increasing the active space, but here we advocate a more appealing approach of capturing the missing correlations via the CC(t,q;3) and CC(t,q;3,4) corrections.

When the CCSDtq results are corrected for just the missing triples via the  $CC(t,q;3)_A$  and  $CC(t,q;3)_D$  approaches, we reduce the maximum error relative to CCSDTQ characterizing the CCSDtq calculations of 2.656 millihartree, to much smaller 1.349 and 1.293 millihartree errors. This clearly shows that the CC(P;Q) corrections can be very effective, while demonstrating that using  $T_1$  and  $T_2$  amplitudes relaxed in the presence of triples and quadruples is a lot better than using their unrelaxed CCSD counterparts exploited in CR-CC(2,4) calculations. If the CCSDtq energies are corrected for missing triples and quadruples with the CC(t,q;3,4) approach, we improve our agreement with CCSDTQ even further. Both the  $CC(t,q;3,4)_{AA}$  and  $CC(t,q;3,4)_{DA}$  methods have maximum errors relative to full CCS-

DTQ of just over 1 millihartree and the corresponding NPE values are equally small. If we constrain ourselves to the  $R_{\text{O-H}} = R_e - 2R_e$  region, the agreement between CC(t,q;3,4) and CCSDTQ calculations is even better, resulting in impressively small differences that do not exceed 0.2–0.4 millihartree. Our cc-pVDZ calculations confirm that the CC(t,q;3) and CC(t,q;3,4) approaches can reproduce the CCSDTQ energies for the double-dissociation of water to within a millihartree or better at the small fraction of the computer costs.

Table 3.2: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCS-DTQ counterparts and the corresponding full CI data for the equilibrium and two displaced geometries of the  $H_2O$  molecule, as described by the DZ basis set [279].

Method	$R_{\text{O-H}} = R_e{}^a$	$R_{\text{O-H}} = 1.5 R_e{}^b$	$R_{\text{O-H}} = 2R_e{}^b$	MUE	NPE
Full $CI^c$	-76.157866	-76.014521	-75.905247		
		CC methods wi	th triples <sup><math>d</math></sup>		
CCSD(T)	0.140(0.574)	-0.008(1.465)	-5.489(-7.699)	5.489(7.699)	5.629(9.164)
$CR-CC(2,3)_A$	0.170(0.604)	0.784(2.257)	2.717(0.507)	2.717(2.257)	2.547(1.750)
$CR-CC(2,3)_D$	-0.140(0.294)	-0.293(1.180)	-0.195(-2.405)	0.293(2.405)	0.153(3.585)
$\mathrm{CCSDt}^{e}$	0.594(1.028)	0.583(2.056)	0.639(-1.571)	0.639(2.056)	0.056(3.627)
$CC(t;3)_A e$	0.073(0.507)	0.105(1.578)	0.295(-1.915)	0.295(1.915)	0.222(3.493)
$CC(t;3)_D e$	-0.066(0.368)	-0.011(1.462)	0.212(-1.998)	0.212(1.998)	0.278(3.460)
CCSDT	0.000(0.434)	0.000(1.473)	0.000(-2.210)	0.000(2.210)	0.000(3.683)
	CC met	hods with triple	es and quadrup	$\log^{f}$	
$\rm CCSD(TQ_f)$	0.151(0.166)	-0.047(0.094)	-6.022(-5.914)	6.022(5.914)	6.173(6.080)
$CR-CC(2,4)_{AA}$	0.041(0.056)	0.238(0.379)	-1.039(-0.931)	1.039(0.931)	1.277(1.310)
$CR-CC(2,4)_{DA}$	-0.269(-0.254)	-0.839(-0.698)	-3.951(-3.843)	3.951(3.843)	3.682(3.589)
$\mathrm{CCSDtq}^e$	0.970(0.985)	1.093(1.234)	1.229(1.337)	1.229(1.337)	0.259(0.352)
$CC(t,q;3)_A{}^e$	0.449(0.464)	0.614(0.755)	0.888(0.996)	0.888(0.996)	0.439(0.532)
$CC(t,q;3)_D^e$	0.309(0.324)	0.497(0.638)	0.805(0.913)	0.805(0.913)	0.496(0.589)
$CC(t,q;3,4)_{AA}^{e}$	0.022(0.037)	0.049(0.190)	0.412(0.520)	0.412(0.520)	0.390(0.483)
$CC(t,q;3,4)_{DA}^{e}$	-0.118(-0.103)	-0.068(0.073)	0.329(0.437)	0.329(0.437)	0.447(0.540)
$\mathbf{CCSDTQ}$	0.000(0.015)	0.000(0.141)	0.000(0.108)	0.000(0.141)	0.000(0.126)

 $^{a}$  The equilibrium geometry and the corresponding full CI result were taken from Ref. [280].

<sup>b</sup> The geometries that represent a simultaneous stretching of both O–H bonds by factors of 1.5 and 2.0 without changing the  $\angle$ (H–O–H) angle and the corresponding full CI results were taken from Ref. [281].

 $^{c}$  The total full CI energies in hartree.

 $^{d}$  For the CC methods with up to triple excitations, the reported energy values, in millihartree, are errors relative to full CCSDT and, in parentheses, relative to full CI.

<sup>e</sup> The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of the  $3a_1$  and  $1b_2$  occupied and  $4a_1$  and  $2b_2$  unoccupied orbitals.

<sup>f</sup> For the CC methods with triples and quadruples, the reported energy values, in millihartree, are errors relative to full CCSDTQ and, in parentheses, relative to full CI.



Figure 3.2: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations, along with their parent full CCSDT and CCSDTQ counterparts, for the equilibrium and two displaced geometries of the  $H_2O$ molecule, as described by the DZ basis set [279]. The numerical values of the errors are found in Table 3.2. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

Table 3.3: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts and the corresponding full CI data for the equilibrium and four displaced geometries of the H<sub>2</sub>O molecule, as described by the spherical cc-pVDZ basis set [237].

Method	$R_{\text{O-H}} = R_e{}^a$	$R_{\text{O-H}} = 1.5 R_e{}^a$	$R_{\rm O-H} = 2R_e{}^a$	$R_{\text{O-H}} = 2.5 R_e{}^a$	$R_{\text{O-H}} = 3R_e{}^a$	MUE	NPE
Full $CI^b$	-76.241860	-76.072348	-75.951665	-75.917991	-75.911946		
			CC method	s with triples <sup><math>c</math></sup>			
$\operatorname{CCSD}(\mathrm{T})$	0.165(0.658)	0.208(1.631)	-2.415(-3.820)	-17.812(-42.564)	-50.386(-90.512)	50.386(90.512)	50.594(92.143)
$CR-CC(2,3)_A$	0.413(0.906)	1.402(2.825)	5.210(3.805)	8.922(-15.830)	7.091(-33.035)	8.922(33.035)	8.509(36.840)
$CR-CC(2,3)_D$	-0.149(0.344)	-0.281(1.142)	0.854(-0.551)	1.652(-23.100)	-0.430(-40.556)	1.652(40.556)	2.082(41.698)
$\mathrm{CCSDt}^d$	2.216(2.709)	1.690(3.113)	1.027(-0.378)	0.860(-23.892)	0.925(-39.201)	2.216(39.201)	1.356(42.314)
$CC(t;3)_A{}^d$	0.261(0.754)	0.130(1.553)	0.168(-1.237)	0.411(-24.341)	0.528(-39.598)	0.528(39.598)	0.398(41.151)
$CC(t;3)_D{}^d$	-0.135(0.358)	-0.159(1.264)	0.023(-1.382)	0.337(-24.415)	0.460(-39.666)	0.460(39.666)	0.619(40.930)
CCSDT	0.000(0.493)	0.000(1.423)	0.000(-1.405)	0.000(-24.752)	0.000(-40.126)	0.000(40.126)	0.000(41.549)
		CC 1	methods with t	riples and quadru	$ples^e$		
$\mathrm{CCSD}(\mathrm{TQ}_{\mathrm{f}})$	0.210(0.229)	0.381(0.502)	-3.933(-3.903)	-11.315(-13.676)	13.957(9.224)	13.957(13.676)	25.272(22.900)
$CR-CC(2,4)_{AA}$	0.249(0.268)	0.779(0.900)	0.729(0.759)	-5.402(-7.763)	-6.101(-10.834)	6.101(10.834)	6.880(11.734)
$CR-CC(2,4)_{DA}$	-0.312(-0.293)	-0.904(-0.783)	-3.627(-3.597)	-12.672(-15.033)	-13.621(-18.354)	13.621(18.354)	13.309(18.061)
$\mathrm{CCSDtq}^d$	2.656(2.675)	2.123(2.244)	1.497(1.527)	1.627(-0.734)	1.653(-3.080)	2.656(3.080)	1.159(5.755)
$\mathrm{CC}(\mathrm{t},\mathrm{q};3)_{\mathrm{A}}{}^{d}$	0.701(0.720)	0.561(0.682)	0.636(0.666)	1.241(-1.120)	1.349(-3.384)	1.349(3.384)	0.788(4.104)
$CC(t,q;3)_D{}^d$	0.305(0.324)	0.271(0.392)	0.491(0.521)	1.177(-1.184)	1.293(-3.440)	1.293(3.440)	1.022(3.961)
$CC(t,q;3,4)_{AA}^{d}$	0.142(0.161)	-0.070(0.051)	0.187(0.217)	0.947(-1.414)	1.104(-3.629)	1.104(3.629)	1.174(3.846)
$CC(t,q;3,4)_{DA}$ <sup>d</sup>	-0.254(-0.235)	-0.360(-0.239)	0.042(0.072)	0.883(-1.478)	1.048(-3.685)	1.048(3.685)	1.408(3.757)
$\mathbf{CCSDTQ}$	0.000(0.019)	0.000(0.121)	0.000(0.030)	0.000(-2.361)	0.000(-4.733)	0.000(4.733)	0.000(4.854)

<sup>*a*</sup> The equilibrium geometry,  $R_{\text{O-H}} = R_e$ , the geometries that represent a simultaneous stretching of both O–H bonds by factors of 1.5, 2.0, 2.5, and 3.0 without changing the  $\angle$ (H–O–H) angle, and the corresponding full CI results were taken from Ref. [282].

<sup>b</sup> The total full CI energies in hartree.

 $^{c}$  For the CC methods with up to triple excitations, the reported energy values, in millihartree, are errors relative to full CCSDT and, in parentheses, relative to full CI.

<sup>d</sup> The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of the  $3a_1$  and  $1b_2$  occupied and  $4a_1$  and  $2b_2$  unoccupied orbitals.

 $^{e}$  For the CC methods with triples and quadruples, the reported energy values, in millihartree, are errors relative to full CCSDTQ and, in parentheses, relative to full CI.



Figure 3.3: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the equilibrium and four displaced geometries of the  $H_2O$  molecule, as described by the spherical cc-pVDZ basis set [237]. The numerical values of the errors are found in Table 3.3. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results.

## 3.2.2 Be + H<sub>2</sub> $\rightarrow$ HBeH $C_{2v}$ Insertion Pathway

The insertion of Be into  $H_2$  is another example where one has to incorporate triply and quadruply excited clusters in a robust manner to obtain an accurate description, i.e., the case well suited for testing the CC(P;Q)-based CC(t,q;3) and CC(t,q;3,4) approaches developed in this thesis. We examined nine points along the  $C_{2v}$ -symmetric Be + H<sub>2</sub>  $\rightarrow$  HBeH insertion pathway, labeled A–I, which were taken from Ref. [283]. The geometries are shown in Table 3.4. Structure A corresponds to the linear HBeH product, while structure I corresponds to the Be + H<sub>2</sub> reactants separated by 6 bohr. Structure E is the transition state, which is described in more detail below. Calculations were performed using the [3s1p/2s] basis set described in Ref. [283] and all electrons were correlated. In order to have access to more decimal places for an accurate error analysis, the corresponding full CI energies were taken from Ref. [48], since the full CI energies reported in in Ref. [283] show only four decimal places, as compared to the six decimal places provided in Ref. [48]. The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of the highest occupied and lowest unoccupied 1b<sub>2</sub> and 3a<sub>1</sub> orbitals.

For the structures A–E, the lowest-energy RHF reference solution is provided by the  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  determinant, while for structures F–I, the lowest-energy RHF solution is given by the  $|(1a_1)^2(2a_1)^2(3a_1)^2|$  determinant, obtained by replacing the highest occupied orbital,  $1b_2$ , by the lowest unoccupied orbital,  $3a_1$ . This is why our choice of active space for the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consists of the  $1b_2$  and  $3a_1$  orbitals. In the full CI wave function expansion, the A–D region is dominated by the  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  configuration, while the F–I region is dominated by the  $|(1a_1)^2(2a_1)^2(3a_1)^2|$  configuration. Structure E represents the transition state where both

configurations become quasi-degenerate, which makes this structure the hardest one to describe. We, thus, used the  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  configuration, further referred to as reference  $|\Phi_1\rangle$  as a reference determinant for structures A–D and the  $|(1a_1)^2(2a_1)^2(3a_1)^2|$  configuration, further referred to as refrence  $|\Phi_2\rangle$ , for structures F–I. To examine the sensitivity of the various SRCC results on the choice of the reference determinant, we performed two sets of calculations for structure E, one using  $|\Phi_1\rangle$  as a reference and another using  $|\Phi_2\rangle$  as a reference. The results of our calculations are given in Table 3.5 and Figure 3.4.

The transition state structure E is the hardest one to describe, exhibiting larger errors than any other geometry for all methods examined, so much of our discussion will focus on it. Although, as shown in Table 3.5 and Figure 3.4, one needs to include quadruples in the SRCC calculations to obtain accurate results, since even CCSDT is insufficient, we begin our discussion with a comparison of the various approximate triples treatments with CCSDT, to see how faithful they can be in reproducing the full CCSDT state. The CCSD(T)scheme struggles at describing the energetics of structure E, producing a large error of 3.569 millihartree relative to CCSDT when determinant  $|\Phi_1\rangle$  is used as reference, and a 1.109 millihartree error relative to full CCSDT when the  $|\Phi_2\rangle$  reference is employed. The  $CR-CC(2,3)_A$  approach has even larger errors relative to CCSDT, but the more complete CR-CC(2,3) treatment, represented by variant D, works better, reducing the 3.569 and 1.109 millihartree errors relative to CCSDT in the  $CCSD(T)/|\Phi_1\rangle$  and  $CCSD(T)/|\Phi_2\rangle$  calculations of the E geometry to 3.305 and 0.824 millihartree, respectively, making the resulting energies somewhat less dependent on the choice of the reference determinant, when compared to full CI. When the  $|\Phi_1\rangle$  reference is employed, the active-space CCSDt approach has a smaller error relative to CCSDT at the E geometry than any of the non-iterative corrections to CCSD and CCSDt competes with the CR-CC(2,3)<sub>D</sub> when the  $|\Phi_2\rangle$  reference is employed, retaining the *ca.* 1 millihartree error relative to CCSDT at the E structure, improving CCSD(T). As for the remaining geometries, the active-space CCSDt calculations tend to give errors relative to CCSDT that are too large, or at the very least similar to the CCSD-based non-iterative triples methods, indicating that some important triples are still missing. Instead of simply increasing the active space, which is one way to go about it, we correct the CCSDt energies for the missing triple excitations using the CC(P;Q)-based CC(t;3) approach.

When the CCSDt energies are corrected for the correlation effects due to the missing triples using the CC(t;3) methodology, the agreement with CCSDT is generally very good. With the exception of structure E, the differences between the  $CC(t;3)_A$  or  $CC(t;3)_D$  and CCSDT energies are on the order of tens of microhartree. When the transition state structure E is examined and reference  $|\Phi_1\rangle$  is employed, the errors in the  $CC(t;3)_A$  and  $CC(t;3)_D$ results relative to CCSDT are 1.469 and 1.199 millihartree, respectively, which is a significant improvement over the 2-5 millihartree error values given by the CCSD(T), CR-CC(2,3), and CCSDt methods. When reference  $|\Phi_1\rangle$  is adopted, the agreement of  $CC(t;3)_A$  or  $CC(t;3)_D$ with CCSDT at structure E is even better, with errors residing in the 0.3–0.5 millihartree range. The  $CC(t;3)_A$  and  $CC(t;3)_D$  approaches also provide substantial improvements in the NPE values relative to CCSDT, compared to CCSD(T), CR-CC(2,3), and CCSDt data, so they are more systematic at describing the triples along the reaction pathway than the standard CCSD-based non-iterative and active-space iterative CC methods. This is, however, not sufficient, since one cannot obtain an accurate description of the Be + H\_2  $\rightarrow$  HBeH reaction without quadruples, especially in the region of the transition state structure E. As shown in Table 3.5 and Figure 3.4, even full CCSDT fails, giving errors relative to full CI that exceed 2 millihartree. One needs connected quadruple excitation to address this problem.

By comparing the CCSDTQ and full CI energies at parts A–I, we can see that full

CCSDTQ is an essentially exact theory in this case, reducing the > 2 millihartree errors in the CCSDT results at structure E to 16 microhartree or less (see Table 3.5 and Figure 3.4). The question then is how accurate various approximate treatments of triples and quadruples are compared to CCSDTQ. It is quite clear from Table 3.5 and Figure 3.4 that traditional perturbative corrections to CCSD, represented here by the  $CCSD(TQ_f)$  approximation, fail, giving errors relative to CCSDTQ and full CI at the transition state structure E that exceed 4.1 millihartree when the  $|\Phi_1\rangle$  is used as a reference or 1.7 millihartree when the  $|\Phi_2\rangle$  reference is employed. The CR-CC(2,4) approaches, especially CR-CC(2,4)<sub>DA</sub>, help bringing the above error to a 1–3 millihartree level, but none of the non-iterative corrections due to triples and quadruples to CCSD work well. The iterative active-space CCSDtq method provides a more robust description, reducing the 1-3 millihartree errors in the CR-CC(2,4) results relative to full CCSDTQ at the E geometry to 1.430 millihartree, when the  $|\Phi_1\rangle$  reference is employed, and 0.908 millihartree, when  $|\Phi_2\rangle$  is used as a reference, reducing the overall NPE values at the same time, but this is done at the expense of losing accuracy in the reactant (structures F–I) and product (structures A–D) regions, where errors in the CCSDtq description can be as high as  $\sim 0.4$  millihartree, since CCSDtq neglects certain classes of dynamical triples and quadruples that a small active space used here cannot capture. It becomes, therefore, important to bring the missing higher-order excitations via the CC(P;Q)-based CC(t,q;3)and CC(t,q;3,4) corrections.

When the CCSDtq energies are corrected for the missing triples via the  $CC(t,q;3)_A$  and  $CC(t,q;3)_D$  approaches, the 0.9–1.4 millihartree errors in the CCSDtq results for structure E relative to CCSDTQ reduce to 0.5–0.8 millihartree when the  $|\Phi_1\rangle$  determinant is used as a reference and to 0.2–0.4 millihartree when the  $|\Phi_2\rangle$  reference is employed. The  $CC(t,q;3)_D$  results, which give 0.2–0.5 millihartree errors, are particularly impressive. For the remaining

geometries A–D and F–I, the CC(t,q;3)<sub>A</sub> approach has a maximum error relative to CCS-DTQ of 0.133 millihartree, while the CC(t,q;3)<sub>D</sub> approximation works even better, giving a maximum error of only 37 microhartree. Very little changes when the CCSDtq energies are corrected for the missing triple as well as quadruple excitations via CC(t,q;3,4)<sub>AA</sub> and CC(t,q;3,4)<sub>DA</sub> approaches, i.e., the CC(t,q;3,4)<sub>AA</sub> and CC(t,q;3,4)<sub>DA</sub> energies are as accurate as their CC(t,q;3)<sub>A</sub> and CC(t,q;3)<sub>D</sub> counterparts, which indicates that it is sufficient to correct CCSDtq energies for the triples missing in CCSDtq in this case, but we must keep in mind that the CC(t,q;3) results, especially those obtained with variant D, are already outstanding. We can, thus, conclude that the CC(P;Q)-based CC(t,q;3) and CC(t,q;3,4) corrections to CCSDtq, especially CC(t,q;3)<sub>D</sub> and CC(t,q;3,4)<sub>DA</sub>, are capable of reproducing the virtually exact full CCSDTQ and exact full CI data for the Be + H<sub>2</sub>  $\rightarrow$  HBeH reaction to within small fractions of a millihartree at the tiny fraction of the computer effort, which is a promising finding for the future applications of the CC(P;Q) formalism.

Table 3.4: Coordinates of points along the sampling path  $C_{2v}$ -symmetric path describing insertion of Be into H<sub>2</sub>, introduced in Ref. [283]. Structure A is the linear HBeH product, while structure I represents the Be + H<sub>2</sub> reactant. Structure E is the transition state.

Point	Coordinates for $H_2$ (X, Y, Z) <sup><i>a</i></sup>
А	$(0.0, \pm 2.54, 0.0)$
В	$(0.0, \pm 2.08, 1.0)$
$\mathbf{C}$	$(0.0, \pm 1.62, 2.0)$
D	$(0.0, \pm 1.39, 2.5)$
Ε	$(0.0, \pm 1.275, 2.75)$
$\mathbf{F}$	$(0.0, \pm 1.16, 3.0)$
G	$(0.0, \pm 0.93, 3.5)$
Н	$(0.0, \pm 0.70, 4.0)$
Ι	$(0.0, \pm 0.70, 6.0)$

<sup>a</sup> Be is located at (0.0, 0.0, 0.0) and all values are in bohr.

Table 3.5: A comparison of the energies resulting from various all-electron CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCS-DTQ counterparts, and the corresponding full CI data for the geometries A–I defining the  $C_{2v}$ -symmetric insertion pathway of Be into H<sub>2</sub>, as described by the [3s1p/2s] basis set, introduced in Ref. [283].<sup>*a*</sup>

	$ \Phi\rangle =  \Phi_1\rangle =  (1a_1)^2 (2a_1)^2 (1b_2)^2 ^b$						
Method	A	В	С	D	Е		
Full $CI^d$	-15.779172	-15.737224	-15.674818	-15.622883	-15.602919		
		CC 1	methods with	$triples^{e}$			
$\operatorname{CCSD}(\mathrm{T})$	0.148(0.169)	0.137(0.140)	0.190(0.149)	0.496(0.392)	3.569(1.195)		
$CR-CC(2,3)_A$	0.150(0.171)	0.140(0.143)	0.188(0.147)	0.547(0.443)	4.828(2.454)		
$CR-CC(2,3)_D$	0.043(0.064)	0.035(0.038)	0.044(0.003)	0.236(0.132)	3.305(0.931)		
$\mathrm{CCSDt}^f$	0.227(0.248)	0.187(0.190)	0.236(0.195)	0.470(0.366)	2.096(-0.278)		
$CC(t;3)_A{}^f$	0.086(0.107)	0.070(0.073)	0.094(0.053)	0.236(0.132)	1.469(-0.905)		
$CC(t;3)_D{}^f$	0.022(0.043)	0.018(0.021)	0.033(-0.008)	0.138(0.034)	1.199(-1.175)		
CCSDT	0.000(0.021)	0.000(0.003)	0.000(-0.041)	0.000(-0.104)	0.000(-2.374)		
		CC methods	with triples a	nd quadruples	$_{3}^{g}$		
$\mathrm{CCSD}(\mathrm{TQ}_{\mathrm{f}})$	0.161(0.161)	0.142(0.141)	0.182(0.181)	0.465(0.464)	-4.104(-4.106)		
$CR-CC(2,4)_{AA}$	0.165(0.165)	0.141(0.140)	0.154(0.153)	0.473(0.472)	3.149(3.147)		
$CR-CC(2,4)_{DA}$	0.058(0.058)	0.036(0.035)	0.010(0.009)	0.162(0.161)	1.626(1.624)		
$\mathrm{CCSDtq}^f$	0.238(0.238)	0.192(0.191)	0.239(0.238)	0.367(0.366)	1.430(1.428)		
$\mathrm{CC}(\mathrm{t},\mathrm{q};3)_{\mathrm{A}}{}^{f}$	0.096(0.096)	0.074(0.073)	0.098(0.097)	0.133(0.132)	0.808(0.806)		
$\mathrm{CC}(\mathrm{t},\mathrm{q};3)_\mathrm{D}{}^f$	0.033(0.033)	0.022(0.021)	0.037(0.036)	0.035(0.034)	0.539(0.537)		
$CC(t,q;3,4)_{AA}^{f}$	0.093(0.093)	0.073(0.072)	0.096(0.095)	0.114(0.113)	0.922(0.920)		
$CC(t,q;3,4)_{DA}^{f}$	0.030(0.030)	0.021(0.020)	0.035(0.034)	0.016(0.015)	0.653(0.651)		
$\mathbf{CCSDTQ}$	0.000(0.000)	0.000(-0.001)	0.000(-0.001)	0.000(-0.001)	0.000(-0.002)		

Table 3.5 (cont'd)

	$ \Phi angle= \Phi_2 angle$	$=  (1a_1)^2 (2a_1) ^2$	$)^{2}(3a_{1})^{2} ^{c}$			
${ m E}$	$\mathbf{F}$	G	Н	Ι	MUE	NPE
-15.602919	-15.624981	-15.693194	-15.736688	-15.760878		
	CC m	ethods with tr	$i p les^{e}$			
-1.109(-3.314)	0.166(0.224)	0.058(0.100)	0.034(0.049)	0.005(0.009)	3.569(3.314)	4.678(4.509)
2.191(-0.014)	0.255(0.313)	0.081(0.123)	0.045(0.060)	0.007(0.011)	4.828(2.454)	4.821(2.468)
0.824(-1.381)	0.037(0.095)	-0.012(0.030)	-0.007(0.008)	-0.001(0.003)	3.305(1.381)	3.317(2.312)
0.981(-1.224)	0.308(0.366)	0.175(0.217)	0.109(0.124)	0.021(0.025)	2.096(1.224)	2.075(1.590)
0.494(-1.711)	0.118(0.176)	0.048(0.090)	0.028(0.043)	0.004(0.008)	1.469(1.711)	1.465(1.887)
0.295(-1.910)	0.026(0.084)	-0.013(0.029)	-0.004(0.011)	-0.001(0.003)	1.199(1.910)	1.212(1.994)
0.000(-2.205)	0.000(0.058)	0.000(0.042)	0.000(0.015)	0.000(0.004)	0.000(2.374)	0.000(2.432)
(	CC methods v	with triples and	d quadruples <sup><math>g</math></sup>			
-1.728(-1.728)	0.182(0.198)	0.075(0.074)	0.040(0.039)	0.006(0.006)	4.104(4.106)	4.569(4.570)
-0.737(-0.737)	0.255(0.271)	0.093(0.092)	0.051(0.050)	0.008(0.008)	3.149(3.147)	3.886(3.884)
-0.936(-0.936)	0.037(0.053)	-0.001(-0.002)	-0.002(-0.003)	0.000(0.000)	1.626(1.624)	2.562(2.560)
0.908(0.908)	0.315(0.331)	0.185(0.184)	0.117(0.116)	0.023(0.023)	1.430(1.428)	1.407(1.405)
0.412(0.412)	0.125(0.141)	0.058(0.057)	0.036(0.035)	0.006(0.006)	0.808(0.806)	0.802(0.800)
0.209(0.209)	0.033(0.049)	-0.003(-0.004)	0.004(0.003)	0.002(0.002)	0.539(0.537)	0.542(0.541)
0.465(0.465)	0.127(0.143)	0.054(0.053)	0.033(0.032)	0.005(0.005)	0.922(0.920)	0.917(0.915)
0.263(0.263)	0.035(0.051)	-0.007(-0.008)	0.001(0.000)	0.001(0.001)	0.653(0.651)	0.660(0.659)
0.000(0.000)	0.000(0.016)	0.000(-0.001)	0.000(-0.001)	0.000(0.000)	0.000(0.016)	0.000(0.018)

<sup>*a*</sup> The [3s1p/2s] basis set and the geometries A through I defining the  $C_{2v}$ -symmetric Be + H<sub>2</sub>  $\rightarrow$  HBeH insertion pathway, where structure A corresponds to the HBeH product, structure I to the Be + H<sub>2</sub> reactants separated by 6 bohr, and structure E to the transition state, were taken from Ref. [283]. In order to have access to more decimal places for an accurate error analysis, the corresponding full CI energies were taken from Ref. [48] (the full CI energies in Ref. [283] show four decimal places, as compared to six decimal places provided in Ref. [48]).

