



LIBRARIES MICHIGAN STATE UNIVERSITY EAST LANSING, MICH 48824-1048

This is to certify that the dissertation entitled

NEW ALTERNATIVES FOR ELECTRONIC STRUCTURE THEORY: THE APPLICATION OF TWO-BODY CLUSTER EXPANSIONS IN HIGH ACCURACY AB INITIO CALCULATIONS

presented by

Peng-Dong Fan

has been accepted towards fulfillment of the requirements for the

Ph.D.	degree in	chemistry		
	Child	Speul		
	Major Pro	fessor's Signature		
May 2, 2005				
		Date		

MSU is an Affirmative Action/Equal Opportunity Institution

.

PLACE IN RETURN BOX to remove this checkout from your record. TO AVOID FINES return on or before date due. MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
WG 2 2 2007		
AUG 2 9 2007		
		2/05 c:/CIRC/DateDue.indd-

1 ...

NEW ALTERNATIVES FOR ELECTRONIC STRUCTURE THEORY: THE APPLICATION OF TWO-BODY CLUSTER EXPANSIONS IN HIGH ACCURACY AB INITIO CALCULATIONS

By

Peng-Dong Fan

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Chemistry

2005

ABSTRACT

NEW ALTERNATIVES FOR ELECTRONIC STRUCTURE THEORY: THE APPLICATION OF TWO-BODY CLUSTER EXPANSIONS IN HIGH ACCURACY AB INITIO CALCULATIONS

By

Peng-Dong Fan

In this thesis, the applicability of exponential cluster expansions involving oneand two-electron operators in high accuracy ab initio calculations is discussed. First, the extended coupled-cluster method with singles and doubles (ECCSD) is tested in the most demanding studies of potential energy surfaces involving multiple bond breaking. The numerical results for a few cases of multiple bond breaking show that the single-reference ECCSD method is capable of providing a qualitatively correct description of entire potential energy surfaces, eliminating, in particular, the failures and the unphysical behavior of all standard coupled-cluster methods in similar cases. It is also demonstrated that one can obtain entire potential energy surfaces with millihartree accuracies by combining the ECCSD theory with the noniterative a posteriori corrections obtained by using the generalized variant of the method of moments of coupled-cluster equations. This is the first time when the relatively simple single-reference formalism, employing one- and two-body clusters only, provides a highly accurate description of the dynamic and significant nondynamic correlation effects characterizing multiply bonded systems. Second, an evidence is presented that one may be able to represent the exact or virtually exact ground- and excited-state wave functions of many-electron systems by exponential cluster expansions employing general two-body operators. Calculations for small many-electron systems indicate the existence of finite two-body parameters that produce the numerically exact wave functions. This finding may have a significant impact on future quantum calculations for many-electron systems, since normally one needs triply excited, quadruply excited, and other higher-than-doubly excited Slater determinants, in addition to all singly and doubly excited determinants, to obtain the exact or virtually exact wave functions.

ACKNOWLEDGMENT

I feel very privileged to have worked with my supervisor, Dr. Piotr Piecuch. To him I owe a great debt of gratitude for his patience, inspiration, and the warmest support. Piotr has taught me a great deal about electronic structure theory and guided me to explore new alternatives in this field.

I also owe a great debt of gratitude to my Guidance Committee, including Dr. Katharine C. Hunt, Dr. Paul F. Mantica, and Dr. James K. McCusker, for advice, support, and setting up and overseeing my graduate study program.

I would also like to thank Dr. Karol Kowalski, a former postdoctoral research associate in our group, who shared the results of his work when I started working on this project. I would like to thank Ian, Mike, Ruth, DJ, and Maricris for their friendship and help in practicing my seminar presentations.

Finally, I would like to acknowledge the Department of Chemistry for providing the teaching assistantship and the Dissertation Completion Fellowship, and the US Department of Energy for the research assistantship.

Table of Contents

Li	st of	Figures	vi			
Li	st of	Tables	vii			
1	Int	roduction	1			
2	Practical Ways of Improving Standard Coupled-Cluster Methods					
	Employing Singly and Doubly Excited Clusters via Extended Coupled-					
	Clu	ster Theory	5			
	2.1	Extended Coupled-Cluster Theory: General Formalism	5			
	2.2	Extended Coupled-Cluster Methods with Singles and Doubles	9			
		2.2.1 The Piecuch-Bartlett ECCSD Approach	9			
		2.2.2 The Arponen-Bishop ECCSD Approach	12			
	2.3	Numerical Results for Multiple Bond Breaking	17			
		2.3.1 The Piecuch-Bartlett ECCSD Approach	17			
		2.3.2 The Arponen-Bishop ECCSD Approach	34			
	2.4	Conclusion	41			
3	No	niterative Corrections to Extended Coupled-Cluster Energies for				
U	Hio	h Accuracy Electronic Structure Calculations: Generalized Meth	od			
	of	Annound of Coupled-Cluster Equations	43			
	3.1	The Standard Method of Moments of Coupled-Cluster Equations	44			
	3.2	The Generalized MMCC Formalism	50			
	3.3	The ECCSD(T), ECCSD(TQ), QECCSD(T), and QECCSD(TQ) Meth-				
	••••	ods and their Performance in Calculations for Triple Bond Breaking in				
		N_2	56			
	3.4	Summary	64			
4	Exa	ctness of Two-Body Cluster Expansions in Many-Body Quantum				
	The	ory	70			
	4.1	Theory	71			
		4.1.1 The $\exp(X)$ Conjecture	72			
		4.1.2 Formal Arguments in Favor of the $exp(X)$ Conjecture (Ground				
		States)	75			
		4.1.3 Extension of the $exp(X)$ Conjecture to Excited States	81			
	4.2	Numerical Results	82			
	4.3	Summary	89			
B	iblio	graphy	90			

•

List of Figures

2.1	Potential energy curves for the MBS S4 system. ECCD represents the Piecuch-Bartlett variant of the ECCD approach (in this case, the ECCD and ECCSD approaches are equivalent).	19
2.2	Potential energy curves for the N_2 molecule, as described by the STO-3G basis set. The LECCSD, BECCSD, QECCSD, and ECCSD