<sup>b</sup> The CC calculations were performed using the  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  reference determinant, which is the lowest-energy RHF solution for the geometries A–E. The  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  configuration dominates the full CI wave function expansion in the A–D region, while becoming quasi-degenerate with the  $|(1a_1)^2(2a_1)^2(3a_1)^2|$  determinant at the transition-state structure E from Ref. [283]

<sup>c</sup> The CC calculations were performed using the  $|(1a_1)^2(2a_1)^2(3a_1)^2|$  reference determinant, which is the lowest-energy RHF solution for the geometries F–I. The  $|(1a_1)^2(2a_1)^2(3a_1)^2|$  configuration dominates the full CI wave function expansion in the F–I region, while becoming quasi-degenerate with the  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  determinant at the transition-state structure E from Ref. [283] <sup>d</sup> The total full CI energies in hartree.

<sup>e</sup> For the CC methods with up to triple excitations, the reported energy values, in millihartree, are errors relative to full CCSDT and, in parentheses, relative to full CI.

<sup>f</sup> The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of the  $1b_2$  and  $3a_1$  orbitals, which are occupied and unoccupied, respectively, in  $|\Phi\rangle =$  $|(1a_1)^2(2a_1)^2(1b_2)^2|$  and unoccupied and occupied, respectively, in  $|\Phi\rangle = |(1a_1)^2(2a_1)^2(3a_1)^2|$ . <sup>g</sup> For the CC methods with triples and quadruples, the reported energy values, in millihartree, are

<sup>9</sup> For the CC methods with triples and quadruples, the reported energy values, in millihartree, are errors relative to full CCSDTQ and, in parentheses, relative to full CI.



Figure 3.4: A comparison of the energies resulting from various all-electron CC calculations, including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the geometries A–I defining the  $C_{2v}$ -symmetric insertion pathway of Be into H<sub>2</sub>, as described by the [3s1p/2s] basis set, introduced in Ref. [283]. The numerical values of the errors are found in Table 3.5. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results. In presenting the results for the transition state structure E, we used the same reference  $|\Phi_2\rangle$  as that used in the F–I region.

## 3.2.3 Singlet-Triplet Gap in HFH<sup>-</sup> Along the $D_{\infty h}$ -Symmetric Double-Dissociation Pathway

Our next example is the linear,  $D_{\infty h}$ -symmetric, (HFH)<sup>-</sup> anion, in which both H–F bonds are simultaneously stretched, and which has been used in the literature as a prototype magnetic system, where two paramagnetic centers, each carrying an unpaired spin, represented by the terminal hydrogen atoms, are linked via a polarizable diamagnetic bridge constituted by F<sup>-</sup> [284]. The spins of the paramagnetic electrons of the H atoms can be parallel or antiparallel, yielding two different spin states, namely, a singlet,  $X^{1}\Sigma_{g}^{+}$ , which is the ground state, and a triplet,  $A^{3}\Sigma_{u}^{+}$ , which is the first excited state. The total electronic energies of these two states and the gap between them, which should approach zero as both H–F bonds are stretched to infinity and which provides information about the magnetic exchange coupling constant J as a function of the H–F distance  $R_{\rm H-F}$ , has been studied using a variety of *ab initio* and density functional theory methods in Refs. [162, 165, 260, 284, 285]. This includes our calculations reported in Ref. [164] and [165]. where we used a wide variety of SRCC methods with up to triple excitations, including CR-CC(2,3) [164] and CC(t;3) [165] comparing the results with CCSDT and full CI. Some of these results, which are relevant to this work, are restated here (see Tables 3.6–3.8).

As shown in Tables 3.6–3.8 and Figures 3.5–3.7, the main challenge for the CC theory, if we are to rely on the SRCC formalism and utilize the spin- and symmetry-adapted RHF (the  $X \ ^{1}\Sigma_{g}^{+}$  state) and ROHF (the  $A \ ^{3}\Sigma_{u}^{+}$  state) references, is the accurate inclusion of higher-than-doubly excited clusters, especially for the  $X \ ^{1}\Sigma_{g}^{+}$  state, which has a manifestly MR character involving the doubly excited determinant corresponding to excitations from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO) (i.e.,  $(\text{HOMO})^2 \rightarrow (\text{LUMO})^2 (\sigma_g^2 \rightarrow \sigma_u^2)$ ), in addition to the RHF configuration. As demonstrated in Ref. [162], the ratio of the full CI expansion coefficients at the  $(HOMO)^2 \rightarrow$  $(LUMO)^2$  and RHF configuration state functions characterizing the  $X \ ^1\Sigma_g^+$  state, which is equivalent to the  $T_2$  cluster amplitude corresponding to the  $(HOMO)^2 \rightarrow (LUMO)^2$  double excitation, since HOMO and LUMO have different symmetries, increases, in absolute value, from 0.38 to 1.17 as  $R_{\rm H-F}$  is varied between 1.5 and 4.0 Å. Thus, the moderately biradical (HFH)<sup>-</sup> system at shorter H–F separations becomes a strong biradical species at larger H–F distances. Because of the significant biradical character of the (HFH)<sup>-</sup> ion at almost all H–F separations shown in Tables 3.6–3.8 and Figures 3.5–3.7, the  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap is already quite small and sensitive to the electron correlation treatment used in the determination of the  $X \ ^1\Sigma_g^+$  and  $A \ ^3\Sigma_u^+$  states in the region of shorter H–F distances, while rapidly decreasing as  $R_{\rm H-F}$  becomes larger, causing troubles to the standard CCSD and CCSD(T) approaches, which are incapable of handling such a challenging situation. As we will discuss it below, as as seen in Tables 3.6–3.8 and Figures 3.5–3.7, the full CCSDT method is a lot more robust, but it is still not accurate enough to provide a fully quantitative description. The full CCSDTQ approach solves the problem, but CCSDTQ is very expensive, so it is important to examine that various approximate treatments of triples and quadruples can do in this regard.

Following the earlier work [162, 165, 260, 284, 285], we employed the 6-31G(d,p) basis set [286, 287] and sampled several values of the H–F distance  $R_{\rm H-F}$  defining the linear  $D_{\infty h}$ symmetric (HFH)<sup>-</sup> system ranging from  $R_{\rm H-F} = 1.5$  Å to  $R_{\rm H-F} = 4.0$  Å. The full CI energies were taken from Ref. [162], but they are also available in Refs. [165, 260, 285]. The CCSD(T), CR-CC(2,3)<sub>A</sub>, CR-CC(2,3)<sub>D</sub>, and CCSDT values can be found in Refs. [162, 165, 260, 285], while the CCSDt, CR-CC(2,3)<sub>A</sub>, and CR-CC(2,3)<sub>D</sub> values can be found in Ref. [165]. The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of two active electrons and two active orbitals corresponding to HOMO and LUMO (the  $\sigma_g$  and  $\sigma_u$  valence orbitals). Following Refs. [162,165,285], the lowest-energy core orbital was frozen in the post-SCF calculations. The results for the  $X \ ^1\Sigma_g^+$  state can be found in Table 3.6 and Figure 3.5. The results for the  $A \ ^3\Sigma_u^+$  state in are shown in Table 3.7 and Figure 3.6, and the  $A \ ^3\Sigma_u^+ - X \ ^1\Sigma_g^+$  gap calculations are summarized in Table 3.8 and Figure 3.7.

As in Sections 3.2.1 and 3.2.2, although our focus here is on the CC methods with triples and quadruples, we begin out discussion with methods truncated at triples. The most popular SRCC approach with an approximate treatment of triple excitations, CCSD(T), which uses arguments originating from MBPT to estimate the  $T_3$  effects, is incapable of handling the  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap in (HFH)<sup>-</sup>. The description of the  $X^{1}\Sigma_{g}^{+}$  state is where CCSD(T) displays a catastrophic failure, as seen in the errors relative to CCSDT, which grow from 0.435 millihartree at  $R_{\rm H-F} = 1.5$  Å to over 40 millihartree at  $R_{\rm H-F} = 4.0$  Å. Already at  $R_{\rm H-F} = 1.875$  Å, CCSD(T) has a larger error relative to CCSDT than any of the other approximate triples methods studied in this work. As for the largely SR A  ${}^{3}\Sigma_{u}^{+}$ state, CCSD(T) does not experience the breakdown, producing rather small errors relative to CCSDT, which do not exceed 0.355 millihartree, but the CCSD(T) results for the  $X \, {}^{1}\Sigma_{g}^{+}$ state are very poor. This unbalanced description of the singlet and triplet states leads to  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  separations that are larger than those given by CCSDT anywhere between 148 cm<sup>-1</sup> at  $R_{\rm H-F} = 1.5$  Å and 8957 cm<sup>-1</sup> at  $R_{\rm H-F} = 4.0$  Å. Clearly, alternative and more robust ways of handling connected triple excitations in the SRCC formalism are needed to reproduce the full CCSDT-quality data for the X  $^{1}\Sigma_{g}^{+}$  and A  $^{3}\Sigma_{u}^{+}$  states in the (HFH)<sup>-</sup> system and the gap between them.

As shown in Refs. [162, 285], the CR-CC(2,3) methodology, including variants A and D

of CR-CC(2,3) examined in Tables 3.6–3.8 and Figures 3.5–3.7, offers great improvements in the results compared to CCSD(T), particularly for the quasi-degenerate  $X^{-1}\Sigma_g^+$  state and especially at larger H–F separations. The large MUE and NPE values relative to full CCSDT characterizing the CCSD(T) energies of the X  ${}^{1}\Sigma_{g}^{+}$  state in the  $R_{\rm H-F} = 1.5 - 4.0$  Å region, of 40.727 and 40.292 millihartree, reduce to the much better values of 2.566 and 2.906 millihartree, respectively, when the  $CR-CC(2,3)_A$  approach is employed, and 2.800 and 2.509 millihartree when the  $CR-CC(2,3)_D$  approximation is utilized. Although CCSD(T) works well for the A  ${}^{3}\Sigma_{u}^{+}$  state, producing small MUE and NPE values relative to full CCSDT of only 0.335 and 0.250 millihartree, respectively, both variants of CR-CC(2,3) improve the CCSD(T) results at all geometries examined for the  $A^{-3}\Sigma_u^+$  state as well. As a result of the very good agreement of the CR-CC(2,3) approaches with the CCSDT method for the two states of  $(HFH)^-$  examined here, the overall performance of the CR-CC(2,3)<sub>A</sub> and CR- $CC(2,3)_D$  methods in describing the  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap in (HFH)<sup>-</sup> is a lot better than the performance of CCSD(T). When compared to CCSDT, the  $CR-CC(2,3)_A$  approach is slightly more accurate than its CR-CC(2,3)<sub>D</sub> counterpart in the  $R_{\rm H-F} = 2.125 - 4.0$  Å region, but CR-CC(2,3)<sub>D</sub> works better when  $R_{\rm H-F} < 2.125$  Å. In the end both CR-CC(2,3) approaches perform in a similar manner, giving the MUE and NPE values relative to full CCSDT characterizing the  $A \ ^{3}\Sigma_{u}^{+} - X \ ^{1}\Sigma_{g}^{+}$  gap in the 500 – 600cm<sup>-1</sup> range, which is a great improvement over CCSD(T), but one does, of course, wonder if further improvements can be made by turning to other approximate treatments of triples.

The active-space treatment of triples via the CCSDt approach, although providing more uniform errors at various H–F separations relative to full CCSDT, reducing the MUE and NPE values characterizing the  $A \ {}^{3}\Sigma_{u}^{+} - X \ {}^{1}\Sigma_{g}^{+}$  gap to 299 and 377 cm<sup>-1</sup>, respectively, is not completely satisfactory either. Indeed, the differences between the CCSDt and CCSDT  $A \,{}^{3}\Sigma_{u}^{+} - X \,{}^{1}\Sigma_{g}^{+}$  gaps increase from 140 cm<sup>-1</sup> at  $R_{\rm H-F} = 1.5$  Å to 299 cm<sup>-1</sup> at  $R_{\rm H-F} = 2.25$ Å, to decrease again to 78 cm<sup>-1</sup> at  $R_{\rm H-F} = 4.0$  Å. While the MUE and NPE values relative to CCSDT, which are 299 and 377 cm<sup>-1</sup>, respectively, are smaller than those given by CCSD(T), CR-CC(2,3)<sub>A</sub>, and CR-CC(2,3)<sub>D</sub>, we are still not obtaining the desirable agreement with CCSDT considering that errors resulting from CCSDt calculations, compared to CCSDT, in the  $R_{\rm H-F} = 2.25 - 4.0$  Å region are comparable to the magnitude of the  $A \,{}^{3}\Sigma_{u}^{+} - X \,{}^{1}\Sigma_{g}^{+}$  gap itself. The inability of the CCSDt approach to more accurately reproduce the CCSDT singlet-triplet gap values in the (HFH)<sup>-</sup> system stems from the rather significant differences between the total energies of the  $X \,{}^{1}\Sigma_{g}^{+}$  and  $A \,{}^{3}\Sigma_{u}^{+}$  states resulting from the CCSDt and CCSDT calculations, which are as large as 2.582 and 1.895 millihartree, respectively. This disagreement could be resolved by expanding the active space in the CCSDt calculations, but here we are more interesting in examining how the CC(*P*;*Q*) corrections, such as the CC(t;3) correction to CCSDt, cope with this issue.

As desired, the  $CC(t;3)_A$  and  $CC(t;3)_D$  approaches improve the total energies of the  $X \, {}^{1}\Sigma_{g}^{+}$  and  $A \, {}^{3}\Sigma_{u}^{+}$  states of (HFH)<sup>-</sup>, when compared to the CR-CC(2,3)<sub>A</sub>, CR-CC(2,3)<sub>D</sub>, and CCSDt calculations, bringing the results to a closer agreement with the full CCSDT data. The MUE values relative to full CCSDT characterizing the CR-CC(2,3)<sub>A</sub>, CR-CC(2,3)<sub>D</sub>, and CCSDt energies of the  $X \, {}^{1}\Sigma_{g}^{+}$  state along the  $R_{H-F}$  coordinate, of 2.566, 2.800, and 2.582 millihartree, respectively, are reduced to a mere 0.357 millihartree, when the CC(t;3)<sub>A</sub> approximation is employed, and 0.335 millihartree, when the CC(t;3)<sub>D</sub> method is used. In addition to the small MUE values, the corresponding NPE values of 0.420 and 0.253 millihartree for the CC(t;3)<sub>A</sub> and CC(t;3)<sub>D</sub> approximations are major improvements over the 2.906, 2.509, and 1.814 millihartree values obtained in the CR-CC(2,3)<sub>A</sub>, CR-CC(2,3)<sub>D</sub>, and CCSDt calculations, not to mention the catastrophically failing CCSD(T) approach, which

gives a 40.292 millihartree error. The situation for the "easier", largely single-configurational,  $A^{3}\Sigma_{u}^{+}$  state is somewhat different because of the already low MUE and NPE values given by the  $CR-CC(2,3)_A$ ,  $CR-CC(2,3)_D$ , and CCSDt calculations. Still, both  $CC(t;3)_A$  and  $CC(t;3)_D$  provide improved description of the  $A^{-3}\Sigma_u^+$  state over their CR-CC(2,3) and CCSDt counterparts, resulting in the NPE and MUE values relative to CCSDT of 0.130 and 0.081 millihartree, respectively, in the case of  $CC(t;3)_A$  and 0.207 and 0.047 millihartree, respectively, in the  $CC(t;3)_D$  case. The highly accurate description of the lowest singlet and triplet states by the CC(t;3) approaches results in the greatly improved description of the  $A \ ^{3}\Sigma_{u}^{+} - X \ ^{1}\Sigma_{g}^{+}$  gap. Both the CC(t;3)<sub>A</sub> and CC(t;3)<sub>D</sub> methods reproduce the CCSDT  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap values to within tens of wavenumbers, with the maximum errors of only 62 cm<sup>-1</sup> for  $CC(t;3)_A$  and 37 cm<sup>-1</sup> for  $CC(t;3)_D$ . These extremely low errors in the total energies of the X  ${}^{1}\Sigma_{g}^{+}$  and A  ${}^{3}\Sigma_{u}^{+}$  states and the corresponding A  ${}^{3}\Sigma_{u}^{+} - X {}^{1}\Sigma_{g}^{+}$ gap given by the CC(t;3) schemes signify that the  $CC(t;3)_A$  and  $CC(t;3)_D$  methods are promising, computationally efficient alternatives to full CCSDT, which may help the various applications where the CCSDT level of theory is sufficient, but, as already pointed out, the (HFH)<sup>-</sup> system has significant correlation effects beyond CCSDT, especially the connected quadruples. Indeed, CCSDT gives errors relative to full CI, which are as high as 2.276 millihartree for the  $X \, {}^{1}\Sigma_{g}^{+}$  state, 0.389 millihartree for the  $A \, {}^{3}\Sigma_{u}^{+}$  state, and 420 cm<sup>-1</sup> for the corresponding  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap. Thus, we move now to SRCC methods that account for triples as well as quadruples correlations.

By comparing the CCSDTQ and full CI values, we immediately see that CCSDTQ is essentially exact, giving errors to full CI that do not exceed 0.148 millihartree for the  $X \, {}^{1}\Sigma_{g}^{+}$ state, 14 microhartree for the  $A \, {}^{3}\Sigma_{u}^{+}$  state, and 30 cm<sup>-1</sup> for the  $A \, {}^{3}\Sigma_{u}^{+} - X \, {}^{1}\Sigma_{g}^{+}$  gap. Unfortunately, the full CCSDTQ calculations are usually prohibitively expensive, so we need

to examine approximate ways of handling triples and quadruples, looking for methods that can reproduce full CCSDTQ data at the fraction of the cost. The completely renormalized methods with triples and quadruples tested in this study, namely  $\operatorname{CR-CC}(2,4)_{\operatorname{AA}}$  and  $\operatorname{CR-cc}(2,4)_{\operatorname{AA}}$  $CC(2,4)_{DA}$ , although not too bad, are not capable of providing accuracies we are interested in. The MUE and NPE values relative to full CCSDTQ of 2.105 and 2.610 millihartree characterizing the CR-CC(2,4)<sub>AA</sub> calculations for the  $X^{1}\Sigma_{g}^{+}$  state and the corresponding 3.267 and 2.889 millihartree MUE and NPE values resulting from the  $CR-CC(2,4)_{DA}$  calculations are acceptable, especially given the challenging nature of this state when both H–F bonds are significantly stretched, but they are not as good as we desire. The situation for the SR  $A^{3}\Sigma_{u}^{+}$  state is much better, but this is not sufficient, since the MUE and NPE values relative to CCSDTQ characterizing the  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap as a function of  $R_{\rm H-F}$  remain quite high, on the order of  $500-600 \text{ cm}^{-1}$ . The active-space CCSDtq approach improves the overall description of the  $A \ ^{3}\Sigma_{u}^{+} - X \ ^{1}\Sigma_{g}^{+}$  gap, reducing the above MUE and NPE values to a 400–500  $\rm cm^{-1}$  level, but it is quite clear from our tables that we need to more if we are reach accuracies at the 0.1 millihartree and  $100 \text{ cm}^{-1}$  levels. We need to find a way to correct the CCSDtq energies for the higher-order dynamical correlations that they cannot describe when smaller active spaces are used. The solution to this is provided by the CC(t,q;3) or CC(t,q;3,4) corrections developed in this work, which we discuss next.

For the more demanding  $X^{1}\Sigma_{g}^{+}$  state, when the CCSDtq energies are corrected for the missing triples, the resulting CC(t,q;3)<sub>A</sub> and CC(t,q;3)<sub>D</sub> values are in very good agreement with CCSDTQ. The MUE values relative to CCSDTQ characterizing the CC(t,q;3)<sub>A</sub> and CC(t,q;3)<sub>D</sub> calculations for the  $X^{1}\Sigma_{g}^{+}$  state are 1.127 and 0.674 millihartree, respectively, which is a significant improvement over the 2.105, 3.267, and 3.349 millihartree MUEs obtained with CR-CC(2,4)<sub>AA</sub>, CR-CC(2,4)<sub>DA</sub>, and CCSDtq. Similar remarks apply to the

NPE values. The  $CC(t,q;3)_D$  approach provides a better overall agreement with CCSDTQ than its  $CC(t,q;3)_A$  counterpart, although both CC(t,q;3) methods work well. Indeed, the MUE and NPE values relative to CCSDTQ resulting from the  $CC(t,q;3)_D$  calculations, of 0.574 and 0.590 millihartree, respectively, are somewhat better than the 1.127 and 0.863 millihartree values obtained with  $CC(t,q;3)_A$ . When the CCSDtq energies are corrected for the missing triple and quadruple excitations via the  $CR-CC(2,4)_{AA}$  and  $CR-CC(2,4)_{DA}$  approaches, the agreement with CCSDTQ is even more impressive. Both approaches have maximum errors of only 0.366 millihartree and the NPE values characterizing the  $CC(t,q;3,4)_{AA}$ and  $CR-CC(2,4)_{DA}$  calculations are equally good, especially in the latter case, where we obtain 0.371 millihartree.

In the case of the  $A^{-3}\Sigma_{u}^{+}$  state, when we examine the CC(P;Q) methods correcting CCSDtq for the missing triples and quadruples, we do not witness the large improvement in results as with the  $X^{-1}\Sigma_{g}^{+}$  state, since the  $A^{-3}\Sigma_{u}^{+}$  state has a SR character and methods such as CR-CC(2,4) already work well for it. The  $CC(t,q;3)_{A}$  and  $CC(t,q;3)_{D}$  approaches have small MUE values relative to CCSDTQ of 0.481 and 0.211 millihartree, respectively, and all energies lie above CCSDTQ for both methods. Furthermore, both approaches behave very systematically in the entire  $R_{H-F} = 1.5 - 4.0$  Å region, as reflected by the low NPE values of 81 and 71 microhartree. When we correct for the missing quadruples as well, via the  $CC(t,q;3,4)_{AA}$  and  $CC(t,q;3,4)_{DA}$  approaches, the energies are lowered slightly below the CCSDTQ ones, but the very small MUE and NPE values, on the order of 0.1–0.4 millihartree, remain.

The highly accurate description of the  $X^{-1}\Sigma_g^+$  state provided by the CC(t,q;3) and CC(t,q;3,4) approximations, combined with the equally small errors obtained for the  $A^{-3}\Sigma_u^+$  state leads to impressively accurate  $A^{-3}\Sigma_u^+ - X^{-1}\Sigma_g^+$  gap predictions. The MUE values rela-

tive to CCSDTQ of 465, 642, and 375 cm<sup>-1</sup> obtained in the CR-CC(2,4)<sub>AA</sub>, CR-CC(2,4)<sub>DA</sub> and CCSDtq calculations are reduced to just 147 and 102 cm<sup>-1</sup> when the CC(t,q;3)<sub>A</sub> and  $CC(t,q;3)_D$  methods are employed, and 108 and 76 cm<sup>-1</sup> when the  $CC(t,q;3,4)_{AA}$  and  $CC(t,q;3,4)_{DA}$  approaches are used. The NPE values improve in a similar manner, so it is hard to tell the difference between the CC(t,q;3) and especially, CC(t,q;3,4) gap values and their virtually exact CCSDTQ counterparts.

In summary, the challenging (HFH)<sup>-</sup> ion, where the degree of biradical character can be continuously varied by simultaneously stretching both H–F bonds, proved to be an excellent system to examine the performance of the CC(P;Q) methods. When it comes to triple excitations, the CR-CC(2,3)<sub>A</sub>, CR-CC(2,3)<sub>D</sub>, and CCSDt approaches improve the erratic CCSD(T) data, but none of them provides satisfactory agreement with CCSDT for total energies of the  $A^{-3}\Sigma_{u}^{+}$  and  $X^{-1}\Sigma_{g}^{+}$  states and the corresponding energy gap. When we correct the CCSDt energies for the correlation effects due to the missing triples using the CC(t;3) methodology, both the CC(t;3)<sub>A</sub> and CC(t;3)<sub>D</sub> variants, which replace the expensive iterative CPU steps of CCSDT that scale with the system size as  $\mathcal{N}^{8}$  by the iterative  $\mathcal{N}^{6}$ type and non-iterative  $\mathcal{N}^{7}$ -type calculations, reproduce the total energies for both states to within a small fraction of a millihartree. As a result, the CC(t;3) approaches accurately describe the corresponding  $A^{-3}\Sigma_{u}^{+} - X^{-1}\Sigma_{g}^{+}$  gap, with errors on the order of only tens of wavenumbers relative to full CCSDT. But, as we have learned above, one must go beyond the CCSDT level to obtain an accurate description of the singlet–triplet gap in (HFH)<sup>-</sup>.

When examining the approximate quadruples methods, we demonstrated that among the CR-CC(2,4)<sub>AA</sub>, CR-CC(2,4)<sub>DA</sub> and CCSDtq approaches, none are capable of reproducing the CCSDTQ data as accurately as desired. However, when the CCSDtq energies are corrected for the missing correlations due to triples or triples and quadruples, we obtain the virtually perfect agreement with full CCSDTQ, both for the total energies of the  $X \ {}^{1}\Sigma_{g}^{+}$  and  $A \ {}^{3}\Sigma_{u}^{+}$  states and the gap between them. The CC(t,q;3,4)<sub>DA</sub> scheme, giving gap values to within tens of cm<sup>-1</sup> from CCSDTQ, turned out to be particularly effective. The excellent agreement between the CC(t,q;3) and CC(t,q;3,4) results and their CCSDTQ counterpart is very encouraging, since CC(t,q;3) and CC(t,q;3,4) replace the iterative CPU steps of CCSDTQ that scale as  $\mathcal{N}^{10}$  with the significantly less expensive iterative  $\mathcal{N}^{6}$ -type and non-iterative  $\mathcal{N}^{7}$ -type calculations.

Table 3.6: A comparison of the energies resulting from various CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the  $X \, {}^{1}\Sigma_{g}^{+}$  state of the linear,  $D_{\infty h}$ -symmetric,  $(\text{HFH})^{-}$  system, as described by the 6-31G(d,p) basis set [286, 287], at a few values of the H–F distance  $R_{\text{H-F}}$  (in Å).<sup>*a*</sup>

			$R_{ ext{H-F}}$		
Method	1.500	1.625	1.750	1.875	2.000
Full $CI^b$	-100.589392	-100.584704	-100.577669	-100.570151	-100.563055
		CC n	nethods with t	$riples^{c}$	
CCSD(T)	-0.435(0.827)	-1.103(0.331)	-2.207(-0.594)	-3.857(-2.071)	-6.122(-4.177)
$CR-CC(2,3)_A$	2.209(3.471)	2.448(3.882)	2.566(4.179)	2.505(4.291)	2.239(4.184)
$CR-CC(2,3)_D$	-0.343(0.919)	-0.467(0.967)	-0.686(0.927)	-1.018(0.768)	-1.455(0.490)
$\mathrm{CCSDt}^d$	2.532(3.794)	2.541(3.975)	2.557(4.170)	2.575(4.361)	2.582(4.527)
$CC(t;3)_A{}^d$	0.261(1.523)	0.293(1.727)	0.323(1.936)	0.347(2.133)	0.357(2.302)
$CC(t;3)_D{}^d$	-0.197(1.065)	-0.164(1.270)	-0.136(1.477)	-0.113(1.673)	-0.098(1.847)
CCSDT	0.000(1.262)	0.000(1.434)	0.000(1.613)	0.000(1.786)	0.000(1.945)
		CC methods	with triples an	nd quadruples <sup><math>e</math></sup>	
$CR-CC(2,4)_{AA}$	1.820(1.908)	2.027(2.131)	2.105(2.225)	2.002(2.136)	1.701(1.844)
$CR-CC(2,4)_{DA}$	-0.732(-0.644)	-0.888(-0.784)	-1.146(-1.026)	-1.521(-1.387)	-1.993(-1.850)
$\mathrm{CCSDtq}^d$	3.231(3.319)	3.284(3.388)	3.326(3.446)	3.349(3.483)	3.346(3.489)
$CC(t,q;3)_A{}^d$	0.963(1.051)	1.039(1.143)	1.096(1.216)	1.125(1.259)	1.127(1.270)
$CC(t,q;3)_D{}^d$	0.506(0.594)	0.583(0.687)	0.638(0.758)	0.668(0.802)	0.674(0.817)
$CC(t,q;3,4)_{AA}^{d}$	0.091(0.179)	0.149(0.253)	0.208(0.328)	0.264(0.398)	0.311(0.454)
$CC(t,q;3,4)_{DA}^{d}$	-0.366(-0.278)	-0.307(-0.203)	-0.249(-0.129)	-0.194(-0.060)	-0.142(0.001)
$\mathbf{CCSDTQ}$	0.000(0.088)	0.000(0.104)	0.000(0.120)	0.000(0.134)	0.000(0.143)

Table 3.6 (cont'd)

		$R_{ m H-F}$				
2.125	2.250	2.500	3.000	4.000	MUE	NPE
-100.556686	-100.551083	-100.542059	-100.531336	-100.526513		
	$\mathbf{C}\mathbf{C}$	methods with tri	$ples^c$			
-8.994(-6.913)	-12.411(-10.226)	-20.360(-18.084)	-34.862(-32.964)	-40.727(-40.115)	40.727(40.115)	40.292(40.942)
1.795(3.876)	1.246(3.431)	0.224(2.500)	-0.340(1.558)	-0.027(0.585)	2.566(4.291)	2.906(3.706)
-1.937(0.144)	-2.378(-0.193)	-2.800(-0.524)	-1.838(0.060)	-0.291(0.321)	2.800(0.967)	2.509(1.491)
2.566(4.647)	2.513(4.698)	2.283(4.559)	1.618(3.516)	0.768(1.380)	2.582(4.698)	1.814(3.318)
0.353(2.434)	0.332(2.517)	0.244(2.520)	-0.021(1.877)	-0.063(0.549)	0.357(2.520)	0.420(1.971)
-0.087(1.994)	-0.082(2.103)	-0.108(2.168)	-0.335(1.563)	-0.244(0.368)	0.335(2.168)	0.253(1.800)
0.000(2.081)	0.000(2.185)	0.000(2.276)	0.000(1.898)	0.000(0.612)	0.000(2.276)	0.000(1.664)
	CC methods	s with triples and	$quadruples^{e}$	. ,		
1.237(1.385)	0.692(0.838)	-0.244(-0.115)	-0.505(-0.434)	-0.113(-0.101)	2.105(2.225)	2.610(2.659)
-2.495(-2.347)	-2.932(-2.786)	-3.267(-3.138)	-2.002(-1.931)	-0.378(-0.366)	3.267(3.138)	2.889(2.772)
3.307(3.455)	3.225(3.371)	2.937(3.066)	2.139(2.210)	1.093(1.105)	3.349(3.489)	2.256(2.384)
1.102(1.250)	1.054(1.200)	0.910(1.039)	0.515(0.586)	0.264(0.276)	1.127(1.270)	0.863(0.994)
0.664(0.812)	0.642(0.788)	0.561(0.690)	0.206(0.277)	0.084(0.096)	0.674(0.817)	0.590(0.721)
0.346(0.494)	0.366(0.512)	0.353(0.482)	0.097(0.168)	-0.157(-0.145)	0.366(0.512)	0.523(0.657)
-0.092(0.056)	-0.046(0.100)	0.005(0.134)	-0.212(-0.141)	-0.337(-0.325)	0.366(0.325)	0.371(0.459)
0.000(0.148)	0.000(0.146)	0.000(0.129)	0.000(0.071)	0.000(0.012)	0.000(0.148)	0.000(0.136)

<sup>a</sup> The full CI energies were taken from Ref. [162]. As in Ref. [162], the lowest-energy core orbital was frozen in the post-SCF calculations and the spherical components of the fluorine d orbital were employed throughout.

<sup>b</sup> The total full CI energies in hartree.

 $^{c}$  For the CC methods with up to triple excitations, the reported energy values, in millihartree, are errors relative to full CCSDT and, in parentheses, relative to full CI.

<sup>d</sup> The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of two active electrons and two active orbitals corresponding to the HOMO and LUMO (the  $\sigma_q$  and  $\sigma_u$  valence orbitals).

 $^{e}$  For the CC methods with triples and quadruples, the reported energy values, in millihartree, are errors relative to full CCSDTQ and, in parentheses, relative to full CI.



Figure 3.5: A comparison of the energies resulting from various CC calculations including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the  $X^{-1}\Sigma_g^+$  state of the linear,  $D_{\infty h}$ -symmetric, (HFH)<sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at several values of the H–F distance  $R_{\text{H-F}}$ . The numerical values of the errors are found in Table 3.6. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results.

Table 3.7: A comparison of the energies resulting from various CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the  $A^{3}\Sigma_{u}^{+}$  state of the linear,  $D_{\infty h}$ -symmetric, (HFH)<sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at a few values of the H–F distance  $R_{\text{H-F}}$  (in Å).<sup>*a*</sup>

			$R_{ m H-F}$		
Method	1.500	1.625	1.750	1.875	2.000
Full $CI^b$	-100.545993	-100.552773	-100.555291	-100.555097	-100.553271
		CC r	methods with t	$riples^{c}$	
$\operatorname{CCSD}(T)$	0.240(0.600)	0.253(0.625)	0.269(0.652)	0.289(0.678)	0.308(0.697)
$CR-CC(2,3)_A$	0.199(0.559)	0.189(0.561)	0.174(0.557)	0.158(0.547)	0.141(0.530)
$CR-CC(2,3)_D$	-0.217(0.143)	-0.197(0.175)	-0.181(0.202)	-0.173(0.216)	-0.172(0.217)
$\mathrm{CCSDt}^d$	1.895(2.255)	1.755(2.127)	1.616(1.999)	1.479(1.868)	1.352(1.741)
$CC(t;3)_A{}^d$	0.130(0.490)	0.121(0.493)	0.109(0.492)	0.095(0.484)	0.080(0.469)
$CC(t;3)_D{}^d$	-0.207(0.153)	-0.188(0.184)	-0.174(0.209)	-0.168(0.221)	-0.167(0.222)
CCSDT	0.000(0.360)	0.000(0.372)	0.000(0.383)	0.000(0.389)	0.000(0.389)
		CC methods	with triples an	nd quadruples <sup><math>e</math></sup>	
$CR-CC(2,4)_{AA}$	-0.001(0.012)	-0.006(0.007)	-0.016(-0.002)	-0.027(-0.014)	-0.042(-0.030)
$CR-CC(2,4)_{DA}$	-0.418(-0.405)	-0.392(-0.379)	-0.371(-0.357)	-0.358(-0.345)	-0.355(-0.343)
$\mathrm{CCSDtq}^d$	2.242(2.255)	2.114(2.127)	1.985(1.999)	1.855(1.868)	1.729(1.741)
$CC(t,q;3)_A{}^d$	0.477(0.490)	0.480(0.493)	0.478(0.492)	0.471(0.484)	0.457(0.469)
$CC(t,q;3)_D{}^d$	0.140(0.153)	0.171(0.184)	0.195(0.209)	0.208(0.221)	0.210(0.222)
$CC(t,q;3,4)_{AA}^{d}$	-0.072(-0.059)	-0.075(-0.062)	-0.082(-0.068)	-0.092(-0.079)	-0.105(-0.093)
$CC(t,q;3,4)_{DA}d$	-0.409(-0.396)	-0.385(-0.372)	-0.365(-0.351)	-0.354(-0.341)	-0.352(-0.340)
CCSDTQ	0.000(0.013)	0.000(0.013)	0.000(0.014)	0.000(0.013)	0.000(0.012)

Table 3.7 (cont'd)

		$R_{\mathrm{H-F}}$				
2.125	2.250	2.500	3.000	4.000	MUE	NPE
-100.550520	-100.547315	-100.540796	-100.531257	-100.526513		
	CC m	ethods with tri	$ples^c$			
0.326(0.711)	0.335(0.712)	0.318(0.679)	0.191(0.537)	0.085(0.434)	0.335(0.712)	0.250(0.278)
0.128(0.513)	0.115(0.492)	0.096(0.457)	0.084(0.430)	0.082(0.431)	0.199(0.561)	0.117(0.131)
-0.166(0.219)	-0.165(0.212)	-0.167(0.194)	-0.169(0.177)	-0.180(0.169)	0.217(0.219)	0.052(0.076)
1.241(1.626)	1.149(1.526)	1.034(1.395)	1.031(1.377)	1.121(1.470)	1.895(2.255)	0.864(0.878)
0.069(0.454)	0.059(0.436)	0.049(0.410)	0.065(0.411)	0.082(0.431)	0.130(0.493)	0.081(0.083)
-0.162(0.223)	-0.160(0.217)	-0.163(0.198)	-0.167(0.179)	-0.180(0.169)	0.207(0.223)	0.047(0.070)
0.000(0.385)	0.000(0.377)	0.000(0.361)	0.000(0.346)	0.000(0.349)	0.000(0.389)	0.000(0.043)
	CC methods v	with triples and	l quadruples <sup><math>e</math></sup>			
-0.056(-0.044)	-0.069(-0.058)	-0.081(-0.070)	-0.054(-0.044)	-0.020(-0.010)	0.081(0.070)	0.080(0.082)
-0.350(-0.338)	-0.348(-0.337)	-0.344(-0.333)	-0.307(-0.297)	-0.282(-0.272)	0.418(0.405)	0.136(0.133)
1.614(1.626)	1.515(1.526)	1.384(1.395)	1.367(1.377)	1.460(1.470)	2.242(2.255)	0.875(0.878)
0.442(0.454)	0.425(0.436)	0.399(0.410)	0.401(0.411)	0.421(0.431)	0.480(0.493)	0.081(0.083)
0.211(0.223)	0.206(0.217)	0.187(0.198)	0.169(0.179)	0.159(0.169)	0.211(0.223)	0.071(0.070)
-0.117(-0.105)	-0.127(-0.116)	-0.130(-0.119)	-0.074(-0.064)	-0.020(-0.010)	0.130(0.119)	0.110(0.109)
-0.347(-0.335)	-0.346(-0.335)	-0.342(-0.331)	-0.306(-0.296)	-0.282(-0.272)	0.409(0.396)	0.127(0.124)
0.000(0.012)	0.000(0.011)	0.000(0.011)	0.000(0.010)	0.000(0.010)	0.000(0.014)	0.000(0.004)

<sup>a</sup> The full CI energies were taken from Ref. [162]. As in Ref. [162], the lowest-energy core orbital was frozen in the post-SCF calculations and the spherical components of the fluorine d orbital were employed throughout.

<sup>b</sup> The total full CI energies in hartree.

 $^{c}$  For the CC methods with up to triple excitations, the reported energy values, in millihartree, are errors relative to full CCSDT and, in parentheses, relative to full CI.

<sup>d</sup> The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations consisted of two active electrons and two active orbitals corresponding to the HOMO and LUMO (the  $\sigma_q$  and  $\sigma_u$  valence orbitals).

 $^{e}$  For the CC methods with triples and quadruples, the reported energy values, in millihartree, are errors relative to full CCSDTQ and, in parentheses, relative to full CI.


Figure 3.6: A comparison of the energies resulting from various CC calculations including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the  $A^{3}\Sigma_{u}^{+}$  state of the linear,  $D_{\infty h}$ -symmetric, (HFH)<sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at several values of the H–F distance  $R_{\text{H-F}}$ . The numerical values of the errors are found in Table 3.7. Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compared various approximate quadruples methods with the parent CCSDTQ results.

Table 3.8: A comparison of the  $A {}^{3}\Sigma_{u}^{+} - X {}^{1}\Sigma_{g}^{+}$  gap values resulting from various CC calculations including up to triple and up to quadruple excitations with their parent CCSDT and CCSDTQ counterparts, and the corresponding full CI data for the linear,  $D_{\infty h}$ -symmetric, (HFH)<sup>-</sup> system, as described by the 6-31G(d,p) basis set [286, 287], at a few values of the H–F distance  $R_{\text{H-F}}$  (in Å).<sup>*a*</sup>

			$R_{\mathrm{H-F}}$		
Method	1.500	1.625	1.750	1.875	2.000
Full CI <sup>b</sup>	9525	7008	4911	3304	2147
		CC m	ethods with	$1 triples^{c}$	
$\operatorname{CCSD}(\mathrm{T})$	148(-50)	298(65)	544(274)	910(603)	1411(1070)
$CR-CC(2,3)_A$	-441(-639)	-496(-729)	-525(-795)	-515(-822)	-461(-802)
$CR-CC(2,3)_D$	28(-170)	59(-174)	111(-159)	186(-121)	281(-60)
$\mathrm{CCSDt}^d$	-140(-338)	-173(-406)	-206(-476)	-240(-547)	-270(-611)
$CC(t;3)_A{}^d$	-29(-227)	-38(-271)	-47(-317)	-55(-362)	-61(-402)
$CC(t;3)_D{}^d$	-2(-200)	-5(-238)	-8(-278)	-12(-319)	-15(-356)
CCSDT	0(-198)	0(-233)	0(-270)	0(-307)	0(-341)
	CC	methods v	with triples	and quadru	$ples^e$
$CR-CC(2,4)_{AA}$	-400(-416)	-446(-466)	-465(-488)	-445(-472)	-383(-411)
$CR-CC(2,4)_{DA}$	68(52)	109(89)	170(147)	256(229)	359(331)
$\mathrm{CCSDtq}^d$	-218(-234)	-256(-276)	-294(-317)	-327(-354)	-355(-383)
$CC(t,q;3)_A{}^d$	-107(-123)	-123(-143)	-135(-158)	-143(-170)	-147(-175)
$CC(t,q;3)_D{}^d$	-81(-97)	-90(-110)	-97(-120)	-100(-127)	-102(-130)
$CC(t,q;3,4)_{AA}^{d}$	-36(-52)	-49(-69)	-64(-87)	-78(-105)	-92(-120)
$CC(t,q;3,4)_{DA}$ <sup>d</sup>	-10(-26)	-17(-37)	-25(-48)	-35(-62)	-47(-75)
CCSDTQ	0(-16)	0(-20)	0(-23)	0(-27)	0(-28)

		$R_{\mathrm{H-F}}$				
2.125	2.250	2.500	3.000	4.000	MUE	NPE
1353	827	277	17	0		
	CC met	hods with t	$riples^{c}$			
2046(1674)	2798(2401)	4538(4118)	7693(7353)	8957(8899)	8957(8899)	8809(8949)
-366(-738)	-248(-645)	-28(-448)	93(-247)	24(-34)	525(822)	618(788)
389(17)	486(89)	578(158)	366(26)	25(-33)	578(174)	553(332)
-291(-663)	-299(-696)	-274(-694)	-129(-469)	78(20)	299(696)	377(716)
-62(-434)	-60(-457)	-43(-464)	19(-321)	32(-26)	62(463)	94(437)
-16(-388)	-17(-414)	-12(-432)	37(-303)	14(-44)	37(432)	54(388)
0(-372)	0(-397)	0(-420)	0(-340)	0(-58)	0(420)	0(362)
CC	methods wi	th triples ar	nd quadruple	$es^e$		
-283(-313)	-167(-197)	36(10)	99(86)	20(20)	465(488)	564(574)
471(441)	567(537)	642(616)	372(359)	21(21)	642(616)	621(595)
-371(-401)	-375(-405)	-341(-367)	-169(-182)	80(80)	375(405)	455(485)
-144(-174)	-138(-168)	-112(-138)	-25(-38)	34(34)	147(176)	181(209)
-99(-129)	-95(-125)	-82(-108)	-8(-21)	16(16)	102(130)	118(146)
-101(-131)	-108(-138)	-106(-132)	-38(-51)	30(30)	108(138)	138(168)
-56(-86)	-65(-95)	-76(-102)	-21(-34)	12(12)	76(102)	88(114)
0(-30)	0(-30)	0(-26)	0(-13)	0(0)	0(30)	0(30)

Table 3.8 (cont'd)

<sup>a</sup> The full CI energies were taken from Ref. [162]. As in Ref. [162], the lowest-energy core orbital was frozen in the post-SCF calculations and the spherical components of the fluorine d orbital were employed throughout.

<sup>b</sup> The full CI values of the  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap in cm<sup>-1</sup>. <sup>c</sup> For the CC methods with up to triple excitations, the reported energy values, in cm<sup>-1</sup>, are errors in the calculated  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gaps relative to full CCSDT and, in parentheses, relative to full CI.

<sup>d</sup> The active space used in the CCSDt, CCSDtq, CC(t;3), CC(t;3), CC(t;3), and CC(t;3,4) calculations consisted of two active electrons and two active orbitals corresponding to the HOMO and LUMO (the  $\sigma_g$  and  $\sigma_u$  valence orbitals).

<sup>*e*</sup> For the CC methods with triples and quadruples, the reported energy values, in cm<sup>-1</sup>, are errors in the calculated  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gaps relative to full CCSDTQ and, in parentheses, relative to full CI.



**R**<sub>H-F</sub> Bond Lengths (Å) Figure 3.7: A comparison of the  $A^{3}\Sigma_{u}^{+} - X^{1}\Sigma_{g}^{+}$  gap values resulting from various CC calculations including up to triple and up to quadruple excitations, along with their parent CCSDT and CCSDTQ counterparts, for the linear,  $D_{\infty h}$ -symmetric, (HFH)<sup>-</sup> system, as described by the 6-31G(d,p) basis set [286,287], at several values of the H–F distance  $R_{\text{H-F}}$ . The numerical values of the errors are found in Table 3.8: Top panel shows a comparison of CCSDT and CCSDTQ with full CI. Middle panel shows a comparison of various approximate triples methods with the parent CCSDT results. Bottom panel compares various approximate quadruples methods with the parent CCSDTQ results.

#### 3.2.4 Singlet–Triplet Gap in BN

While it may seem simple, accurate determination of the ground states of the 12-electron isoelectronic series, including, but not limited to BN, C<sub>2</sub>, BeO, CN<sup>+</sup>, and BO<sup>+</sup>, has historically been a very difficult task for both experiment and theory. This is because the two lowest-energy electronic states,  ${}^{1}\Sigma^{+}$  and  ${}^{3}\Pi$  ( ${}^{1}\Sigma_{g}^{+}$  and  ${}^{3}\Pi_{u}$  in the case of C<sub>2</sub>), are nearly degenerate. With the exception of BN, the  ${}^{1}\Sigma^{+}$  ( ${}^{1}\Sigma_{g}^{+}$  for C<sub>2</sub>) state has been determined to be slightly lower in energy. In the case of BN, the ground state is of  ${}^{3}\Pi$  symmetry and is lower than the  ${}^{1}\Sigma^{+}$  state by less than 200 cm<sup>-1</sup>. As shown in this section, arriving at a proper and balanced description of the lowest-energy  ${}^{3}\Pi$  ( $X {}^{3}\Pi$ ) and  ${}^{1}\Sigma^{+}$  ( $a {}^{1}\Sigma^{+}$ ) states of BN and the gap between them is a Herculean task.

From the beginning, BN gave experiment and theory trouble. The initial work by Douglas and Herzberg focused on investigating the  ${}^{3}\Pi$  state, and while they were able to observe the  ${}^{1}\Sigma^{+}$  state in their study, they could only speculate that  ${}^{3}\Pi$  was the lower-energy state [288]. In addition, their estimated bond length of the  ${}^{3}\Pi$  state was a gross underestimation of the true value. It was not until several years later that additional experiments [289–292] and the first theoretical calculations [293–297] offered support that the  ${}^{3}\Pi$  state is the ground state. However, none of the experiments provided an estimated energy difference and theory was severely overestimating the  $a {}^{1}\Sigma^{+} - X {}^{3}\Pi$  gap, even as high as 2.60 eV [295]. While these early calculations could not properly describe the energetics of the  $X {}^{3}\Pi$  and  $a {}^{1}\Sigma^{+}$  states, few of them did suggest that the bond length of the  $X {}^{3}\Pi$  state was significantly longer than that proposed by Douglas and Herzberg [294–297]. This prompted Bredohl and coworkers to perform the analysis which revealed a longer bond length that was in good agreement with the early theoretical estimates [292]. Soon after that, the focus was shifted toward accurately determining the  $a \, {}^{1}\Sigma^{+} - X \, {}^{3}\Pi$  gap by both experiment [298–300] and theory [298, 301–307], which explored several methods in an attempt to describe the low-lying states of BN. All of these calculations provided a lot of intuition on the important correlation effects that must be taken into account, such as the requirement of a sufficiently large basis set (some small basis sets can produce a wrong ordering of states) and the need for extremely accurate methods (low-level methods do not produce correct energetics and in some cases get the wrong ordering of states, even with a sufficiently large basis set). As experiment was settling on an adiabatic electronic  $(T_e) a {}^{1}\Sigma^{+} - X {}^{3}\Pi$  gap of less than 200 cm<sup>-1</sup>, giving results, such as 15–182 cm<sup>-1</sup> [299] or 158  $\pm$  36 cm<sup>-1</sup> [300], computational resources started increasing and previously unattainable desired calculations became a reality. In the most recent years, high-level calculations employing MRCI schemes [303–305, 308–311], SRCC methods with up to hextuple excitations [312–319], RMRCC approximations [315,316], quantum Monte Carlo approach [320], and even full CI [314,317,318] have been performed for the  $a \ ^1\Sigma^+$  and  $X \ ^3\Pi$ states. Many of these studies have shown that one can obtain an accurate description of both states and a corresponding  $a {}^{1}\Sigma^{+} - X {}^{3}\Pi$  separation on the order of about 200 cm<sup>-1</sup>, but one has to work very hard to come close to the best available experimental estimates.

After the exhaustive investigations of its low-lying states, as summarized above, BN now serves as a "torture" molecule to test new quantum chemistry methods. In this study, we explore the performance of our CC(P;Q) methodology [163–165], with a focus on CC(t;3), CC(t,q;3), and CC(t,q;3,4) approaches by investigating the low-lying X <sup>3</sup>II and a <sup>1</sup> $\Sigma$ <sup>+</sup> states of BN and the corresponding adiabatic gap between them. The spherical cc-pVDZ and cc-pVTZ basis sets [237] were employed and the results are compiled in Tables 3.9 and 3.10, respectively. We did not use larger basis sets, since our goal has been to compare our CC(P;Q) calculations with the full CCSDT and CCSDTQ data and it is quite hard to

perform full CCSDTQ calculations when larger basis sets are employed. Using the cc-pVDZ basis set, the equilibrium bond lengths for the two states were determined at the CCSD, CCSDt, CCSDT, CCSDtq, and CCSDTQ levels of theory, and the results are summarized in Table 3.9. For the CCSDt and CCSDtq optimizations, as well as the CC(t;3), CC(t;3), and CC(t,q;3,4) single-point calculations, the active space consisted of the valence  $1\pi$ ,  $5\sigma$ ,  $2\pi$ , and  $6\sigma$  orbitals of BN corresponding to the 2p subshells of the B and N atoms. In terms of these orbitals, the ROHF and RHF configurations, which we used in our various CC calculations, are  $|\{\operatorname{core}\}(1\pi)^3(5\sigma)^1(2\pi)^0(6\sigma)^0|$  for the X <sup>3</sup> $\Pi$  state and  $|\{\operatorname{core}\}(1\pi)^4(5\sigma)^0(2\pi)^0(6\sigma)^0|$  for the  $a^1 \Sigma^+$  state. The CCSDt geometries were used for all approximate triples methods, i.e, all SRCC methods with triples other than CCSDT, while the CCSDtq geometries were exploited in the calculations including quadruples other than CCSDTQ. As for the cc-pVTZ basis set, equilibrium bond lengths of the  $a^{1}\Sigma^{+}$  and  $X^{3}\Pi$  states were optimized at the CCSD, CCSDt, and CCSDT levels of theory and, once again, the CCSDt geometries were used for all approximate triples methods (see Table 3.10). The CCSDT/cc-pVTZ equilibrium bond lengths can be found in Ref. [312], but they were recalculated by us in this study as well. Since the CCSDt, CCSDT, CCSDtq, and CCSDTQ geometries are all in close agreement with one another, when using the cc-pVDZ basis set, and since the SRCC geometry optimizations with quadruples using cc-pVTZ basis set are rather expensive, we used the CCSDt/cc-pVTZ geometries in the calculations with approximate treatments of quadruples and the CCSDT/cc-pVTZ geometries in the calculations using CCSDTQ. For all post-SCF calculations, the lowest-energy molecular orbitals that correlate with the 1s orbitals of the B and N atoms were kept frozen.

The results in Tables 3.9 and 3.10 show that independent of the basis set, the CCSD approach is incapable of providing a proper description of the BN molecule. Although

this is not a new finding, it is worth commenting on it. As shown in Tables 3.9 and 3.10, CCSD provides noticeably smaller bond lengths than the higher-order CCSDT and CCSDTQ methods, with results for the X  ${}^{3}\Pi$  state being particularly bad. The CCSD approach overestimates the adiabatic  $a {}^{1}\Sigma^{+} - X {}^{3}\Pi$  splitting by about 3500 cm<sup>-1</sup> when compared to CCSDT and nearly 4000 cm<sup>-1</sup> when compared to CCSDTQ. Clearly, the CCSD method cannot provide an accurate and balanced description of many-electron correlation effects for the two states. So, as demonstrated in several earlier studies, we must turn to the higher-order approaches in order to properly describe the  $a {}^{1}\Sigma^{+}$  and  $X {}^{3}\Pi$  states of BN and the corresponding splitting.

As can be seen in Tables 3.9 and 3.10 the CCSDT method provides significant improvements over CCSD, reducing the  $a^{1}\Sigma^{+} - X^{3}\Pi$  splitting from 4196 cm<sup>-1</sup> obtained with CCSD to 799 cm<sup>-1</sup> when the cc-pVDZ basis set is employed, and from 4391 cm<sup>-1</sup> to 834  $\rm cm^{-1}$  when one uses cc-pVTZ. When compared with the CCSDTQ data, none of these results is good yet, but before discussing the performance of various CC approaches with quadruples, let us comment on the approximate treatments of triples to see how well they do when compared to their CCSDT parent. We start with the CR-CC(2,3) approaches, since it is well established that CCSDT(T) gives wrong state ordering, placing the singlet below the triplet [315-317]. As we examine the CR-CC(2,3) calculations, we can see that, although they provide a much better description of the  $a^{1}\Sigma^{+}$  and  $X^{3}\Pi$  states of BN than CCSD, the results are strongly dependent on the variant of CR-CC(2,3) used. The CR- $CC(2,3)_A$  approach produces energies that are about 540–570 and about 800 cm<sup>-1</sup> above CCSDT for the X  ${}^{3}\Pi$  and  $a {}^{1}\Sigma^{+}$  states, respectively, when the cc-pVDZ and cc-pVTZ basis sets are employed. This imbalance in describing correlation effects for the two states results in  $a \ ^{1}\Sigma^{+} - X \ ^{3}\Pi$  gaps of about 1100 cm<sup>-1</sup>, almost 300 cm<sup>-1</sup> above the corresponding CCSDT values. Going from the CR-CC $(2,3)_A$  approach to the CR-CC $(2,3)_D$  method, the total energies of the X  ${}^{3}\Pi$  state lower by about 550 cm<sup>-1</sup> when the cc-pVDZ basis set is used and by about  $480 \text{ cm}^{-1}$  when we use the cc-pVTZ basis, bringing the resulting energies to a much better agreement with CCSDT. Unfortunately, the CR-CC(2,3) energies of the  $a \ ^{1}\Sigma^{+}$  state lower by about 900–1160 cm<sup>-1</sup> when we go from variant A to D, so much of the imbalance in electron correlation effects between the two states remains, although the CR-CC(2,3)<sub>D</sub> value of the  $a \, {}^{1}\Sigma^{+} - X \, {}^{3}\Pi$  gap obtained using the cc-pVTZ basis set, of 670  $cm^{-1}$ , is in reasonable agreement with CCSDT, which gives 834  $cm^{-1}$ . One would, however, like to improve this result. Given that  $T_1$  and  $T_2$  clusters that enter the equation defining the non-iterative corrections for the CR-CC(2,3) methods originate from the underlying poorly performing CCSD calculations, the CR-CC(2,3) approaches, even the more complete variant D, struggle. Let us then turn to the active-space CCSDt calculations, where  $T_1$  and  $T_2$  are iterated in the presence of the dominant triples, and CC(t;3) corrections. It is clear from Tables 3.9 and 3.10, that CCSDt calculations are capable of improving CCSD results, bringing them closer to a full CCSDT level compared to CCSD, confirming our expectations that  $T_1$  and  $T_2$  amplitudes obtained with CCSDt are better than those resulting from CCSD calculations, but the CCSDt energies of the X  ${}^{3}\Pi$  and  $a {}^{1}\Sigma^{+}$  states and the gap between them are still quite inaccurate. We could increase the active space to improve this situation, but we believe that it is more efficient to correct the CCSDt results for the triples outside of the "little t" set using the CC(P;Q)-based CC(t;3) methodology. This is confirmed by the results in our tables. Indeed, the  $CC(t;3)_A$  and  $CC(t;3)_D$  schemes provide an excellent agreement with the parent CCSDT results for the X  ${}^{3}\Pi$  and  $a {}^{1}\Sigma^{+}$  states of BN and adiabatic gap between them. The CC(t;3)<sub>A</sub> scheme places the X  ${}^{3}\Pi$  and a  ${}^{1}\Sigma^{+}$  states only 56–61 and 36–48  $\rm cm^{-1}$  above the corresponding CCSDT energies, when the cc-pVDZ and c-pVTZ basis sets are employed. This perfectly balanced description of the two states given by the CC(t;3)<sub>A</sub> approach leads to  $a \ {}^{1}\Sigma^{+} - X \ {}^{3}\Pi$  gaps that are a mere 8–25 cm<sup>-1</sup> below their values given by CCSDT. The CC(t;3)<sub>D</sub> scheme works slightly worse, but the overall agreement between the CC(t;3)<sub>D</sub> and full CCSDT data is still very good. The CC(t;3)<sub>D</sub> values of the  $a \ {}^{1}\Sigma^{+} - X \ {}^{3}\Pi$  gap, of 756–758 cm<sup>-1</sup>, are in excellent agreement with their CCSDT counterparts, which are 799 cm<sup>-1</sup> when the cc-pVDZ basis set is used and 834 cm<sup>-1</sup> when we use cc-pVTZ basis. The observed excellent agreement of the CC(t;3)<sub>A</sub> and CC(t;3)<sub>D</sub> schemes with the parent CCSDT method in the case of the challenging  $X \ {}^{3}\Pi$  and  $a \ {}^{1}\Sigma^{+}$  states of BN is very encouraging, since the CC(t;3) calculations replace  $\mathcal{N}^{8}$  steps of CCSDT by the much less expensive  $\mathcal{N}^{6}$ -like iterative steps of CCSDt and non-iterative  $\mathcal{N}^{7}$ -like operations associated with the CC(t;3) energy corrections.

We now turn to the various methods with quadruple excitations. As already pointed out above, one needs to incorporate  $T_4$  clusters in a robust manner to bring the  $a {}^{1}\Sigma^{+} - X {}^{3}\Pi$ gap values close to the available experimental estimates. This is because  $T_4$  effects in BN are huge. Indeed, when the cc-pVDZ basis set is employed, the differences between the CCSDT and CCSDTQ energies are 171 cm<sup>-1</sup> for the  $X {}^{3}\Pi$  state, 501 cm<sup>-1</sup> for the  $a {}^{1}\Sigma^{+}$  state, and 330 cm<sup>-1</sup> for the adiabatic  $a {}^{1}\Sigma^{+} - X {}^{3}\Pi$  gap. When we use the cc-pVTZ basis set the differences between the CCSDT and CCSDTQ energies for the  $X {}^{3}\Pi$  and  $a {}^{1}\Sigma^{+}$  states and the gap between them are 195, 663, and 468 cm<sup>-1</sup>, respectively. Let us then examine how various approximate treatments of the triples and quadruples perform.

The CR-CC(2,4) corrections to CCSD are not robust enough to provide trustworthy data. The CR-CC(2,4) approach, which we used in the earlier study [316], creates an impression that it works, bringing the  $a \ ^{1}\Sigma^{+} - X \ ^{3}\Pi$  gap values to a reasonable agreement with CCSDTQ and experiment, giving gaps in the 250-300 cm<sup>-1</sup> range, but this may be misleading, since the CR-CC(2,4)<sub>AA</sub> energies of the individual X  ${}^{3}\Pi$  and  $a {}^{1}\Sigma^{+}$  states differ from the CCSDTQ counterparts by 399 and 184  $\rm cm^{-1}$ , respectively, when the cc-pVDZ basis set is used and 458 and  $383 \text{ cm}^{-1}$ , respectively, when we use cc-pVTZ. CR-CC(2,3)<sub>DA</sub> is even worse, providing the incorrect order of both states. The CCSDtq calculations improve the CCSD and CCSDt results, but errors in the resulting energies of the  $X^{3}\Pi$  and  $a^{1}\Sigma^{+}$  states and the adiabatic gap between them are still rather large, indicating the presence of significant high-order dynamical correlation effects beyond the CCSDtq level. Indeed, when the cc-pVDZ basis set is employed, the CCSDtq energies of the X  $^{3}\Pi$  and  $a^{1}\Sigma^{+}$  states and the adiabatic gap between them deviate from the corresponding CCSDTQ data by 555, 874, and 319  $\rm cm^{-1}$ , respectively. These differences increase to 1144, 1867, and 724 cm<sup>-1</sup>, respectively, when the larger cc-pVTZ basis set is used, pointing to massive dynamical correlations that grow with the basis set size. Once again, we could try to improve the CCSDtq results using larger active spaces, but here we are more interested in the effectiveness of the CC(P;Q)based CC(t,q;3) and CC(t,q;3,4) corrections, so we turn our attention to the CC(t,q;3) and CC(t,q;3,4) approaches.

As we can see in Tables 3.9 and 3.10, the performance of the CC(t,q;3) and CC(t,q;3,4)methods is excellent. In the case of the cc-pVDZ basis set, the  $CC(t,q;3)_A$  and  $CC(t,q;3)_D$ calculations correcting the CCSDtq energies for triples outside the "little t" set reduce the 555, 874, and 319 cm<sup>-1</sup> errors relative to CCSDTQ for the energies of the X<sup>3</sup>II and  $a^{-1}\Sigma^{+}$  states, and the gap between them to 138, 193, and 55 cm<sup>-1</sup>, respectively, in the  $CC(t,q;3)_A$  case, and 47, 67, and 21 cm<sup>-1</sup>, in the case of  $CC(t,q;3)_D$ . The  $CC(t,q;3,4)_{AA}$ and  $CC(t,q;3,4)_{DA}$  approaches, which also include the quadruples outside of the "little q" set, work in a similar way, with the  $CC(t,q;3,4)_{AA}$  approach being somewhat more accurate than  $CC(t,q;3,4)_{DA}$ . The use of the larger cc-pVTZ basis set does not change these observations. The CC(t,q;3)<sub>A</sub> and CC(t,q;3)<sub>D</sub> methods reduce the 1144, 1867, and 724 cm<sup>-1</sup> errors in the energies of the X <sup>3</sup> $\Pi$  and a <sup>1</sup> $\Sigma$ <sup>+</sup> states, and the gap between them to 181, 337, and 156 cm<sup>-1</sup>, respectively, when the former approach is used, and 37, 139, and 102 cm<sup>-1</sup>, when the latter method is exploited. Once again, the CC(t,q;3,4)<sub>AA</sub> and CC(t,q;3,4)<sub>DA</sub> approaches work equally well, with the CC(t,q;3,4)<sub>AA</sub> method being most accurate, giving 45, 11, and 56 cm<sup>-1</sup> errors relative to CCSDTQ in describing the X <sup>3</sup> $\Pi$  and a <sup>1</sup> $\Sigma$ <sup>+</sup> energies and the gap between them.

In summary, we investigated the nearly degenerate  $X^{-3}\Pi$  and  $a^{-1}\Sigma^{+}$  states and the corresponding adiabatic  $a^{-1}\Sigma^{+} - X^{-3}\Pi$  gap for the difficult BN molecule. We evaluated the ability of the CR-CC(2,3), CCSDt, and CC(t;3) approaches to capture the correlation effects due connected triple excitations by comparing them against CCSDT. We were able to show that neither CR-CC(2,3) nor CCSDt provide accurate and reliable results, but when the CCSDt energies are corrected for the correlation effects due to the missing triples outside of the "little t" set via the CC(t;3) methodology, we are able to reproduce the total electronic energies of the  $X^{-3}\Pi$  and  $a^{-1}\Sigma^{+}$  states and the gap between them obtained with CCSDT to within tens of wavenumbers.

The sizable correlation contributions due to connected quadruple excitations for the  $X^{3}\Pi$ and  $a^{1}\Sigma^{+}$  states provided us with an excellent opportunity to evaluate the performance of the CR-CC(2,4), CCSDtq, CC(t,q;3), and CC(t,q;3,4) approximations, which we tested against CCSDTQ. We illustrated that the CCSDtq approach misses a significant amount of the correlation effects due to triple and quadruple excitations, especially for the larger cc-pVTZ basis set, whereas the CR-CC(2,4) methodology has difficulties balancing the two states. When the CCSDtq energies are corrected for the missing triple or the missing triple and quadruple excitations, the corresponding CC(t,q;3) and CC(t,q;3;4) methods provide the reliable and systematic behavior we were hoping to find. Both schemes recover the effects due to connected triple and quadruple excitations very well, as evidenced by the typically small errors with respect to CCSDTQ, on the order of tens of wavenumbers for the X <sup>3</sup> $\Pi$  and a <sup>1</sup> $\Sigma$ <sup>+</sup> states and the adiabatic gap between them. Given the observed accuracies and the fact that the CC(t,q;3) and CC(t,q;3,4) approaches reduce the iterative  $\mathcal{N}^{10}$  steps of CCSDTQ to iterative  $\mathcal{N}^{6}$  and non-iterative  $\mathcal{N}^{7}$  levels, and considering the challenging nature of the lowest energy singlet and triplet states of BN, we can conclude that CC(t,q;3) and CC(t,q;3,4) methods developed in this work represent and important advance in electronic structure calculations.

Table 3.9: Equilibrium bond lengths (in Å) for the lowest triplet and singlet states of BN, the corresponding adiabatic singlet-triplet splittings  $T_e$  (in cm<sup>-1</sup>), and energies relative to CCSDT and CCSDTQ (in cm<sup>-1</sup>), as obtained with the cc-pVDZ basis set.<sup>*a*</sup>

	Bond Le	ngth (Å)		Relativ	e to CCSDT	Relative	to CCSDTQ
Method	$X {}^{3}\Pi$	$a \ ^1\Sigma^+$	$T_e$	$X {}^{3}\Pi$	$a \ ^1\Sigma^+$	$X {}^{3}\Pi$	$a \ ^1\Sigma^+$
CCSD	1.3360	1.2934	4196.3	2368.9	5766.3	2540.0	6267.0
$\mathrm{CCSDt}^{b}$	$1.3487^{c}$	$1.2970^{c}$	1054.0	471.8	726.8	642.9	1227.4
$\operatorname{CR-CC}(2,3)_A$			1083.7	570.0	791.8	678.1	1292.4
$\operatorname{CR-CC}(2,3)_{\mathrm{D}}$			410.2	18.6	-370.2	189.7	130.5
$CC(t;3)_A{}^b$			791.1	55.6	47.7	226.8	548.4
$CC(t;3)_D{}^b$			757.9	-35.8	-76.9	135.4	423.7
CCSDT	$1.3491^{d}$	$1.2962^{d}$	799.0			171.1	500.7
$\mathrm{CCSDtq}^b$	$1.3496^{e}$	$1.2983^{e}$	788.3			554.7	873.5
$CR-CC(2,4)_{AA}$			254.5			399.3	184.3
$CR-CC(2,4)_{DA}$			-453.2			-57.8	-980.4
$CC(t,q;3)_A{}^b$			523.9			138.3	192.7
$CC(t,q;3)_D{}^b$			490.0			46.7	67.3
$CC(t,q;3,4)_{AA}{}^b$			371.9			33.8	-63.7
$CC(t,q;3,4)_{DA}b$			338.1			-57.8	-189.2
CCSDTQ	$1.3503^{f}$	$1.2976^{f}$	469.4				

 $^{a}$  The lowest-energy molecular orbitals that correlate with the 1s orbitals of the B and N atoms were frozen in all calculations.

 $^b$  The active space used consisted of the valence orbitals that correlate with the 2p subshells of the B and N atoms.

<sup>c</sup> The CCSDt geometries were used for the CCSDt, CR-CC(2,3), and CC(t;3) calculations.

 $^{d}$  The CCSDT geometries were optimized in this work with the cc-pVDZ basis set.

 $^e$  The CCSDtq geometries were used for the CCSDtq, CR-CC(2,4), CC(t,q;3), and CC(t,q;3,4) calculations.

f The CCSDTQ geometries were optimized in this work with the cc-pVDZ basis set.

Table 3.10: Equilibrium bond lengths (in Å) for the lowest triplet and singlet states of BN, the corresponding adiabatic singlet-triplet splittings  $T_e$  (in cm<sup>-1</sup>), and energies relative to CCSDT and CCSDTQ (in cm<sup>-1</sup>), as obtained with the cc-pVTZ basis set.<sup>*a*</sup>

	Bond Le	ngth (Å)		Relativ	e to CCSDT	Relative	e to CCSDTQ
Method	$X \ ^3\Pi$	$a \ ^1\Sigma^+$	$T_e$	$X {}^{3}\Pi$	$a \ ^1\Sigma^+$	$X \ ^3\Pi$	$a \ ^1\Sigma^+$
CCSD	1.3177	1.2732	4391.3	3188.0	6744.9	3382.9	7407.8
$\mathrm{CCSDt}^{b}$	$1.3365^{c}$	$1.2848^{c}$	1372.0	1023.2	1560.8	1218.1	2223.6
$\operatorname{CR-CC}(2,3)_A$			1095.7	544.2	805.5	739.1	1468.3
$\operatorname{CR-CC}(2,3)_{\mathrm{D}}$			670.2	68.4	-95.9	263.2	567.0
$CC(t;3)_A{}^b$			809.5	60.7	35.8	255.6	698.6
$CC(t;3)_D{}^b$			756.3	-83.3	-161.4	111.5	501.4
CCSDT	$1.3367^{d}$	$1.2826^{d}$	834.4			194.9	662.8
$\mathrm{CCSDtq}^{b}$	$1.3365^{c}$	$1.2848^{c}$	1089.9			1143.9	1867.3
$CR-CC(2,4)_{AA}$			291.8			457.8	383.1
$CR-CC(2,4)_{DA}$			-133.7			-18.1	-518.2
$\mathrm{CC}(\mathbf{t},\mathbf{q};3)_{\mathbf{A}}{}^{b}$			522.4			181.1	337.0
$CC(t,q;3)_D{}^b$			468.2			36.9	138.7
$CC(t,q;3,4)_{AA}{}^b$			310.7			44.7	-11.0
$CC(t,q;3,4)_{DA}b$			256.5			-99.4	-209.4
$\mathbf{CCSDTQ}$	$1.3367^{e}$	$1.2826^{e}$	366.4				

 $^a$  The lowest-energy molecular orbitals that correlate with the 1s orbitals of the B and N atoms were frozen in all calculations.

 $^b$  The active space used consisted of the valence orbitals that correlate with the 2p subshells of the B and N atoms.

<sup>c</sup> The CCSDt/cc-pVTZ optimized geometries were used for the CCSDt, CR-CC(2,3), CC(t;3), CCSDtq, CR-CC(2,4), CC(t,q;3), and CC(t,q;3,4) calculations.

 $^{d}$  The CCSDT optimized geometries were computed by us and replicate the previous CCSDT/cc-pVTZ geometries in Ref. [312].

 $^{e}$  The CCSDT optimized geometries were used for the CCSDTQ calculations.

## Chapter 4

# Algorithmic Advances: Efficient Automated Implementation of Active-Space CCSDtq and CCSDTq Methods, and Their Full CCSDTQ Counterpart

The CC(t;3) approach was first implemented by Dr. Jun Shen from our group who employed a spin-integrated CCSDt program he created, coupled with non-iterative corrections of the CR-CC type obtained from already efficient modified in-house programs. The spin-integrated CCSDt equations came from an automatic derivation code written by Dr. Jun Shen, which also factorized these equations and generated the corresponding FORTRAN code. The CC(t,q;3) and CC(t,q;3,4) schemes required a CCSDtq program. However, this automatic derivation process was initially developed to only handle CC methods up to CCSDT. For the double dissociation of water, the Be+H<sub>2</sub>  $\rightarrow$  HBeH insertion, and the singlet-triplet gap in the strongly biradical (HFH)<sup>-</sup> system, we were able to perform the CC(t,q;3) and CC(t,q;3,4) computations by using a pilot CCSDtq program in conjunction with the aforementioned non-iterative corrections. Unfortunately, the pilot code was written implemented using spinorbitals and the CC equations were unfactorized, and due to the memory and CPU time step requirements of the program, we quickly realized that without an efficient CCSDtq code we would be computationally restricted to systems with similar numbers of electrons and orbitals as those described above. I revised the automatic derivation code, originally written by Dr. Jun Shen, to derive the set of spin-integrated CCSDTQ equations and the initial working code, which was modified to obtain efficient CCSDtq and CCSDTq codes.

This section starts by discussing the powerful technique of spin integration for closed and open shells through an example and then describes basic rules for quickly obtaining the spin-integrated terms through diagrammatic techniques. Then we describe how the spinintegrated CCSDTQ equations are obtained using an automatic derivation routine and how these equations are automatically turned into working FORTRAN codes. After that we discuss how the reorganization of a set of loops in the resulting code reduced the number of unnecessary CPU operations and lead to a significant improvement in the speed of the code. We also describe how the CCSDTQ code was altered to obtain CCSDtq and CCSDTq codes.

### 4.1 Spin Integration for Closed and Open Shells

We have seen the structure of the CC equations in Section 3.1.1, and one thing to notice is that the Hamiltonian, consisting of one- and two-body operators, does not depend on the spin of the electrons. What spin-free operators allow us to do is integrate out the spin functions from our equations, and as a result we no longer have to carry arrays in our CC codes with all combinations of spin-orbitals, but rather arrays for spin-free operators, which have a smaller number of elements, leading to a significant reduction in the computational effort and memory when properly implemented.

This chapter makes use of diagrams, which are powerful tools to derive and organize numerous algebraic expressions that almost any accurate many-body theory generates, so we want to give a brief overview of diagrams and how to interpret them algebraically. Historically, the use of diagrams originated in quantum field theory using the time-dependent formalism. However, as advocated by Čížek and Paldus in the late 1960s and 1970s [3,4,321] (cf. [322–327] for additional remarks and further details), the time-independent formulation is sufficient in the development of quantum chemical and other many-body methods that rely on the time-independent Schrödinger equation. Diagrams are a graphical representation of Wick's theorem, which is a basic theorem for the algebraic manipulations involving operators in the second-quantized form.

The brief discussion of diagrammatic methods in this dissertation focuses on the timeindependent formulation. It is important to note that the sequence in which the operators act (i.e., right to left) is important; this is indicated in the diagram by means of a so-called formal time axis as shown below:



If we want to represent the operator product  $V_N T_1$  diagramatically, we begin with a diagrammatic representation of  $T_1$  on the right, followed by a diagram representing the operator  $V_N$  drawn to the left of the  $T_1$  diagram. The subscript N signifies that were are using the normal-order form or these operators.

In deriving the explicit algebraic expressions for the ground-state CC equations (Eqs. (3.11) and (3.12)) we obtain resulting diagrams by contractions of fermion lines represent-

ing the relevant creation and annihilation operators that enter the second-quantized forms of the operators, and applying the diagrammatic rules to convert the resulting diagrams back into algebraic expressions. We use the Hugenholtz and the corresponding Brandow diagrams [324–327] to derive the explicit many-body expressions for all terms corresponding to the ground-state CC equations (Eqs. (3.11) and (3.12)). Other representations, such as that of Goldstone, could be used as well, but Goldstone diagrams do not make use of antisymmetrized matrix elements. Hugenholtz/Brandow diagrams are preferable whenever we rely on second-quantized operators using antisymmetrized matrix elements, which is associated with fewer distinct resulting diagrams than the Goldstone representation. One other common technique when deriving the CC equations, in which we project  $(H_N e^{T^{(A)}})_C |\Phi\rangle$  in the excited determinants  $|\Phi_{i_1...i_n}^{a_1...a_n}\rangle$ , is that we do not draw the diagrams representing the bra state  $\langle \Phi_{i_1...i_n}^{a_1...a_n} |$ . Instead, we draw all permissible resulting diagrams for  $(H_N e^{T(A)})_C |\Phi\rangle$ with n incoming and n outgoing external fermion lines labeled by fixed indices  $i_1, \ldots, i_n$  and  $a_1, \ldots, a_n$ , corresponding to the determinant  $|\Phi_{i_1...i_n}^{a_1...a_n}\rangle$  on which we project [324–327]. This greatly facilitates the process of drawing the resulting diagrams and makes the diagrams much less complicated [324–327].

The diagrams may be interpreted algebraically using the following rules [321, 324–327]:

a. Each external line pointing to the left is labeled with a "particle" (unoccupied) spinorbital label  $a, b, c, d, \ldots$  and each external line pointing to the right with a "hole" (occupied) spin-orbital label  $i, j, k, l, \ldots$  In the CC equations and diagrams, external lines should always be labeled in a canonical sequence defining the particle-hole excitation in the  $|\Phi_{i_1...i_n}^{a_1...a_n}\rangle$  determinants on which we project, i.e., as a, i ; b, j ; c, k ; etc. The internal hole lines are labeled with  $m, n, \ldots$ , whereas the internal particle lines with  $e, f, \ldots$ . b. The one-body vertex representing the one-body component  $F_N = f_p^q N[a^p a_q]$  of  $H_N$ , carries the numerical value of the Fock matrix element  $\langle p|f|q \rangle = f_p^q$ , where p is an outgoing line and q is an incoming line. For example, the following figure carries a value of matrix element  $f_a^b$ .



Figure 4.1: The one-body vertex representing the one-body component  $F_N = f_p^q N[a^p a_q]$  of  $H_N$ .

c. The two-body vertex representing the two-body component  $V_N = \frac{1}{4}v_{pq}^{rs}N[a^pa^qa_sa_r]$ of  $H_N$  carries the numerical value of the antisymmetrized interaction matrix element  $v_{pq}^{rs} = \langle pq|v|rs \rangle - \langle pq|v|sr \rangle$ , where p and q are outgoing lines and r and s are incoming lines. For example, the following figure is a Brandow diagram that carries a value of  $v_{ab}^{cd} = \langle ab|v|cd \rangle - \langle ab|v|dc \rangle$ . In general,  $v_{pq}^{rs} = -v_{pq}^{sr} = -v_{qp}^{rs} = v_{qp}^{sr}$ .



Figure 4.2: The two-body vertex representing the two-body component  $V_N = \frac{1}{4}v_{pq}^{rs}N[a^pa^qa_sa_r]$  of  $H_N$ .

d. The one-, two-, three-, and four-body vertices representing the  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$  cluster operators, i.e., carry the numerical values of the  $t_a^i$ ,  $t_{ab}^{ij}$ ,  $t_{abc}^{ijk}$ , and  $t_{abcd}^{ijkl}$  amplitudes, respectively. The two-, three- and four-body amplitudes are antisymmetric with respect to the permutation of indices  $i_1, \ldots, i_n$  and  $a_1, \ldots, a_n$ .



Figure 4.3: The one-, two-, three-, and four-body vertices representing the  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$  cluster operators.

- e. All the spin-orbital labels are summed over internal lines, which are obtained by contracting the external lines of  $F_N$ ,  $V_N$ ,  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$ .
- f. The sign of the diagram is determined from  $(-1)^{l+h}$ , where l is the number of loops and h is the number of internal hole lines in a Brandow representation.
- g. The combinatorial weight factor of the connected diagram is specified by  $(\frac{1}{2})^z$ , where z is the number of pairs of "equivalent" lines. A pair of equivalent lines is defined as being two lines beginning at the same vertex and ending at another, but also same vertex, and going in the same direction. This weight rule is specific to a Hugenholtz/Brandow representation. Lines that carry fixed labels (such as the external lines defining the  $\langle \Phi_{i_1...i_n}^{a_1...a_n} |$  bra state on which we project the CC equations) are always regarded as non-equivalent.
- h. The algebraic expression for each diagram should be preceded by a suitable complete or partial antisymmetrization operator, permuting the external lines in all distinct ways to keep the full antisymmetry of a final expression for a quantity, such as the cluster

amplitudes, which are antisymmetric with respect to permutations of indices  $i_1, \ldots, i_n$ and  $a_1, \ldots, a_n$ .

While the Hugenholtz representation produces fewer distinct resulting diagrams than the Goldstone one, the Goldstone representation is useful in developing a spin-adapted formalism for spin-free Hamiltonians [321,324–326]. We will use non-antisymmetrized matrix elements and cluster amplitudes, just as with the Goldstone formulation, to carry out an example of spin-integration and to build an understanding of the technique. In the end, we will revisit the above diagrammatic rules for Hugenholtz/Brandow diagrams and extend them for the development of spin-integrated equations.



Figure 4.4: Diagram representing one of the  $\frac{1}{2}v_nt_2^2$  terms that appear in the projection of the CC equations on to doubly excited determinants.

Let us consider the term corresponding to the diagram in Figure 4.4, which represents one of the  $\frac{1}{2}v_nt_2^2$  terms that appear in the projection of the CC equations on to doubly excited determinants. If we were considering the Hugenholtz/Brandow formalism, then the complete expression for this term is

$$\mathcal{A}^{ij} v_{mn}^{ef} t_{ae}^{im} t_{fb}^{nj}, \tag{4.1}$$

where  $v_{mn}^{ef}$ ,  $t_{ae}^{im}$ , and  $t_{fb}^{nj}$  are antisymmetrized matrix elements and  $\mathcal{A}^{ij}$  is an antisymmetrizer defined as

$$\mathcal{A}^{ij} = 1 - (ij), \tag{4.2}$$

where (ij) is the transposition of indices i and j. However, for this spin integration example, we need to consider non-antisymmetrized matrix elements and also drop the antisymmetrizer  $\mathcal{A}^{ij}$ . The non-antisymmetrized integral form of the three elements in Eq. (4.1) are

$$v_{mn}^{ef} = \langle mn | \hat{v} | ef \rangle = \int \psi_m^*(x_1) \psi_n^*(x_2) \hat{v} \psi_e(x_1) \psi_f(x_2) dx_1 dx_2, \qquad (4.3)$$

$$t_{ae}^{im} = \langle ae | \hat{t}_2 | im \rangle = \int \psi_a^*(x_1) \psi_e^*(x_2) \hat{t}_2 \psi_i(x_1) \psi_m(x_2) dx_1 dx_2, \qquad (4.4)$$

$$t_{fb}^{nj} = \langle fb|\hat{t}_2|nj\rangle = \int \psi_f^*(x_1)\psi_b^*(x_2)\hat{t}_2\psi_n(x_1)\psi_j(x_2)dx_1dx_2, \tag{4.5}$$

where  $\psi_p(x_q)$  is a spin-orbital and  $x_q = \{\vec{r}_q, \omega_q\}$  is the composition of the three spatial coordinates  $(\vec{r}_q)$  and an arbitrary spin variable  $(\omega_q)$  for a given electron q. The eight indices in these expressions, each labeling a spin-orbital (indicated with a lower-case letter) can be rewritten as a product of a spatial orbital (indicated with an upper-case letter) and a spin function,

$$\psi_p(x_q) = \phi_P(\vec{r}_q)\sigma_P(\omega_q), \quad \sigma_P = \alpha \text{ or } \beta,$$
(4.6)

or in Dirac notation,

$$|p\rangle = |P\rangle \otimes |\sigma_P\rangle = |P\sigma_P\rangle. \tag{4.7}$$

We can therefore rewrite Eqs. (4.3-4.5) using this relationship as follows:

$$v_{mn}^{ef} = \langle M\sigma_M N\sigma_N | \hat{v} | E\sigma_E F\sigma_F \rangle, \qquad (4.8)$$

$$t_{ae}^{im} = \langle A\sigma_A E\sigma_E | \hat{t}_2 | I\sigma_I M\sigma_M \rangle, \qquad (4.9)$$

$$t_{fb}^{nj} = \langle F\sigma_F B\sigma_B | \hat{t}_2 | N\sigma_N J\sigma_J \rangle.$$
(4.10)

Since none of the operators depend on spin, then these integrals can be separated in integrals that contain only the spatial orbitals and integrals for the spin functions as follows:

$$v_{mn}^{ef} = \langle MN | \hat{v} | EF \rangle \langle \sigma_M \sigma_N | \sigma_E \sigma_F \rangle, \qquad (4.11)$$

$$t_{ae}^{im} = \langle AE | \hat{t}_2 | IM \rangle \langle \sigma_A \sigma_E | \sigma_I \sigma_M \rangle, \qquad (4.12)$$

$$t_{fb}^{nj} = \langle FB|\hat{t}_2|NJ\rangle \langle \sigma_F \sigma_B|\sigma_N \sigma_J\rangle.$$
(4.13)

The last integrals over spin functions in these equations will simply reduce to Kronecker deltas giving

$$v_{mn}^{ef} = \langle MN | \hat{v} | EF \rangle \delta_{ME} \delta_{NF}, \qquad (4.14)$$

$$t_{ae}^{im} = \langle AE | \hat{t}_2 | IM \rangle \delta_{AI} \delta_{EM}, \qquad (4.15)$$

$$t_{fb}^{nj} = \langle FB|\hat{t}_2|NJ\rangle\delta_{FN}\delta_{BJ}.$$
(4.16)

Two of the Kronecker deltas are repeated which leaves only four unique Kronecker deltas that have to correspond to the two spin cases ( $\alpha$  and  $\beta$ ). When we started with Eqs. (4.3)– (4.5) we had eight indices for each spin-orbital. That means if one were to naively code this term based on spin-orbitals, then for a given combination of separate spatial orbitals, each containing an  $\alpha$  and  $\beta$  spin-orbital, one would carry out 2<sup>8</sup> or 256 products for this expression. But as one separates out the spin functions and evaluates the corresponding integrals many of these combinations are simply zero and in the end we end up with a set of Kronecker deltas that correspond to the  $\alpha$  and  $\beta$  spin cases, which are less in number than the number of indices for a term. In this example, since there are only four unique Kronecker deltas, then for a given combination of separate spatial orbitals one would only carry out 2<sup>4</sup> or 16 products, a drastic reduction from the 256 products based on a spin-orbital implementation. In addition, we went from matrix elements for spin-orbitals to matrix elements that depend only on the spatial orbitals, for which there are half as many spatial orbitals as spin-orbitals. This means there is the added benefit of saving on memory in the calculations.

From our example, we can make some general observations so we can establish diagrammatic rules for quickly generating expressions which will allow us to efficiently implement the spin-integrated CC equations. The key observation from our example is that every path, whether closed or open, has a Kronecker delta, or a set of Kronecker deltas that together simplify to a single Kronecker delta, associated with the two spin cases. If we were dealing with a closed-shell system, we can further simplify our equations by taking advantage of the fact that the  $\alpha$  and  $\beta$  spin cases for each loop are equivalent. In this case we can multiply the expression by the factor  $2^l$ , where l is the number of loops in a given Goldstone diagram. Then one would simply evaluate the spin cases for the open paths corresponding to the remaining Kronecker deltas. In this thesis work, we wanted to develop a general approach so we can deal with closed- and open-shell molecules, so, we need to keep all of the spin cases for each path. So for every operator, one will need separate spin cases. Table 4.1 lists the matrix elements for the one- and two-body Hamiltonian operators and the amplitudes for the cluster operators up to  $T_4$  corresponding to the various spin cases. In Table

Spin Orbit	Spin Integrated	Antisymmetrizers
$f_p^q$	$f_p^q$	
	$f_{\widetilde{p}}^{\widetilde{q}}$	
$v_{pq}^{rs}$	$v_{pq}^{rs}$	$\mathcal{A}_{pq}\mathcal{A}^{rs}$
11	$v_{p ilde{q}}^{r ilde{s}}$	
	$v_{\widetilde{p}\widetilde{q}}^{\widetilde{r}\widetilde{s}}$	$\mathcal{A}_{pq}\mathcal{A}^{rs}$
$t^i_a$	$t^i_a$	
	$t { ilde{i}} { ilde{a}}$	
$t_{ab}^{ij}$	$t^{ij}_{ab}$	${\cal A}^{ij}{\cal A}_{ab}$
	$t_{a ilde{b}}^{ec{i ilde{j}}}$	
	$t^{\widetilde{i}\widetilde{j}}_{\widetilde{a}\widetilde{b}}$	${\cal A}^{ij}{\cal A}_{ab}$
$t_{abc}^{ijk}$	$t_{abc}^{ijk}$	$\mathcal{A}^{ijk}\mathcal{A}_{abc}$
uoc	$t^{ij\tilde{k}}_{ab\tilde{c}}$	${\cal A}^{ij}{\cal A}_{ab}$
	$t^{i\tilde{j}\tilde{k}}_{\tilde{a}\tilde{b}\tilde{a}}$	$\mathcal{A}^{jk}\mathcal{A}_{bc}$
	$t^{\widetilde{i}\widetilde{j}\widetilde{k}}_{\widetilde{z}\widetilde{1}\widetilde{z}\widetilde{1}\widetilde{z}}$	$\mathcal{A}^{ijk}\mathcal{A}_{abc}$
$t_{abcd}^{ijkl}$	$t^{ijkl}_{abcd}$	$\mathcal{A}^{ijkl}\mathcal{A}_{abcd}$
uocu	$t^{ijk\tilde{l}}_{l,\tilde{l}}$	$\mathcal{A}^{ijk}\mathcal{A}_{abc}$
	$t^{abcd}_{tj  ilde{k}  ilde{l}_{ ilde{-}}}$	$A^{ij}A^{kl}A_{ab}A_{ad}$
	$ab\tilde{c}d\ _{4}i\tilde{j}\tilde{k}\tilde{l}$	
		$\mathcal{A}^{f} \mathcal{A}_{bcd}$
	$t^{ij\kappa\iota}_{ ilde{a} ilde{b} ilde{c} ilde{d}}$	$\mathcal{A}^{\imath\jmath kl}\mathcal{A}_{abcd}$

Table 4.1: Matrix elements for the standard Hamiltonian and amplitudes for the cluster operators that appear in CCSDTQ, along with their corresponding cases for spin-integrated equations, and the respective antisymmetrizers for the spin cases.

\_

4.1, the indices for the  $\alpha$  spin do not have accents, whereas the indices for the  $\beta$  spin are differentiated by a tilde. Additionally, we recall that we have applied the spin-integration

technique to a Goldstone diagram, which correlates to non-antisymmetrized matrix elements. If we generate all necessary Goldstone diagrams for a given term and then apply the spin integration technique, then we can sum the resulting terms to give matrix elements that are antisymmetrized for indices that share the same particle-hole character and spin assignment. In conventional spin-orbital implementations, this is avoided because one can generate expressions with fully antisymmetrized matrix elements by employing Hugenholtz/Brandow diagrammatic techniques. However, in our situation, we have different spin cases for each operator, and the indices of these operators are antisymmetrized only with other indices that have the same particle-hole character and the same spin assignment. Table 4.1 lists the corresponding antisymmetrization character describing how each matrix element transforms for the different spin cases of each operator. The antisymmetrizers that appear in Table 4.1 are defined as follows:

$$\mathcal{A}_{pq} \equiv \mathcal{A}^{pq} = 1 - (pq), \tag{4.17}$$

$$\mathcal{A}_{pqr} \equiv \mathcal{A}^{pqr} = 1 - (pq) - (pr) - (qr) + (pqr) + (prq), \qquad (4.18)$$

$$\mathcal{A}_{pqrs} \equiv \mathcal{A}^{pqrs} = 1 - (pq) - (pr) - (ps) - (qr) - (qs) - (rs) + (pq)(rs) + (pr)(qs) + (ps)(qr) + (pqr) + (prq) + (qrs) + (qsr) + (rsp) + (rsp) + (rsp) + (rsp) + (spq) - (pqrs) - (pqrs) - (pqsr) - (prqs) - (prsq) - (psqr) - (psrq), \qquad (4.19)$$

where (pq) is the transposition of indices p and q, while (pqr) and (pqrs) designate the threeand four-index cyclic permutations. The other key observation is that cluster amplitudes are solved for by projecting the CC equations onto excited determinants corresponding to the specific cluster operators entering a given level of theory (see Eq. 3.11). In the case of spin integrated equations, we project onto excited determinants corresponding to the spin cases of the cluster operators in Table 4.1. For the CCSDTQ method, the set of excited determinants on which we project are listed in Table 4.2.

Table 4.2: Excited determinants entering into Eq. (3.11) for spin-orbital implementations of CCSDTQ and the corresponding spin variants for the spin-integrated implementations.

Spin-Orbit	Spin-Integrated
$\langle \Phi^a_i  $	$\langle \Phi^a_i ,\langle \Phi^{ ilde{a}}_{ ilde{i}} $
$\langle \Phi^{ab}_{ij} $	$\langle \Phi^{ab}_{ij}  , \ \langle \Phi^{a  ilde{b}}_{i  ilde{j}}  , \ \langle \Phi^{ ilde{a}  ilde{b}}_{ ilde{i}  ilde{j}}  $
$\langle \Phi^{abc}_{ijk} $	$\langle \Phi^{abc}_{ijk}  , \ \langle \Phi^{ab\tilde{c}}_{ij\tilde{k}}  , \ \langle \Phi^{ab\tilde{c}}_{i\tilde{j}\tilde{k}}  , \ \langle \Phi^{a\tilde{b}\tilde{c}}_{i\tilde{j}\tilde{k}}  $
$\langle \Phi^{abcd}_{ijkl} $	$\langle \Phi^{abcd}_{ijkl} , \ \langle \Phi^{abc\tilde{d}}_{ijk\tilde{l}} , \ \langle \Phi^{ab\tilde{c}\tilde{d}}_{ij\tilde{k}\tilde{l}} , \ \langle \Phi^{a\tilde{b}\tilde{c}\tilde{d}}_{i\tilde{j}\tilde{k}\tilde{l}} , \ \langle \Phi^{\tilde{a}\tilde{b}\tilde{c}\tilde{d}}_{\tilde{i}\tilde{j}\tilde{k}\tilde{l}} $

As mentioned above, if we sum the spin-integrated terms which use non-antisymmetrized matrix elements we can rewrite the equations in terms of partially, or in some cases fully, antisymmetrized matrix elements (see Table 4.1). Instead of considering all Goldstone diagrams, which for CCSDTQ is a discouraging task due to the sheer number of diagrams, we can reformulate the Hugenholtz/Brandow diagrammatic rules mentioned previously. Diagrams corresponding to spin-integrated equations may be obtained and interpreted using the following rules:

a. Each external line pointing to the left is labeled with a "particle" (unoccupied) label  $a, b, c, d, \ldots$  for the  $\alpha$  spin cases or  $\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}, \ldots$  for the  $\beta$  spin cases and each external line pointing to the right with a "hole" (occupied) label  $i, j, k, l, \ldots$  for the  $\alpha$  spin cases or  $\tilde{i}, \tilde{j}, \tilde{k}, \tilde{l}, \ldots$  for the  $\beta$  spin cases. In the CC equations and diagrams, external lines should always be labeled in a canonical sequence defining the particle-hole excitation in the  $\langle \Phi_{i_1 \ldots i_m \tilde{i}_{m+1} \ldots \tilde{i}_n}^{a_1 \ldots a_m \tilde{a}_{m+1} \ldots \tilde{i}_n} |$  determinants on which we project, i.e., as a, i; b, j; c, k;

etc. (see Table 4.2). The internal hole lines are labeled with  $m, n, \ldots$  (or  $\tilde{m}, \tilde{n}, \ldots$ ), whereas the internal particle lines with  $e, f, \ldots$  (or  $\tilde{e}, \tilde{f}, \ldots$ ).

- b. The one-body vertex representing the one-body component  $F_N$  of  $H_N$  is split into two spin cases:  $F_N = f_p^q N[a^p a_q]$  and  $F_N = f_{\tilde{p}}^{\tilde{q}} N[a^{\tilde{p}} a_{\tilde{q}}]$ . The former carries the numerical value of the Fock matrix element  $f_p^q = \langle p | f | q \rangle$ , while the latter carries the numerical value of the Fock Matrix element  $f_{\tilde{p}}^{\tilde{q}} = \langle \tilde{p} | f | \tilde{q} \rangle$ , where p (or  $\tilde{p}$ ) is an outgoing line and q (or  $\tilde{q}$ ) is an incoming line.
- c. The two-body vertex representing the two-body component  $V_N$  of  $H_N$  is split into three spin cases:  $V_N = \frac{1}{4} v_{pq}^{rs} N[a^p a^q a_s a_r], V_N = v_{p\tilde{q}}^{r\tilde{s}} N[a^p a^{\tilde{q}} a_{\tilde{s}} a_r]$ , and  $V_N = \frac{1}{4} v_{\tilde{p}\tilde{q}}^{\tilde{r}\tilde{s}} N[a^{\tilde{p}} a^{\tilde{q}} a_{\tilde{s}} a_{\tilde{r}}]$ . The first carries the numerical value of the antisymmetrized interaction matrix element  $v_{pq}^{rs} = \langle pq | v | rs \rangle - \langle pq | v | sr \rangle$ , the second  $v_{p\tilde{q}}^{r\tilde{s}} = \langle p\tilde{q} | v | r\tilde{s} \rangle$ , and the third  $v_{\tilde{p}\tilde{q}}^{\tilde{r}\tilde{s}} = \langle \tilde{p}\tilde{q} | v | \tilde{r}\tilde{s} \rangle - \langle \tilde{p}\tilde{q} | v | \tilde{s}\tilde{r} \rangle$ , where p (or  $\tilde{p}$ ) and q (or  $\tilde{q}$ ) are outgoing lines and r (or  $\tilde{r}$ ) and s (or  $\tilde{s}$ ) are incoming lines.
- d. The one-, two-, three-, and four-body vertices representing the  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$  cluster operators are split into several spin cases. The matrix elements carried by the different spin cases are listed in Table 4.1 along with their corresponding antisymmetrizers that describe how each element transforms with a permutation of indices.
- e. All the spin-orbital labels are summed over internal lines, which are obtained by contracting the external lines of the different spin cases of  $F_N$ ,  $V_N$ ,  $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$ .
- f. The sign of the diagram is determined from  $(-1)^{l+h}$ , where l is the number of loops and h is the number of internal hole lines in a Brandow representation.
- g. The combinatorial weight factor of the connected diagram is specified by  $(\frac{1}{2})^z$ , where

z is the number of pairs of "equivalent" lines. A pair of equivalent lines is defined as being two lines beginning at the same vertex and ending at another, but also same vertex, going in the same direction, and having the same spin assignment. If all three requirements are not met then the lines are regarded as non-equivalent. Lines that carry fixed labels (such as the external lines defining the  $\langle \Phi_{i_1...i_m\tilde{i}_{m+1}...\tilde{i}_n}^{a_1...a_m\tilde{a}_{m+1}...\tilde{i}_n} |$  bra state on which we project the CC equations (see Table 4.2) are always regarded as nonequivalent.

h. The algebraic expression for each diagram should be preceded by a suitable complete or partial antisymmetrization operator, permuting the external lines in all distinct ways to keep the full or partial antisymmetry of a final expression for a quantity defined by its corresponding spin case, such as cluster amplitudes, which are antisymmetric with respect to permutations of indices  $i_1, \ldots, i_m$  and  $\tilde{i}_{m+1}, \ldots, \tilde{i}_n$  as well as  $a_1, \ldots, a_m$  and  $\tilde{a}_{m+1}, \ldots, \tilde{a}_n$  (see Table 4.1).

These rule can greatly facilitate the process of drawing the resulting diagrams for the spin-integrated equations and allows for easy translation back into the algebraic language.

# 4.2 Automated Approach to Derivation and Computer Implementation of Coupled-Cluster Methods in Factorized Form

From the last section we learned that deriving the spin-integrated equations for CCSDTQ can be a daunting task. Compared to the four different cluster operators used in standard spin-orbital-based implementations of CCSDTQ, the spin-integrated approach requires fourteen cluster operators for the various spin cases. As a consequence, we must project the connected cluster form of the Schrödinger equation onto fourteen classes of excited determinants compared to projecting onto four classes of excited determinants corresponding to the standard CCSDTQ approach. Also, instead of two operators for the Hamiltonian, we now consider five separate operators after spin integration. If we had to derive every terms from the possible combinations from these sets of operators and projections manually, not only would it take an incredibly large amount of time, but also be error-prone and still requires the intimidating task of translating the equations into a computer program. As a note, the use of the word 'projections' here and throughout refers to the excited determinant onto which we project the connected cluster form of the Schrödinger equation  $\langle \Phi_{i_1...i_m\tilde{i}_{m+1}...\tilde{i}_n}^{a_1...a_m} \rangle$  listed in Table 3.1 for CCSDTQ.

The code that Dr. Jun Shen wrote, and which was modified by myself to allow for the development of CCSDTQ, provides a meticulously systematic way of deriving the spinintegrated CCSDTQ equations and subsequently translating the equation into ready-touse code. The automatic derivation program is designed to only take the desired level of CC theory one wants to derive, denoted by  $m_A$  (recall that  $m_A = 2$  for CCSD,  $m_A = 3$ for CCSDT,  $m_A = 4$  for CCSDTQ, etc.) as an input. From that single variable, the code is able to generate all of the terms for that level of theory by evaluating all possible combinations of the cluster and Hamiltonian operators for projections onto each of the different excited determinants corresponding to spin cases of the cluster operators. This is done by generating a fully contracted expressions (fully connected diagrams) while obeying the algebraic/diagrammatic rules of many-body methods.

The initial part of the program assigns various parameters to the operators, projections, and their corresponding indices in order to systematically track and edit them throughout the code. To start, each spin case of the cluster operators is assigned an individual number from 1 to  $n_T$ . That number is used to distinguish the cluster operators from each other and is related to the rank,  $r_T$ , of the cluster operator and to the value,  $s_T$ , which is the number of occupied and unoccupied indices that correspond to spin  $\beta$ . For example, the number 1, in the series from 1 to  $n_T$ , corresponds to the  $T_1$  operator with amplitudes  $t_a^i$ , the number 2 corresponds to the  $T_1$  operator with amplitudes  $t_{\tilde{a}}^{\tilde{i}}$ , number 3 corresponds to the  $T_2$  operator with amplitudes  $t_{ab}^{ij}$ , and so on down the list of cluster operators in Table 4.1. In general, the total number of different cluster operators  $n_T$  is

$$n_T = \sum_{r_T=1}^{m_A} \sum_{s_T=0}^{r_T} 1.$$
(4.20)

This can be verified by examining Table 4.1 ( $n_T = 5$  for CCSD,  $n_T = 9$  for CCSDT,  $n_T = 14$  for CCSDTQ). Since there are only up to two-body components in the Hamiltonian, labeled by four orbital indices, then we can only contract up to four cluster operators of which there are  $n_T$  options, plus one more for choosing not to contract with a cluster operator but the projection instead. That leaves  $(n_T + 1)^4$  combinations that must be evaluated in junction with the five spin cases for the Hamiltonian operators for each projection corresponding to the  $n_T$  cluster operators. While checking each combination the program employs a nifty trick in recognizing that any number from 0 to  $(n_T + 1)^4 - 1$  can be represented as:

$$\operatorname{Ind}_{T}(1)(n_{T}+1)^{3} + \operatorname{Ind}_{T}(2)(n_{T}+1)^{2} + \operatorname{Ind}_{T}(3)(n_{T}+1) + \operatorname{Ind}_{T}(4), \qquad (4.21)$$

where  $\operatorname{Ind}_T(i)$  is the index for the *i*<sup>th</sup> cluster operator and has values from 0 to  $n_T$ , where 0 corresponds to no cluster operator, so we can easily determine which set of spin-integrated

cluster operators characterize any of the  $(n_T + 1)^4$  combinations. Since the generating code is written to be general for any standard level of CC theory, this math devise proves to be a useful tool in many combinatorics circumstances, such as counting operators and indices, while systematically deriving the CC equations. Another observation to make is that we do not have to consider all combinations because cluster operators commute with each other, but rather those in which  $\operatorname{Ind}_T(1) \leq \operatorname{Ind}_T(2) \leq \operatorname{Ind}_T(3) \leq \operatorname{Ind}_T(4)$ .

After setting up all of the possible combinations of cluster operators, what is left is to evaluate the mix of Hamiltonian operators, cluster operators, and projections in an attempt to obtain a fully contracted expression with these three groups. However, most of these combinations can be quickly eliminated based on the number of orbital indices that label the operators and projections. For example, if the total number of orbital indices for any one of the three groups mentioned above (projection, Hamiltonian, or cluster operators) is more than half of the total number of indices for all three groups in any of the combinations, that mix can be skipped, because in order to obtain a fully contracted expression, one would have to start contracting among that group of indices, which is not allowed. So any of the three groups can contribute up to half of the total number of indices, but even then as long as there is a cluster operator in the term being considered, the number of indices from the projection cannot be half of the total number of indices, otherwise the Hamiltonian and cluster operators would have to be uncontracted, so those combinations can be tossed as well. Of the remaining possibilities, we have to start probing the hole-particle character and the spin assignments for indices before we can eliminate a term.

In order to analyze the indices and start considering contractions, we have to first design a way of carefully monitoring all necessary information about a given index. An array is allocated that has as many elements as the total number of indices for the projection,

Hamiltonian, and cluster operators associated with a particular combination. Each element of the array, which corresponds to a particular index, is assigned a three digit number. The first number tells if the index is  $\alpha$  or  $\beta$  spin ( $\alpha = 1, \beta = 2$ ). The second number tells the set of characters that describe whether the index is a particle or hole and whether it is fixed, free, or undeclared  $(1:\{a, b, c, d\}, 2:\{i, j, k, l\}, 3:\{e, f, g, h\}, 4:\{m, n, o, p\}, 5:\{u, v\}, 6:\{v, w\})$ . The third number tells which specific character in the set corresponds to that index. For example, if the index was given the label b, and corresponded to a  $\beta$  spin case, then the three digit number assigned to the particular index is 212; the first number is 2 because it corresponds to  $\beta$  spin, the second number is 1 because we are describing an unoccupied fixed index, and the third number is 2 because b is the second letter in that list of characters. All of the indices for the projection are fixed (character lists 1 and 2) and always remain fixed. Initially, all of the indices for the cluster operators are free (character lists 3 and 4), but can change to fixed indices if they correspond to a contraction with a fixed index from the projection. At the start the indices for the one- and two-body terms of the Hamiltonian operator are assigned characters from lists 5 and 6, because depending on whether it contracts with an index from the project or cluster operators determines if the index corresponds to a particle or hole and if it is fixed or free.

Now that the code has identified all possible combinations of cluster operators, eliminated those that obviously do not appear in the CC equations, and then assigned the indices to the remaining combinations, what is left is to carry out Wick's theorem on the different combinations. The program systematically and carefully caries out the contractions in an attempt to create a fully contracted expression. It is careful to consider the necessary permutations of indices to retain the antisymmetry of the cluster amplitude being calculated. The program tracks the sign and weight of the term, and if it can form a fully contracted

expression, then that term is saved by writing it to a Latex file for the given projection for easy viewing. Figure 4.5 shows an example of the output by the code for a specific  $H_N T_3$ combination that appears in the projection of the connected cluster form of the Schrödinger equation onto  $\langle \Phi_{i\tilde{j}}^{a\tilde{b}} |$  for CCSDTQ. When writing the terms out, the cluster operators are labeled with their ranks, but also a letter, which is necessary for successive alterations of the terms. The letter A indicates that all indices correspond to the  $\alpha$  spin. Each successive character after A coincide with replacing an occupied and unoccupied pair of indices by the  $\beta$  complement, just as illustrated in the order of matrix elements listed in Table 4.1. As an example, 't\_ $\{3B\}$ (aebimj)', which is found in Figure 4.5, corresponds to the cluster amplitude  $t_{ae\tilde{b}}^{imj}$ . If there is more than one cluster operator for a term, then they are written successively, one after the next. The different spin cases for the one- and two-body matrix elements of  $H_N$  are also distinguished when written to the Latex file. The one-body matrix elements  $f_p^q$  and  $f_{\tilde{p}}^{\tilde{q}}$  are written out to the Latex file in the form "F\_{qp}", and for the case where both indices correspond to the  $\beta$  spin this term is followed by a letter 'B' that is commented out in order to distinguish it. The two-body matrix element  $v_{p\tilde{q}}^{r\tilde{s}}$  is written out in the form "< pq|rs >", while the matrix elements  $v_{pq}^{rs}$  and  $v_{\tilde{p}\tilde{q}}^{\tilde{r}\tilde{s}}$  are written to the Latex file with the form "< pq||rs >", where once again, if all the indices followed to the  $\beta$  spin then this term is proceeded by a letter 'B' that is commented out. With these formats, a term can then be picked up any time and indices and their corresponding spin can be assigned instantly for the operators.

If we were to naively program the various terms of the CC equations derived by the code at this point, i.e., one by one and with the explicit loops what correspond to summations over indices that label internal lines, we would have a CCSDTQ routine that has CPU time step that scale as  $\mathcal{N}^{12}$ , rather than  $\mathcal{N}^{10}$ , which is obtained with an efficient implementation

```
1 \begin{eqnarray}
          2 <\Psi_{i\tilde{j}}^{a\tilde{b}}|\hat{H}|\hat{T}_{3B}\Psi_0>
           3 +\sum_{em}
          4 F_{me}
          5 t_{3B}(aebimj)
          7 -\int \{1\} \{2\} \setminus \{
          8 <mn||ie>
          9 t_{3B}(aebmnj)
10 \nonumber \\&&
11 +\frac{1}{2}\sum_{efm}
12 <am||ef>
13 t_{3B}(efbimj)
14 \nonumber \\&
15 - \sum \{emn\}
16 <mn|ej>
17 t_{3B}(aebimn)
18 \nonumber \\&&
19 + \sum_{efm}
20 <mb|ef>
21 t_{3B}(aefimj)
22 \nonumber \\&
23 \setminus end \{eqnarray\}
```

Figure 4.5: Sample output of the Latex files corresponding to the  $H_N T_{3B}$  term generated by the automatic derivation code. See text for a description of the notation.

of CCSDTQ. In order to obtain the lower-order scaling the equations have to be factorized in which the equations are rewritten in terms of binary tensor products. The program that derives the equations also performs this factorization procedure, breaking down the terms into binary products. When there are multiple cluster operators contracted with the Hamiltonian operators, the program will form the product between the Hamiltonian and cluster operator with the most computationally intensive summations over indices that label internal lines first. For example, for the term  $\mathcal{A}^{ij}\mathcal{A}_{ab}v_{mn}^{ef}t_{ae}^{im}t_{f}^{j}t_{b}^{n}$  the product between  $v_{mn}^{ef}$  and  $t_{ae}^{im}$  would be carried out first, since the summation over indices e and m is more demanding then either the summation over f corresponding to the second cluster operator or the summation over
*n* corresponding to the third cluster operator. The result of the product of  $v_{mn}^{ef}$  and  $t_{ae}^{im}$  is the intermediate  $x_{an}^{if}$ . This process is repeated, forming a product between the intermediate and cluster operator corresponding to the next most computationally intensive summation, until the last cluster operator in the series is reached. At this point, the program will look for other terms that share the same cluster operator and same antisymmetrizers applied to the term (in this example the antisymmetrizers are  $\mathcal{A}^{ij}$  and  $\mathcal{A}_{ab}$ ), and before carrying out the product, the program sums the corresponding intermediates and Hamiltonian elements for the terms that only have one cluster operator into a collective intermediate in order to reduce the number of operations that would otherwise be carried out.

After obtaining the factorized spin-integrated CC equations, then the set of terms is converted into FORTRAN routines that can be read into a code predesigned for the iterative procedure. Every product of two operators determined through the factorization is carried out in the same general manner which is laid out in Figure 4.6. The first steps are to reorganize the indices labeling the two operators of the binary product in order to take advantage of efficient linear algebra libraries that can be intrinsically parallelized making use of shared-memory systems. We want to take advantage of the fast vector-matrix and matrix-matrix multiplication routines, but in order to do so every array of a Hamiltonian, intermediate, or cluster operator has to be arranged as either a vector (one-dimensional array) or matrix (two-dimensional array). If all of the indices of an operator in a given product are internal indices that are summed over, one can represent it as a vector whose size is the product of the individual dimensions of the original array. So, for example, the  $f_m^e$  matrix elements in the product  $f_m^e t_{ae}^{im}$  term would be transformed to a vector in this procedure with the dimension  $n_{o,\alpha} \times n_{u,\alpha}$  corresponding to the number of elements for indices e and m. If the indices of an operator are both internal and external, one can represent it as as a matrix, where one dimension of the matrix is for the internal indices being summed over and has an extent equal to the product of each dimension corresponding to the internal indices. The other dimension of the matrix is for the external indices and has an extent equal to the product of each dimension corresponding to the external indices. For example, the  $t_{ae}^{im}$  matrix elements in the product  $f_m^e t_{ae}^{im}$  term would be transformed to a matrix, where one dimension corresponds to indices e and m and has an extent equal to  $n_{o,\alpha} \times n_{u,\alpha}$ , and the other dimension of the matrix corresponds to indices a and i and has an extent equal to

 $n_{o,\alpha} \times n_{u,\alpha}$ .

```
1
     ALLOCATE array A
2
     CALL "REORDER" routine for array A
3
     ALLOCATE array B
     CALL "REORDER" routine for array B
4
5
     ALLOCATE array C
6
     Declare I1
7
     Declare I2
8
     Declare I3
9
     CALL DGEMM or DGEMV routine
10
     DEALLOCATE array A
11
     DEALLOCATE array B
12
     CALL "SUM" routine(s)
13
     DEALLOCATE array C
```

Figure 4.6: The general structure of each binary tensor product as produced by the automatic derivation and implementation program.

For a more specific description of how each product of two elements is carried out we can examine the general layout of the code for a binary product (Figure 4.6) and an actual example of the code (Figure 4.7) in detail. For reasons that will soon be realized, the operator arrays must be arranged such that the first dimensions are for internal indices of that operator and the remaining dimensions are for the external indices of the operator. The order of internal and external indices among themselves does not matter, with one exception.

```
ALLOCATE (D1 (N0+1:N1,N0+1:N1,N1+1:N3,N0+1:N1))
1
\mathbf{2}
     CALL REORDER1243 (NO, N3, NO, N3, NO, N3, NO, N3,
   & NO, N1, NO, N1, N1, N3, NO, N1, IntR, D1)
3
4
     ALOCATE (D2 (N0+1:N1,N0+1:N1,N1+1:N3,N1+1:N3))
5
     CALL REORDER4312(N1,N3,N1,N3,N0,N1,N0,N1,
6
   & NO, N1, NO, N1, N1, N3, N1, N3, t2A, D2)
7
     ALLOCATE (S14 (N1+1:N3,N1+1:N3,N1+1:N3,N0+1:N1))
8
     I1 = K1 * K3
9
     12 = K3 * K3
10
     I3 = K1 * K1
11
     CALL DGEMMROUT (I1, I2, I3, D1, D2, S14)
12
     DEALLOCATE (D1)
13
     DEALLOCATE (D2)
14
     CALL SUM2314(N1,N3,N1,N3,N1,N3,N0,N1,X2,S14, 0.500)
15
     DEALLOCATE (S14)
```

Figure 4.7: Example of binary tensor product between a two-body Hamiltonian operator and a  $T_{2A}$  cluster operator that produces and intermediate labeled S14, which is summed with other intermediates and Hamiltonian matrix elements in a collective intermediate labeled X2.

Since the internal indices that are being summed over are common between the two operators in the product term, the order of internal indices in both arrays have to be the same. In order to rearrange the elements to achieve this reordering, we take a trivial approach and first allocate a new array with the same rank and size as the original, and then call a subroutine "REORDER" that copies the elements in the desired order in the new array. An example of one of the "REORDER" routines can be seen in Figures 4.8. The elements of the first and second operator are respectively copied to array "A" and "B" in the desired order. The result of the binary tensor product is stored in array "C" which is allocated after the reordering of indices for the two operators is complete. At this point, arrays "A", "B" and "C" still range in rank from two- to eight-dimensions and still need to transformed down to one- or two-dimensional arrays. Luckily, this can be accomplished effortlessly through the process of calling a subroutine that is a wrapper for calling fast vector-matrix or matrix-matrix

```
1
     SUBROUTINE REORDER4312(M1,N1,M2,N2,M3,N3,M4,N4,
2
   & K4,L4,K3,L3,K1,L1,K2,L2,A,B)
3
     REAL *8 A (M1+1:N1, M2+1:N2, M3+1:N3, M4+1:N4)
4
     REAL*8 B(K4+1:L4,K3+1:L3,K1+1:L1,K2+1:L2)
5
     REAL*8 C
6
7
     DO I1 = K1 + 1, L1
8
        DO I2=K2+1, L2
9
          DO I3=K3+1,L3
10
             DO I4=K4+1,L4
               B(I4, I3, I1, I2) = A(I1, I2, I3, I4)
11
12
             ENDDO
13
          ENDDO
14
        ENDDO
15
     ENDDO
16
     END
```

Figure 4.8: An example of one of the "REORDER" routines. The number in the title of the routine reflects the order that the indices are permuted to.

multiplication routines. In the subroutine, arrays "A", "B", and "C" are declared as one- or two-dimensional arrays so the program will automatically associate the elements of "A", "B", and "C", which are multidimensional outside the subroutine, with one- or two-dimensional arrays inside the subroutine by simply calling the subroutine. In order to do this the code has to know what are the dimensions of the one- or two-dimensional arrays for which arrays "A", "B", and "C" will translated into, so three numbers are declared (I1, I2, and I3). The variable I1 is the dimension for the external indices of array "A", I2 is the dimension for the external indices of array "B", and I3 is the dimension for the internal indices of the two arrays. If either array "A" or "B" has no external indices, then the corresponding variable is not declared. After the vector-matrix or matrix-matrix multiplication is carried out, then arrays "A" and "B" are deallocated since they are no longer needed. The results of the binary tensor product may be added with other intermediates if further products with additional cluster operators are needed for the term, or is added to array for the newest iteration cluster

```
SUBROUTINE SUM2314 (K1, L1, K2, L2, K3, L3, K4, L4, A, B, C)
1
2
      REAL*8 A(K1+1:L1,K2+1:L2,K3+1:L3,K4+1:L4)
3
      REAL *8 B(K2+1:L2,K3+1:L3,K1+1:L1,K4+1:L4)
4
      REAL C
5
6
      DO I1 = K1 + 1, L1
7
        DO I2 = K2 + 1, L2
8
          DO I3=K3+1,L3
9
             DO I4 = K4 + 1, L4
10
               A(I1,I2,I3,I4) = A(I1,I2,I3,I4) + C * B(I2,I3,I1,I4)
11
             ENDDO
12
          ENDDO
13
        ENDDO
14
      ENDDO
15
      END
```

Figure 4.9: An example of one of the "SUM" routines. The number in the title of the routine reflects the order that the indices are permuted to.

amplitudes if the product is with the last cluster operator in the series of cluster operators for a term or set of terms, and this is done by calling a "SUM" subroutine. An example of a "SUM" routine can be seen in Figure 4.9. If the result of the product is added to the array for the newest iteration of the cluster amplitude, then a series of "SUM" routines is called the necessary amount of times to satisfy the antisymmetrizer that precedes the term in order to retain the antisymmetry of the cluster amplitude being updated (see Table 4.1). With this general layout, it is easy to pick up any of the binary tensor products and instantly generate the corresponding lines of FORTRAN code.

## 4.3 Improvements in Efficiency via Loop Reorganization

The generalized layout of each product of two arrays, illustrated in Fig. 4.6, is an easy way to quickly generate the code, but restricting oneself to just this format reduces efficiency, and there is one particular situation where efficiency is greatly effected by this form. After a binary tensor product is carried out, the "SUM" routine adds the result to another array and this is done for all elements of the array and a coupled of options may happen. If the result of the product is an intermediate that can be summed with other intermediates, then this routine is only called once. Another possibility is that the intermediate formed by the product may not be summed with others, but still used in later terms, and therefore the "SUM" routine is not necessary. However, if the product is between a Hamiltonian operator, intermediate, or sum of intermediates and the last cluster operator for a term or set of terms in the factorized CC equations, then the "SUM" routine is called as many times as necessary to satisfy the antisymmetrizer associated with those terms, as the result is added to the array for the newest iteration of the cluster operators. The "SUM" routine in this instance is called anywhere from one to over a hundred times in CCSDTQ, depending on the term, and each time the entire array corresponding to the result of the product just carried out is added to the array for the newest iteration of the cluster operator being calculated. But, we recognized that not all indices need to be updated for every term. Rather, we can consider updating only the unique terms based on the antisymmetry of the cluster operator and then employ that antisymmetry at the end of the routine to obtain the remaining cluster amplitudes. For example, if consider updating the  $t_{ijkl}^{abcd}$  amplitudes, then rather update all values of i, j, k, land a,b,c,d we can choose to update just those that satisfy the inequality  $i\,<\,j\,<\,k\,<\,l$  and a < b < c < d, then replicate these amplitudes for the remaining indices at the end of the routine, while being mindful of their antisymmetry. In this example, the number of operators for each summation is reduced by a factor of  $(4!)^2$  or 576. The different spin cases for the cluster operators have different corresponding antisymmetrizers that must be taken into account when deciding to update just the unique terms. As another example, if we were to update the  $t_{ij\tilde{k}\tilde{l}}^{ab\tilde{c}\tilde{d}}$  amplitudes, then the program only needs to update indices i < j,  $\tilde{k} < \tilde{l}$ , a < b, and  $\tilde{c} < \tilde{d}$  the number of operators for each summation is reduced by a factor of  $(2!)^4$  or 16. An example of this restructuring can be seen in Figure 4.10 where the "SUM" routines are replaced by explicit DO loops over the unique elements (in this case i < j < land a < b < c corresponding to the  $t_{ijk\tilde{l}}^{abc\tilde{d}}$  amplitudes). This technique was put into action only for the binary products formed with the last cluster operators for a term, because those instances are the ones for which the result of the product is defined only by fixed indices corresponding to the excited determinant bra state which we project the connected form of the Schrödinger equation onto. In addition, this technique was applied to just the update routines for the projections onto quadruply excited determinants. These cases are where the code benefits the most from this rewriting of the "SUM" routines since we can take the most use of the antisymmetry of the cluster operators and also because the "SUM" routines are usually called several times in these sections of the code. In the following section, when we discuss transforming the CCSDTQ routine to obtain the CCSDtq and CCSDTq codes, we incorporate this technique as we rewrite the binary product in a completely different way than the general layout presented in Figure 4.6.

The benefit of this restructuring can be seen in Table 4.3, where we compare timings of calculations for the  $X \ ^1\Sigma_g^+$  and  $A \ ^3\Sigma_u^+$  states of (HFH)<sup>-</sup> before and after the changes in the "SUM" routines. Timing obtained using NWCHEM are also included for comparison.

```
ALLOCATE (F1 (
1
2
  & NO+1:N1,N2+1:N3,NO+1:N2,N2+1:N3,NO+1:N2,NO+1:N1))
3
     CALL REORDER546123 (N2, N3, N0, N2, N0, N1, N2, N3, N0, N1, N0, N2,
4 & NO, N1, N2, N3, NO, N2, N2, N3, NO, N2, NO, N1, U41, F1)
5
     ALLOCATE (H2 (N0+1:N1, N2+1:N3, N0+1:N2, N1+1:N3,
6 & N1+1:N3,N1+1:N3,N0+1:N1,N0+1:N1))
7
     CALL REORDER61523478 (N2, N3, N1, N3, N1, N3, N1, N3,
8 & NO,N2,NO,N1,NO,N1,NO,N1,NO,N1,N2,N3,NO,N2,N1,N3,
9 & N1, N3, N1, N3, N0, N1, N0, N1, t4B, H2)
     ALLOCATE (Z140 (N1+1:N3, N1+1:N3, N1+1:N3, N0+1:N1,
10
11 & NO+1:N1,N2+1:N3,NO+1:N2,NO+1:N1))
12
     I1 = K1 * K2 * K4
13
     I2=K1*K1*K3*K3*K3
14
     I3 = K2 * K4 * K1
     CALL DGEMMROUT (I1, I2, I3, F1, H2, Z140)
15
16
     DEALLOCATE (F1)
17
     DEALLOCATE(H2)
18
     D0 i=N0+1,N1-2;D0 j=i+1,N1-1;D0 k=j+1,N1;D0 l=N0+1,N2
       DO a=N1+1,N3-2;DO b=a+1,N3-1;DO c=b+1,N3;DO d=N2+1,N3
19
20
          V4B(d,c,b,a,l,k,j,i) = V4B(d,c,b,a,l,k,j,i)
21 &
          -Z140(c,b,a,k,j,d,l,i)
         +Z140(c,b,a,k,i,d,l,j)
22 &
23 &
          -Z140(c,b,a,j,i,d,l,k)
24
       ENDDO; ENDDO; ENDDO; ENDDO
25
     ENDDO; ENDDO; ENDDO; ENDDO
26
     DEALLOCATE (Z140)
```

Figure 4.10: An example of the product between an intermediate labeled U41 and the cluster operator  $T_{4B}$ , which is projected onto the excited determinant  $\langle \Phi_{ijk\bar{l}}^{abc\bar{d}} |$ . The array corresponding to the product formed is labeled Z140 and is added to the residual named V2B.

For the original CCSDTQ program, before the "SUM" routines were changed, the iterations took an average of 2491 and 3601 seconds for the  $X \, {}^{1}\Sigma_{g}^{+}$  and  $A \, {}^{3}\Sigma_{u}^{+}$  states, respectively. As a comparison, for NWCHEM the iterations took an average of 1888 and 2068 seconds, so the original implementation of CCSDTQ was between 1.5 and 1.8 times slower than NWCHEM for these set of calculations. This changes significantly by simply rewriting the "SUM" routines as described above. After the technique described above is implemented

Table 4.3: Average iteration time, in seconds, for calculations of the  $X \, {}^{1}\Sigma_{g}^{+}$  and  $A \, {}^{3}\Sigma_{u}^{+}$  states of (HFH)<sup>-</sup>, computed using our original implementation of CCSDTQ, the improved version of CCSDTQ, and NWCHEM for comparison.<sup>*a*</sup>

Program	$X \ ^1\Sigma_g^+$	$A {}^{3}\Sigma_{u}^{+}$
CCSDTQ (Original)	2941	3601
CCSDTQ (Improved)	553	558
NWCHEM	1888	2068

 $^{a}$  Calculations were performed with the 6-31G(d,p) basis set on a single computer core and the core electrons were frozen in all calculations.

the iteration reduced to an average of 553 and 558 seconds, respective. So the new version of the spin-integrated CCSDTQ approach is between 5.3 and 6.5 times faster than the original implementation and is 3.4–3.7 times faster than NWCHEM for these set of calculations, and is generally true for other calculations as well. Its important to note since the code is general for open- and closed-shell molecules, that the calculations for the  $X^{-1}\Sigma_{g}^{+}$  still go trough same number of routines as the open-shell calculation, even though many of the terms are redundant. As discussed in Section 4.1, if we were dealing with a closed-shell system, we can simplify our equations by taking advantage of the fact that the  $\alpha$  and  $\beta$ spin cases for each loop are equivalent. Despite going through extra steps, the CCSDTQ code with the updated "SUM" routines is still faster than NWCHEM. When we translated the CCSDTQ code to the obtain the CCSDtq and CCSDTq methods, we abandon calling the "SUM" routines, but still retain the lesson we learned from this restructuring, which is discussed in the following section.

# 4.4 Transformation from CCSDTQ to the CCSDtq and CCSDTq Methods

The automatic derivation process described in the previous sections was done for the full treatment of triply- and quadruply-excited cluster operators. In other words, the binary tensor product for two operators is carried out for all indices corresponding to the operators. In order to obtain the active-space CCSDtq and CCSDTq methods we had to find a mechanism for reducing the number of operators so as not to carry out the product for all indices, but rather those that obey the active-space logic set in place by the user. To do this, we abandoned the general layout seen in Figures 4.6 and 4.7, and replace the code with explicit DO loops. This is only done for the binary products involving the last cluster operator in the set of cluster operators for a given term, just as previous done with the "SUM" routines described in the last section. Examples of the explicit loops for projections onto triples and quadruples can be seen in Figures 4.11 and 4.12 respectively.

There are several advantages to rewriting the products, which again are only those that involve the last cluster operator in the set of cluster operators for a given term, in the format seen in Figures 4.11 and 4.12. By structuring the products with DO loops, we can easily introduce logic statements (lines 2 and 4 in Figures 4.11 and 4.12) that allow us to bypass the product if a given set of indices do not obey the active-space logic set by the user. Before running the calculation, the user defines the active space by declaring the number of the occupied and unoccupied  $\alpha$  and  $\beta$  electrons and orbitals in the active space along with how many indices for the triple and quadruple excitations they want to restrict to only belonging to the active space. The number of restricted occupied and unoccupied indices for the triple excitations is declared separately from the number occupied and unoccupied restricted

```
1
     DO i=NO+1,N1-2;DO j=i+1,N1-1;DO k=j+1,N1
2
       IF(indocc(k,j,i).eq.1)CYCLE
3
         DO a=N1+1,N3-2;DO b=a+1,N3-1;DO c=b+1,N3
4
            IF(indunocc(c,b,a).eq.1)CYCLE
5
            SUM = 0.0d0
6
              DO e = N1 + 1, N3
7
              SUM=SUM
8
              X2(e,c,b,i)*t2A(e,a,k,j)
   &
           -
              X2(e,c,a,i)*t2A(e,b,k,j)
9
   &
           +
10 &
              X2(e,b,a,i)*t2A(e,c,k,j)
           -
              X2(e,c,b,j)*t2A(e,a,k,i)
11
  &
           +
              X2(e,c,a,j)*t2A(e,b,k,i)
12 &
           -
              X2(e,b,a,j)*t2A(e,c,k,i)
13 &
           +
              X2(e,c,b,k)*t2A(e,a,j,i)
14 &
           -
15 &
              X2(e,c,a,k)*t2A(e,b,j,i)
           +
              X2(e,b,a,k)*t2A(e,c,j,i)
16 &
17
              ENDDO
18
             V3A(c,b,a,k,j,i) = V3A(c,b,a,k,j,i) + SUM
19
          ENDDO; ENDDO; ENDDO
20
     ENDDO: ENDDO: ENDDO
```

Figure 4.11: An example of the product between an intermediate labeled X2 and the cluster operator  $T_{2A}$ , which is projected onto the excited determinant  $\langle \Phi_{ijk}^{abc} |$ .

indices for the quadruple excitations, giving the user flexibility in their choice of the particular active-space approach. For example, if one occupied and unoccupied index is restricted to the active space for triple excitations and two occupied and unoccupied indices are restricted to the active space for quadruple, we obtain the CCSDtq approach presented in Section 3.1.2. If the active space includes all orbitals for the triple excitations and one occupied and unoccupied index is restricted to the active space for quadruple excitations, we obtain the CCSDTq approach. This flexibility also allows us to consider "in between" approximations. For example, one may consider only restricting one occupied and unoccupied index to the active space for the triple excitations, while restricting three occupied and unoccupied indices to the active space for quadruple excitations, while restricting three occupied and unoccupied indices

```
1
     D0 i=N0+1,N1-3;D0 j=i+1,N1-2;D0 k=j+1,N1-1;D0 l=k+1,N1
2
       IF(indocc(l,k,j,i).eq.1)CYCLE
3
         D0 a=N1+1,N3-3;D0 b=a+1,N3-2;D0 c=b+1,N3-1;D0 d=c+1,N3
4
            IF(indunocc(d,c,b,a).eq.1)CYCLE
5
            SUM = 0.0d0
6
              DO m=NO+1, N1; DO n=NO+1, N1
7
              SUM=SUM
8
   &
         +
              (S5(j,m,i,n)*t4A(d,c,b,a,n,m,l,k)
9
              S5(k,m,i,n)*t4A(d,c,b,a,n,m,l,j)
         _
  &
              S5(1,m,i,n)*t4A(d,c,b,a,n,m,k,j)
10 &
         +
              S5(i,m,j,n)*t4A(d,c,b,a,n,m,l,k)
11 &
         -
12 &
         +
              S5(i,m,k,n)*t4A(d,c,b,a,n,m,l,j)
13 &
         _
              S5(i,m,l,n)*t4A(d,c,b,a,n,m,k,j)
14 &
              S5(k,m,j,n)*t4A(d,c,b,a,n,m,l,i)
         +
              S5(1,m,j,n)*t4A(d,c,b,a,n,m,k,i)
15 &
         -
16 &
         -
              S5(j,m,k,n)*t4A(d,c,b,a,n,m,l,i)
17 &
              S5(j,m,l,n)*t4A(d,c,b,a,n,m,k,i)
         +
18 &
              S5(l,m,k,n)*t4A(d,c,b,a,n,m,j,i)
         +
19 &
         _
              S5(k,m,l,n)*t4A(d,c,b,a,n,m,j,i))/2.0d0
20
              ENDDO; ENDDO
            V4A(d,c,b,a,l,k,j,i)=V4A(d,c,b,a,l,k,j,i)+SUM
21
22
         ENDDO; ENDDO; ENDDO; ENDDO
23
     ENDDO; ENDDO; ENDDO; ENDDO
```

Figure 4.12: An example of the product between an intermediate labeled S5 and the cluster operator  $T_{4A}$ , which is projected onto the excited determinant  $\langle \Phi_{ijkl}^{abcd} |$ .

excitations play a significant role and we want to include a large subset of them in our activespace calculation, while at the same time incorporating a very small subset of quadruples in the active-space consideration. After inputting this information, the program constructs two arrays for the occupied and unoccupied indices which tell if a particular combination of occupied and unoccupied indices is allowed based on the active-space defined by the user. If the combination of occupied indices does not meet the active space requirements, then the loops cycle to the next combination (line 2 in Figures 4.11 and 4.12), and if the combination of unoccupied indices does not meet the active space requirements, then the loops cycle to the next combination (line 4 in Figures 4.11 and 4.12). Only when sets of indices satisfy

the active space is the binary product carried out. There are other benefits to this new structure as well. From our earlier lesson with the "SUM" routines discussed in Section 4.3, we know we do not have to update all values of the indices, but just the unique ones, based on the antisymmetry of the of the amplitudes we are updating. Figure 4.11 is an example of a product in the update routine for the  $t_{ijk}^{abc}$  amplitudes, which computed using just the set of indices i < j < k and a < b < c, while Figure 4.12 is an example a product in the update routine for the  $t^{abcd}_{ijkl}$  amplitudes, which computed using just the set of indices i < j < k < land a < b < c < d. In addition, we incorporate the "SUM" routines in the loop by explicitly summing the product of the two tensors with the other products corresponding to applying the antisymmetrizer that is needed to retain the antisymmetry of the newest iteration of the amplitudes being calculated. Another big benefit of this new format is that we remove the need to reorder the indices, therefore we no longer need to allocate arrays "A", "B", and "C". As a result of not having to allocate these arrays, the memory needed to carry out this product is reduced by more than half of the amount. Lastly, since this loop format is primitive, it allows us to easily introduce other aspects such as symmetry adaptation and explicit parallelization in future development work.

### Chapter 5

## **Conclusions and Future Outlook**

In this dissertation, we presented the CC(P;Q) formalism, which provides a systematic approach to correcting energies obtained in the active-space CC and EOMCC calculations that recover much of the non-dynamical and some dynamical many-electron correlation effects for the remaining, mostly dynamical, correlation effects missing in the active-space CC and EOMCC considerations. We discussed the details of the CC(t;3), CC(t,q;3), CC(t,q;3,4), and CC(q;4) methods, which use the CC(P,Q) formalism to correct energies obtained with the CC and EOMCC approaches with singles, doubles, and active-space triples (CCSDt/EOM-CCSDt) for missing triple excitations (CC(t;3)), or to correct energies obtained with the CC and EOMCC approaches with singles, doubles, and active-space triples and quadruples (CCSDtq/EOMCCSDtq) for missing triples (CC(t,q;3)) or missing triples and quadruples (CC(t,q;3,4)), and to correct energies obtained with the CC and EOMCC approaches with singles, doubles, triples, and active-space quadruples (CCSDTq/EOMCCSDTq) for correlation effect due to the missing quadruple excitations (CC(q;4)). By examining the double dissociation of water, the Be +  $H_2 \rightarrow HBeH$  insertion, and the singlet-triplet gaps in the strongly biradical (HFH)<sup>-</sup> system and the BN molecule, we further established that the previously developed CC(t;3) scheme is capable reproducing the total and relative energies obtained with the parent CCSDT approach to within fractions of a millihartree at a tiny fraction of the computer cost. In addition, we were also able to show, for the first time, that the CC(t,q;3) and CC(t,q;3,4) methods accurately and reliably reproduce the total and relative energies obtained with the parent CCSDTQ approach, once again to within fractions of a millihartree, and at a tiny fraction of the computer cost, even when the electronic quasi-degeneracies become substantial.

In our future endeavors, we would like to continue examining other difficult molecular systems to further test the performance of the CC(P;Q) methods, especially those involving quadruple excitations, and in particular the CC(q;4) scheme which was not investigated in this study. Current efforts are underway to study the performance of all of the CC(P;Q)approaches presented in this dissertation by examining the potential energy surface for the beryllium dimer, Be<sub>2</sub>. These methods also provide the opportunity to study other multireference molecular problems involving biradicals, bond breaking, and other instances characterized by quasi-degenerate states for which the full treatment of higher-order correlation is too expensive. The future application of the CC(P;Q) are contingent on having efficient algorithms, especially for the underlying active-space calculations which continuously being improved.

In this study we described how the spin-integrated CCSDTQ equations were derived, factorized, and translated into FORTRAN code using a program that carries out these procedures automatically. We learned that the performance of the resulting program could be significantly improved by reorganizing some loops, removing unnecessary operations. Still, there is a lot of room for further improvement. In the future we would like to examine the code for additional unnecessary operations and determine other bottlenecks in the program. For example, in the case of closed-shell systems, we can take advantage that the  $\alpha$  and  $\beta$ spin cases for each loop are equivalent and generate a separate and efficient code for just closed-shell calculations. The new loop structure introduced for the active-space CCSDtq and CCSDTq also provides the opportunity for easy further improvement. For example, we would like to introduce spatial symmetry-adaptation to further speed up the code. We would also like to introduce explicit parallelization to take advantage of shared and distributed memory systems, which the DO loop structure of the active-space CCSDtq and CCSDTq methods allows us to implement more easily. Finally, we want to address the fact that the  $T_3$ and  $T_4$  cluster operators are allocated in memory for all values of the indices that correspond to the operators, not just those that satisfy the active-space logic set by the user. In other words, the memory requirements for the CCSDtq and CCSDTq calculations are the same as the full CCSDTQ method. In the future, we want to replace the arrays for  $T_3$  and  $T_4$  that contain all values of the indices that correspond to the operators with those that contain only the elements defined by the active-space opening up many more applications of the CC(P;Q) methods.

#### BIBLIOGRAPHY

#### BIBLIOGRAPHY

- [1] F. Coester, Nucl. Phys. 7, 421 (1958).
- [2] F. Coester and H. Kümmel, Nucl. Phys. 17, 477 (1960).
- [3] J. Cížek, J. Chem. Phys. 45, 4256 (1966).
- [4] J. Cížek, Adv. Chem. Phys. 14, 35 (1969).
- [5] J. Cížek and J. Paldus, Int. J. Quantum Chem. 5, 359 (1971).
- [6] J. Paldus, J. Cížek, and I. Shavitt, Phys. Rev. A 5, 50 (1972).
- [7] B. Jeziorski and H. J. Monkhorst, Phys. Rev. A 24, 1668 (1981).
- [8] B. Jeziorski and J. Paldus, J. Chem. Phys. 88, 5673 (1988).
- [9] L. Meissner, K. Jankowski, and J. Wasilewski, Int. J. Quantum Chem. 34, 535 (1988).
- [10] J. Paldus, L. Pylypow, and B. Jeziorski, in Many-Body Methods in Quantum Chemistry Lecture Notes in Chemistry, Vol. 52, edited by U. Kaldor (Springer, Berlin, 1989) p. 151.
- [11] P. Piecuch and J. Paldus, Theor. Chim. Acta 83, 69 (1992).
- [12] J. Paldus, P. Piecuch, B. Jeziorski, and L. Pylypow, in *Recent Progress in Many-Body Theories*, Vol. 3, edited by T. L. Ainsworthy, C. E. Campbell, B. E. Clements, and E. Krotschek (Plenum Press, New York, 1992) p. 287.
- [13] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A 47, 2738 (1993).
- [14] P. Piecuch, R. Toboła, and J. Paldus, Chem. Phys. Lett. **210**, 243 (1993).
- [15] P. Piecuch and J. Paldus, Phys. Rev. A 49, 3479 (1994).
- [16] P. Piecuch and J. Paldus, J. Chem. Phys. **101**, 5875 (1994).
- [17] P. Piecuch and J. Paldus, J. Phys. Chem. **99**, 15354 (1995).
- [18] K. Kowalski and P. Piecuch, Phys. Rev. A **61**, 052506 (2000).
- [19] P. Piecuch and J. I. Landman, Parallel Comp. 26, 913 (2000).

- [20] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **334**, 89 (2001).
- [21] K. Kowalski and P. Piecuch, J. Molec. Struct.: THEOCHEM 547, 191 (2001).
- [22] P. Piecuch and K. Kowalski, Int. J. Mol. Sci. 3, 676 (2002).
- [23] K. Kowalski and P. Piecuch, Mol. Phys. **102**, 2425 (2004).
- [24] J. Pittner and P. Piecuch, Mol. Phys. **131**, 1209 (2009).
- [25] L. Meissner, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **91**, 6187 (1989).
- [26] L. Meissner and R. J. Bartlett, J. Chem. Phys. **92**, 561 (1990).
- [27] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. 95, 8227 (1991).
- [28] A. Balková, S. A. Kucharski, L. Meissner, and R. J. Bartlett, Theor. Chim. Acta 80, 335 (1991).
- [29] A. Balková, S. A. Kucharski, L. Meissner, and R. J. Bartlett, J. Chem. Phys. 95, 4311 (1991).
- [30] A. Balková, S. A. Kucharski, and R. J. Bartlett, Chem. Phys. Lett. 182, 511 (1991).
- [31] A. Balková and R. J. Bartlett, J. Chem. Phys. 101, 8972 (1994).
- [32] S. Berkovic and U. Kaldor, Chem. Phys. Lett. 42, 199 (1992).
- [33] S. Berkovic and U. Kaldor, J. Chem. Phys. **98**, 3090 (1993).
- [34] J. Paldus and X. Li, J. Chem. Phys. **118**, 6769 (2003).
- [35] X. Li and J. Paldus, J. Chem. Phys. **119**, 5320 (2003).
- [36] X. Li and J. Paldus, J. Chem. Phys. **119**, 5334 (2003).
- [37] X. Li and J. Paldus, J. Chem. Phys. **119**, 5346 (2003).
- [38] X. Li and J. Paldus, Int. J. Quantum Chem. **99**, 914 (2004).
- [39] X. Li and J. Paldus, J. Chem. Phys. **120**, 5890 (2004).
- [40] X. Li and J. Paldus, J. Chem. Phys. **124**, 034112 (2006).
- [41] X. Li and J. Paldus, J. Chem. Phys. **133**, 184106 (2010).
- [42] X. Li and J. Paldus, Chem. Phys. Lett. **496**, 183 (2010).

- [43] X. Li and J. Paldus, Int. J. Quantum Chem. **110**, 2734 (2010).
- [44] X. Li and J. Paldus, J. Chem. Phys. **134**, 214118 (2011).
- [45] K. R. Shamsundar and S. Pal, J. Chem. Phys. **114**, 1981 (2001); **115**, 1979 (2001)
   [Erratum].
- [46] F. A. Evangelista, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. 125, 154113 (2006).
- [47] K. Kowalski, J. Chem. Phys. **134**, 194107 (2011).
- [48] U. S. Mahapatra, B. Datta, and D. Mukherjee, J. Chem. Phys. 110, 6171 (1999).
- [49] U. S. Mahapatra, B. Datta, and D. Mukherjee, Chem. Phys. Lett. **299**, 42 (1999).
- [50] J. Masik and I. Hubač, Adv. Quantum Chem. **31**, 75 (1998).
- [51] J. Pittner, P. Nachtigall, P. Carsky, J. Masik, and I. Hubač, J. Chem. Phys. 110, 10275 (1999).
- [52] I. Hubač, J. Pittner, and P. Carsky, J. Chem. Phys. **112**, 8779 (2000).
- [53] J. Pittner, J. Chem. Phys. **118**, 10876 (2003).
- [54] S. Chattopadhyay, D. Pahari, D. Mukherjee, and U. S. Mahapatra, J. Chem. Phys. 120, 5968 (2004).
- [55] D. Pahari, S. Chattopadhyay, A. Deb, and D. Mukherjee, Chem. Phys. Lett. 386, 307 (2004).
- [56] S. Chattopadhyay, P. Ghosh, and U. S. Mahapatra, J. Phys. B: At. Mol. Opt. Phys. 37, 495 (2004).
- [57] M. Hanrath, J. Chem. Phys. **123**, 084102 (2005).
- [58] M. Hanrath, Chem. Phys. Lett. **420**, 426 (2006).
- [59] F. A. Evangelista, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. 127, 024102 (2007).
- [60] F. A. Evangelista, A. C. Simmonett, W. D. Allen, H. F. Schaefer III, and J. Gauss, J. Chem. Phys. 128, 124104 (2008).
- [61] E. Prochnow, F. A. Evangelista, H. F. Schaefer III, W. D. Allen, and J. Gauss, J. Chem. Phys. 131, 064109 (2009).

- [62] T.-C. Jagau, E. Prochnow, F. A. Evangelista, and J. Gauss, J. Chem. Phys. 132, 144110 (2010).
- [63] F. A. Evangelista, E. Prochnow, J. Gauss, and H. F. Schaefer III, J. Chem. Phys. 132, 074107 (2010).
- [64] F. A. Evangelista and J. Gauss, J. Chem. Phys. **133**, 044101 (2010).
- [65] K. Bhaskaran-Nair, O. Demel, and J. Pittner, J. Chem. Phys. **129**, 184105 (2008).
- [66] S. Das, M. Kállay, and D. Mukherjee, J. Chem. Phys. **133**, 234110 (2010).
- [67] K. Kowalski and P. Piecuch, Int. J. Quantum Chem. 80, 757 (2000).
- [68] D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. 112 182 (2012).
- [69] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. 76, 1910 (1982).
- [70] J. M. Cullen and M. C. Zerner, J. Chem. Phys. 77, 4088 (1982).
- [71] P. Piecuch and J. Paldus, Int. J. Quantum Chem. **36**, 429 (1989).
- [72] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer III, J. Chem. Phys. 86, 2881 (1987).
- [73] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
- [74] W. Li, J. R. Gour, P. Piecuch, and S. Li, J. Chem. Phys. **131**, 114109 (2009).
- [75] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 6721 (2010).
- [76] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).
- [77] M. Schütz, and H.-J. Werner, Chem. Phys. Lett. **318** 370 (2000).
- [78] M. Schütz, and H.-J. Werner, J. Chem. Phys. **114** 661 (2001).
- [79] F. Neese, A. Hansen, and D. G. Liakos, J. Chem. Phys. **131** 064103 (2009).
- [80] D. G. Liakos and F. Neese, J. Chem. Theory Comput. **11** 4054 (2015).
- [81] J. Geertsen, M. Rittby, and R. J. Bartlett, Chem. Phys. Lett. 164, 57 (1989).
- [82] D. C. Comeau and R. J. Bartlett, Chem. Phys. Lett. 207, 414 (1993).
- [83] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. 98, 7029 (1993).

- [84] H. Nakatsuji and K. Hirao, Chem. Phys. Lett. 47, 569 (1977).
- [85] H. Nakatsuji, Chem. Phys. Lett. **59**, 362 (1978).
- [86] H. Nakatsuji, in Computational Chemistry: Reviews of Current Trends, Vol. 2, edited by J. Leszczyński, (World Scientific, Singapore, 1997) p. 62.
- [87] H. Nakatsuji, Bull. Chem. Soc. Jpn. **78**, 1705 (2005).
- [88] H. Monkhorst, Int. J. Quantum Chem. Symp. 11, 421 (1977).
- [89] E. Dalgaard and H. Monkhorst, Phys. Rev. A 28, 1217 (1983).
- [90] D. Mukherjee and P. K. Mukherjee, Chem. Phys. **39**, 325 (1979).
- [91] M. Takahashi and J. Paldus, J. Chem. Phys. 85, 1486 (1986).
- [92] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).
- [93] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. 93, 3345 (1990).
- [94] 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).
- [95] K. Kowalski, S. Krishnamoorthy, O. Villa, J. R. Hammond, and N. Govind, J. Chem. Phys. 132, 154103 (2010).
- [96] G. Fradelos, J. J. Lutz, T. A. Wesołowski, P. Piecuch, and M. Włoch, J. Chem. Theory Comput. 7, 1647 (2011).
- [97] K. Kornobis, N. Kumar, P. Lodowski, M. Jaworska, P. Piecuch, J. J. Lutz, B. M. Wong, and P. M. Kozlowski, J. Comp. Chem. 34, 987 (2013).
- [98] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. 128, 134110 (2008).
- [99] S. P. A. Sauer, M. Schreiber, M. R. Silva-Junior, and W. Thiel, J. Chem. Theory Comput. 5, 555 (2009).
- [100] M. R. Silva-Junior, S. P. Sauer, M. Schreiber, and W. Thiel, Mol. Phys. 108, 453 (2010).
- [101] T. J. Watson, Jr., V. F. Lotrich, P. G. Szalay, A. Perera, and R. J. Bartlett, J. Phys. Chem. A 117, 2569 (2013).

- [102] J. Sous, P. Goel, and M. Nooijen, Mol. Phys. **112**, 616 (2014).
- [103] D. Kánnár and P. G. Szalay, J. Chem. Theory Comput. 10, 3757 (2014).
- [104] P. Piecuch, J. A. Hansen, and A. O. Ajala, Mol. Phys. **113**, 3805 (2015).
- [105] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, Int. Rev. Phys. Chem. 21, 527 (2002).
- [106] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P.-D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kuś, and M. Musiał, Theor. Chem. Acc. **112**, 349 (2004).
- [107] K. Kowalski and P. Piecuch, J. Chem. Phys. **115**, 643 (2001).
- [108] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **347**, 237 (2001).
- [109] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 8490 (2000).
- [110] T. Shiozaki, K. Hirao, and S. Hirata, J. Chem. Phys. **126**, 224106 (2007).
- [111] K. Kowalski and P. Piecuch, J. Chem. Phys. **115**, 2966 (2001).
- [112] K. Kowalski and P. Piecuch, J. Chem. Phys. **116**, 7411 (2002).
- [113] K. Kowalski and P. Piecuch, J. Chem. Phys. **120**, 1715 (2004).
- [114] M. Włoch, J. R. Gour, K. Kowalski, and P. Piecuch, J. Chem. Phys. 122, 214107 (2005).
- [115] M. D. Lodriguito, K. Kowalski, M. Włoch, and P. Piecuch, J. Mol. Struct.: THEOCHEM 771, 89 (2006).
- [116] M. Włoch, M. D. Lodriguito, P. Piecuch, and J. R. Gour, Mol. Phys. 104, 2149 (2006);
   104, 2991 (2006) [Erratum].
- [117] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. 109, 3268 (2009).
- [118] K. Kowalski, S. Hirata, M. Włoch, P. Piecuch, and T. L. Windus, J. Chem. Phys. 123, 074319 (2005).
- [119] P. Piecuch, Mol. Phys. **108**, 2987 (2010).
- [120] J. Noga and R. J. Bartlett, J. Chem. Phys. 86, 7041 (1987); 89, 3401 (1988) [Erratum].
- [121] G. E. Scuseria and H. F. Schaefer III, Chem. Phys. Lett. 152, 382 (1988).

- [122] N. Oliphant and L. Adamowicz, J. Chem. Phys. **95**, 6645 (1991).
- [123] Stanislaw A. Kucharski and Rodney J. Bartlett, Theor. Chim. Acta 80, 387 (1991)
- [124] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. 97, 4282 (1992).
- [125] P. Piecuch and L. Adamowicz, J. Chem. Phys. **100**, 5792 (1994).
- [126] M. Kállay and P. Surjan, J. Chem. Phys. **113**, 1359 (2000).
- [127] S. Hirata, M. Nooijen, and R. J. Bartlett, Chem. Phys. Lett. **326**, 255 (2000).
- [128] S. A. Kucharski, M. Włoch, M. Musiał, and R. J. Bartlett, J. Chem. Phys. 115, 8263 (2001).
- [129] S. Hirata, J. Chem. Phys. **121**, 51 (2004).
- [130] M. Kállay and J. Gauss, J. Chem. Phys. **121**, 9257 (2004).
- [131] Y. S. Lee and R. J. Bartlett, J. Chem. Phys. 80, 4371 (1984).
- [132] Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. 81, 5906 (1984); 82, 5761 (1985) [Erratum].
- [133] J. Noga, R. J. Bartlett, and M. Urban, Chem. Phys. Lett. **134**, 126 (1987).
- [134] G. W. Trucks, J. Noga, and R. J. Bartlett, Chem. Phys. Lett. 145, 548 (1988).
- [135] P. Piecuch and J. Paldus, Theor. Chim. Acta 78, 65 (1990).
- [136] P. Piecuch, S. Zarrabian, J. Paldus, and J. Cížek, Phys. Rev. B 42, 3351 (1990).
- [137] S. A. Kucharski and R. J. Bartlett, Chem. Phys. Lett. **158**, 550 (1989).
- [138] M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, J. Chem. Phys. 83, 4041 (1985).
- [139] S. J. Cole and R. J. Bartlett, J. Chem. Phys. 86, 873 (1987).
- [140] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. 102, 479 (1989).
- [141] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. 108, 9221 (1998).
- [142] P. Piecuch, R. Toboła, and J. Paldus, Int. J. Quantum Chem. 55, 133 (1995).
- [143] P. Piecuch, R. Toboła, and J. Paldus, Phys. Rev. A 54, 1210 (1996).

- [144] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 5243 (1998).
- [145] K. Raghavachari, J. Chem. Phys. 82, 4607 (1985).
- [146] R. J. Bartlett, J. D. Watts, and S. A. Kucharski, Chem. Phys. Lett. 165, 513 (1990).
- [147] J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. 233, 81 (1995).
- [148] J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. **258**, 581 (1996).
- [149] J. D. Watts and R. J. Bartlett, J. Chem. Phys. 101, 3073 (1994).
- [150] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. 244, 75 (1995).
- [151] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. 103, 7429 (1995).
- [152] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **105**, 1451 (1996).
- [153] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. 256, 185 (1996).
- [154] P. Piecuch and K. Kowalski, in Computational Chemistry: Reviews of Current Trends, Vol. 5, edited by J. Leszczyński, (World Scientific, Singapore, 2000) p. 1.
- [155] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).
- [156] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 5644 (2000).
- [157] P. Piecuch, S. A. Kucharski, and K. Kowalski, Chem. Phys. Lett. **344**, 176 (2001).
- [158] K. Kowalski and P. Piecuch, J. Chem. Phys. **122**, 074107 (2005).
- [159] P.-D. Fan, K. Kowalski, and P. Piecuch, Mol. Phys. **103**, 2191 (2005).
- [160] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).
- [161] P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, Chem. Phys. Lett. 418, 467 (2006).
- [162] M. Włoch, J. R. Gour, and P. Piecuch, J. Phys. Chem. A **111**, 11359 (2007).
- [163] J. Shen and P. Piecuch, Chem. Phys. **401**, 180 (2012).
- [164] J. Shen and P. Piecuch, J. Chem. Phys. **136**, 144104 (2012).
- [165] J. Shen and P. Piecuch, J. Chem. Theory Comput. 8, 4968 (2012).

- [166] P. Piecuch, S. A. Kucharski, V. Spirko, and K. Kowalski, J. Chem. Phys. 115, 5796 (2001).
- [167] M. J. McGuire, K. Kowalski, and P. Piecuch, J. Chem. Phys. **117**, 3617 (2002).
- [168] C. J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, and L. Gagliardi, J. Phys. Chem. A 110, 1991 (2006); 111, 4871 (2007) (Addition/Correction).
- [169] Y. Z. Song, A. Kinal, P. J. S. B. Caridade, A. J. C. Varandas, and P. Piecuch, J. Mol. Struct.: THEOCHEM 859, 22 (2008).
- [170] A. Kinal and P. Piecuch, J. Phys. Chem. A **110**, 367 (2006).
- [171] A. Kinal and P. Piecuch, J. Phys. Chem. A **111**, 734 (2007).
- [172] X. Li and J. Paldus, J. Chem. Phys. **124**, 174101 (2006).
- [173] 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).
- [174] A. G. Taube, Mol. Phys. **108**, 2951 (2010).
- [175] P. M. Kozlowski, M. Kumar, P. Piecuch, W. Li, N. P. Bauman, J. A. Hansen, P. Lodowski, and M. Jaworska, J. Chem. Theory Comput. 8, 1870 (2012).
- [176] C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).
- [177] M. Z. Zgierski, S. Patchkovskii, and E. C. Lim, J. Chem. Phys. **123**, 081101 (2005).
- [178] M. Z. Zgierski, S. Patchkovskii, T. Fujiwara, and E. C. Lim, J. Phys. Chem. A 109, 9384 (2005).
- [179] S. Nangia, D. G. Truhlar, M. J. McGuire, and P. Piecuch, J. Phys. Chem. A 109, 11643 (2005).
- [180] M. Valiev and K. Kowalski, J. Phys. Chem. A **110**, 13106 (2006).
- [181] M. Valiev and K. Kowalski, J. Chem. Phys. **125**, 211101 (2006).
- [182] D. M. Chipman, J. Chem. Phys. **124**, 044305 (2006).
- [183] M. Z. Zgierski, S. Patchkovskii, and E. C. Lim, Can. J. Chem. 85, 124 (2007).
- [184] S. Tokura, K. Yagi, T. Tsuneda, and K. Hirao, Chem. Phys. Lett. 436, 30 (2007).
- [185] B. Ginovska, D. M. Camaioni, and M. Dupuis, J. Chem. Phys. **127**, 084309 (2007).

- [186] P.-D. Fan, M. Valiev, and K. Kowalski, Chem. Phys. Lett. 458, 205 (2008).
- [187] D. M. Chipman, J. Phys. Chem. A **112**, 13372 (2008).
- [188] A. Kvaran, H. Wang, K. Matthiasson, A. Bodi, and E. Jónsson, J. Chem. Phys. 129, 164313 (2008).
- [189] D. I. Lyakh, V. V. Ivanov, and L. Adamowicz, J. Chem. Phys. **128**, 074101 (2008).
- [190] J. R. Hammond, W. A. de Jong, and K. Kowalski, J. Chem. Phys. **128**, 224102 (2008).
- [191] N. Govind, P. V. Sushko, W. P. Hess, M. Valiev, and K. Kowalski, Chem. Phys. Lett. 470, 353 (2009).
- [192] L. A. Burns, D. Murdock, and P. H. Vaccaro, J. Chem. Phys. **130**, 144304 (2009).
- [193] S. I. Bokarev, E. K. Dolgov, V. A. Bataev, and I. A. Godunov, Int. J. Quantum Chem. 109, 569 (2009).
- [194] K. R. Glaesemann, N. Govind, S. Krishnamoorthy, and K. Kowalski, J. Phys. Chem. A 114, 8764 (2010).
- [195] J. R. Gour, P. Piecuch, and M. Włoch, Mol. Phys. **108**, 2633 (2010).
- [196] P. N. Day, K. A. Nguyen, and R. Pachter, J. Chem. Theory Comput. 6, 2809 (2010).
- [197] K. Lopata, R. Reslan, M. Kowalska, D. Neuhauser, N. Govind, and K. Kowalski, J. Chem. Theor. Comput. 7, 3686 (2011).
- [198] K. Kowalski, R. M. Olson, S. Krishnamoorthy, V. Tipparaju, and E. Aprá, J. Chem. Theor. Comput. 7, 2200 (2011).
- [199] G. Fradelos, J. J. Lutz, T. A. Wesołowski, P. Piecuch, and M. Włoch, in Advances in the Theory of Quantum Systems in Chemistry and Physics, Vol 22, Progress in Theoretical Chemistry and Physics, edited by P. E. Hoggan, E. J. Brändas, J. Maruani, P. Piecuch, and G. Delgado-Barrio, (Springer, Dordrecht, 2012) p. 219.
- [200] P. D. Dau, J. Su, H.-T. Liu, J.-B. Liu, D.-L. Huang, J. Li, and L.-S. Wang, Chem. Sci. 3, 1137 (2012).
- [201] G. Murdachaew, M. Valiev, S. M. Kathmann, and X.-B. Wang, J. Phys. Chem. A 116, 2055 (2012).
- [202] P. D. Dau, J. Su, H.-T. Liu, D.-L. Huang, J. Li, and L.-S. Wang, J. Chem. Phys. 137, 064315 (2012).

- [203] N. De Silva, S. Y. Willow, and M. S. Gordon, J. Phys. Chem. A **117**, 11847 (2013).
- [204] J. Su, P. D. Dau, C.-F. Xu, D.-L. Huang, H.-T. Liu, F. Wei, L.-S. Wang, and J. Li, Chem. Asian J. 8, 2489 (2013).
- [205] J. Su, W.-H. Xu, C.-F. Xu, W. H. E. Schwarz, and J. Li, Inorg. Chem. 52, 9867 (2013).
- [206] E. Berardo, H.-S. Hu, K. Kowalski, and M. A. Zwijnenburg, J. Chem. Phys. 139, 064313 (2013).
- [207] P. Tecmer, N. Govind, K. Kowalski, W. A. de Jong, and L. Visscher, J. Chem. Phys. 139, 034301 (2013).
- [208] P. Piecuch, J. A. Hansen, D. Staedter, S. Faure, and V. Blanchet, J. Chem. Phys. 138, 201102 (2013).
- [209] J. Su, P. D. Dau, Y.-H. Qiu, H.-T. Liu, C.-F. Xu, D.-L. Huang, L.-S. Wang, and J. Li, Inorg. Chem. 52, 6617 (2013).
- [210] F. Trinter, J. B. Williams, M. Weller, M. Waitz, M. Pitzer, J. Voigtsberger, C. Schober, G. Kastirke, C. Müller, C. Goihl, P. Burzynski, F. Wiegandt, R. Wallauer, A. Kalinin, L. P. H. Schmidt, M. S. Schöffler, Y.-C. Chiang, K. Gokhberg, T. Jahnke, and R. Dörner, Phys. Rev. Lett. **111**, 233004 (2013).
- [211] J. J. Lutz and P. Piecuch, Comput. Theor. Chem. **1040-1041**, 20 (2014).
- [212] P. Zhou, J. Liu, K. Han, and G. He, J. Comput. Chem. **35**, 109 (2014).
- [213] G. Schoendorff and A. K. Wilson, J. Chem. Phys. 140, 224314 (2014).
- [214] H.-S. Hu, K. Bhaskaran-Nair, E. Aprá, N. Govind, and K. Kowalski, J. Phys. Chem. A 118, 9087 (2014).
- [215] J. Su, P. D. Dau, H.-T. Liu, D.-L. Huang, F. Wei, W. H. E. Schwarz, J. Li, and L.-S. Wang, J. Chem. Phys. 142, 134308 (2015).
- [216] P. Arora, W. Li, P. Piecuch, J. W. Evans, M. Albao, and M. S. Gordon, J. Phys. Chem. C 114, 12649 (2010).
- [217] X. Li and J. Paldus, J. Chem. Phys. **131**, 114103 (2009).
- [218] J. Shen, T. Fang, S. Li, and Y. Jiang, J. Phys. Chem. A **112**, 12518 (2008).
- [219] O. Demel, K. R. Shamasundar, L. Kong, and M. Nooijen, J. Phys. Chem. A 112, 11895 (2008).

- [220] D. I. Lyakh, V. F. Lotrich, and R. J. Bartlett, Chem. Phys. Lett. 501, 166 (2011).
- [221] D. W. Whitman and B. K. Carpenter, J. Am. Chem. Soc. **104**, 6473 (1982).
- [222] B. K. Carpenter, J. Am. Chem. Soc. **105**, 1700 (1983).
- [223] B. A. Hess, P. Carsky, and L. J. Schaad, J. Am. Chem. Soc. 105, 695 (1983).
- [224] A. F. Voter and W. A. Goddard III, J. Am. Chem. Soc. **108**, 2830 (1986).
- [225] P. Čarsky, R. J. Bartlett, G. Fitzgerald, J. Nova, and V. Špirko, J. Chem. Phys. 89, 3008 (1988).
- [226] B. R. Arnold and J. Michl, J. Phys. Chem. 97, 13348 (1993).
- [227] W. Wu, Y. Mo, Z. Cao, and Q. Zhang, Theor. Comput. Chem. 10, 143 (2002).
- [228] S. V. Levchenko and A. I. Krylov, J. Chem. Phys. **120**, 175 (2004).
- [229] O. Demel and J. Pittner, J. Chem. Phys. **124**, 144112 (2006).
- [230] S. Saddique and G. A. Worth, Chem. Phys. **329**, 99 (2006).
- [231] M. Eckert-Maksić, M. Vazdar, M. Barbatti, H. Lischka, and Z. B. Maksić, J. Chem. Phys. 125, 064310 (2006).
- [232] P. B. Karadakov, J. Phys. Chem. **112**, 7303 (2008).
- [233] K. A. Nguyen and M. S. Gordon, J. Am. Chem. Soc. **117**, 3835 (1995)
- [234] J. Lutz and P. Piecuch, J. Chem. Phys. **128**, 154116 (2008).
- [235] D. A. Mazziotti, J. Phys. Chem. A **112**, 13684 (2008).
- [236] R. Berner and A. Lüchow, J. Phys. Chem. A **114**, 13222 (2010).
- [237] T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).
- [238] N. Oliphant and L. Adamowicz, J. Chem. Phys. 94, 1229 (1991).
- [239] N. Oliphant and L. Adamowicz, J. Chem. Phys. 96, 3739 (1992).
- [240] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. 99, 1875 (1993).
- [241] P. Piecuch and L. Adamowicz, J. Chem. Phys. **102**, 898 (1995).
- [242] K. B. Ghose, P. Piecuch, and L. Adamowicz, J. Chem. Phys. **103**, 9331 (1995).

- [243] P. Piecuch, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. 110, 6103 (1999).
- [244] P. Piecuch, S. A. Kucharski, and V. Spirko, J. Chem. Phys. **111**, 6679 (1999).
- [245] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **344**, 165 (2001).
- [246] J. R. Gour, P. Piecuch, and M. Włoch, J. Chem. Phys. **123**, 134113 (2005).
- [247] J. R. Gour, P. Piecuch, and M. Włoch, Int. J. Quantum Chem. 106, 2854 (2006).
- [248] J. R. Gour and P. Piecuch, J. Chem. Phys. **125**, 234107 (2006).
- [249] J. Shen and P. Piecuch, J. Chem. Phys. **138**, 194102 (2013).
- [250] J. Shen and P. Piecuch, Mol. Phys. **112**, 868 (2014).
- [251] P. Piecuch, J. Shen, N. P. Bauman, and M. Ehara, in preparation.
- [252] N. P. Bauman, J. Shen, and P. Piecuch, in preparation.
- [253] I. Magoulas, N. P. Bauman, J. Shen, and P. Piecuch, in preparation.
- [254] M. Eckert-Maksić, M. Vazdar, M. Barbatti, H. Lischka, and Z. B. Maksić, J. Chem. Phys. 125, 064310 (2006).
- [255] X. Li and J. Paldus, J. Chem. Phys. **124** 174101 (2006).
- [256] X. Li and J. Paldus, J. Chem. Phys. **129** 054104 (2008).
- [257] X. Li and J. Paldus, J. Chem. Phys. **107** 6257 (1997).
- [258] J. Shen, E. Xu, Z. Kou, and S. Li, J. Chem. Phys. **132**, 114115 (2010).
- [259] J. Shen, Z. Kou, E. Xu, and S. Li, J. Chem. Phys. **133**, 234106 (2010).
- [260] J. Shen, Z. Kou, E. Xu, and S. Li, J. Chem. Phys. **134**, 044134 (2011).
- [261] J. Shen, E. Xu, Z. Kou, and S. Li, Phys. Chem. Chem. Phys. 13, 8795 (2011).
- [262] Z. Kou, J. Shen, E. Xu, and S. Li, J. Chem. Phys. **136**, 194105 (2012).
- [263] J. F. Stanton, Chem. Phys. Lett. **281** 130 (1997).
- [264] A. G. Taube and R. J. Bartlett, J. Chem. Phys. **128** 044110 (2008).
- [265] A. G. Taube and R. J. Bartlett, J. Chem. Phys. **128** 044111 (2008).

- [266] S. Hirata, P.-D. Fan, A. A. Auer, M. Nooijen, and P. Piecuch, J. Chem. Phys. 121 12197 (2004).
- [267] A. Melnichuk and R. J. Bartlett, J. Chem. Phys. **137** 214103 (2012).
- [268] A. Melnichuk and R. J. Bartlett, J. Chem. Phys. **140** 064113 (2012).
- [269] T. Kinoshita, O. Hino, and R. J. Bartlett, J. Chem. Phys. **123** 074106 (2005).
- [270] "O. Hino, T. Kinoshita, G. K.-L. Chan, and R. J. Bartlett, J. Chem. Phys. 124 114311 (2006).
- [271] J. J. Eriksen, K. Kristensen, T. Kjærgaard, P. Jørgensen, and J. Gauss, J. Chem. Phys. 140 064108 (2014).
- [272] J. J. Eriksen, P. Jørgensen, J. Olsen, and J. Gauss, J. Chem. Phys. **140** 174114 (2014).
- [273] J. J. Eriksen, P. Jørgensen, and J. Gauss, J. Chem. Phys. **142** 014102 (2015).
- [274] Y. Ohtsuka, P. Piecuch, J. R. Gour, M. Ehara, and H. Nakatsuji, J. Chem. Phys. 126, 164111 (2007).
- [275] P. Piecuch, S. Hirata, K. Kowalski, P.-D. Fan, and T. L. Windus, Int. J. Quantum Chem. 106, 79 (2006).
- [276] P.-D. Fan and S. Hirata, J. Chem. Phys. **124**, 104108 (2006).
- [277] 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).
- [278] 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, (Elsevier, Amsterdam, 2005) p. 1167.
- [279] T. H. Dunning Jr., J. Chem. Phys. 53, 2823 (1970).
- [280] P. Saxe, H. F. Schaefer III, and N. C. Handy, Chem. Phys. Lett. **79**, 202 (1981).
- [281] R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **95**, 386 (1983).
- [282] J. Olsen, F. Jorgensen, H. Koch, A. Balkova, and R. J. Bartlett, J. Chem. Phys. 104, 8007 (1996).
- [283] G. D. Purvis III, R. Shepard, F. B. Brown, and R. J. Bartlett, Int. J. Quantum Chem. 23, 835 (1983).

- [284] J. R. Hart, A. K. Rappe, S. M. Gorun, and T. H. Upton, J. Phys. Chem. 96, 6264 (1992).
- [285] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. 108, 2128 (2008).
- [286] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. 56, 2257 (1972).
- [287] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta 28, 213 (1973).
- [288] A. E. Douglas and G. Herzberg, Can. J. Res. 18, 179 (1940).
- [289] B. A. Thrush, Nature **186**, 1044 (1044).
- [290] O. A. Mosher and R. P. Frosch, J. Chem. Phys. 52, 5781 (1970).
- [291] H. Bredohl, I. Dubois, Y. Houbrechts, and P. Nzohabonayo, Phys. B: At. Mol. Phys. 17, 95 (1984).
- [292] H. Bredohl, I. Dubois, Y. Houbrechts, and P. Nzohabonayo, J. Mol. Spectrosc. 112, 430 (1985).
- [293] J. L. Masse and M. Bärlocher, Helvetica Chimica Acta 47, 314 (1964).
- [294] G. Verhaegen, W. G. Richards, and C. M. Moser, J. Chem. Phys. 46, 160 (1967).
- [295] M. P. Melrose and D. Russell, J. Chem. Phys. 55, 470 (1971).
- [296] Y. U. Khait and V. I. Baranovskii, J. Struct. Chem. 21, 266 (1980).
- [297] P. K. Shashi and G. Friedrich, Chem. Phys. **98**, 207 (1985).
- [298] R. C. Mawhinney, P. J. Bruna, and F. Grein, Can. J. Chem. **71**, 1581 (1993).
- [299] M. Lorenz, J. Agreiter, A. M. Smith, and V. E. Bondybey, J. Chem. Phys. 104, 3143 (1996).
- [300] K. R. Asmis, T. R. Taylor, and D. M. Neumark, Chem. Phys. Lett. 295, 75 (1998).
- [301] J. M. L. Martin, J. P. François, and R. Gijbels, Z. Phys. D 21, 47 (1991).
- [302] J. M. L. Martin, T. J. Lee, G. E. Scuseria, and P. R. Taylor, J. Chem. Phys. 97, 6549 (1992).
- [303] K. A. Peterson, J. Chem. Phys. **102**, 262 (1995).
- [304] P. J. Bruna, R. C. Mawhinney, and F. Grein, Int. J. Quant. Chem. 56, 455 (1995).

- [305] C. W. Bauschlicher and H. Partridge, Chem. Phys. Lett. 257, 601 (1996).
- [306] R. R. Reddy, Y. Nazeer Ahammed, K. Rama Gopal, P. Abdul Azeem, and S. Anjaneyulu, Astrophys. Space Sci. 262, 223 (1998).
- [307] Z.-L. Cai and J. P. François, Chem. Phys. Lett. **300**, 69 (1999).
- [308] G. Chambaud, M. Guitou, and S. Hayashi, Chem. Phys. **352**, 147 (2008).
- [309] M. Ončák and M. Srnec, J. Comp. Chem. 29, 233 (2008).
- [310] D. Shi, W. Xing, H. Liu, J. Sun, Z. Zhu, and Y. Liu, Spectrochim. Acta Part A 93, 367 (2012).
- [311] S. Mahmoud, M. Bechelany, P. Miele, and M. Korek, J. Quant. Spectrosc. Radiat. Transfer 151, 58 (2015).
- [312] J. D. Watts, in Computational Chemistry : Reviews of Current Trends, Vol 7, edited by J. Leszczynski (World Scientific, Singapore, 2002), p. 89.
- [313] P. A. Denis, Chem. Phys. Lett. **395**, 12 (2004).
- [314] Z. Gan, D. J. Grant, R. J. Harrison, and D. A. Dixon, J. Chem. Phys. 125, 124311 (2006).
- [315] X. Li and J. Paldus, Chem. Phys. Lett. **431**, 179 (2006).
- [316] X. Li, J. R. Gour, J. Paldus, and P. Piecuch, Chem. Phys. Lett. 461, 321 (2008).
- [317] A. Karton and J. M. L Martin, J. Chem. Phys. **125**, 144313 (2006).
- [318] J. M. L. Martin, Mol. Phys. **112**, 785 (2014).
- [319] D. Kats, D. Kreplin, H-J Werner, and F. R. Manby, J. Chem. Phys. 142, 064111 (2015).
- [320] S-I Lu, J. Chem. Phys. **123**, 174313 (2005).
- [321] J. Paldus and J. Cížek, Adv. Quantum Chem. 9, 105 (1975).
- [322] J. Paldus in Methods in Computational Molecular Physics, NATO Advanced Study Institute, Series B: Physics, Vol. 293, edited by S. Wilson and G. H. F. Diercksen, (Plenum, New York, 1992) p. 99.
- [323] J. Paldus in *Encyclopedia of Computational Chemistry*, Vol. 1, edited by S. Wilson (Wiley, Chichester, 2003) p. 615.

- [324] J. Paldus, in Diagrammatic Methods for Many-Fermion system (University of Nijmegen, 1981)
- [325] J. Paldus, J. Chem. Phys. 67, 303 (1976).
- [326] S. A. Kucharski and R. J. Bartlett, Adv. Quantum Chem. 18, 281 (1986).
- [327] M. D. Lodriguito and P. Piecuch in Progress in Theoretical Chemistry and Physics, Vol 18, Fontiers in Quantum Systems in Chemistry and Physics, edited by S. Wilson, P. Grout, J. Maruani, G. Delgado-Barrio, and P. Piecuch (Springer, Dordrecht, 2008) p. 67.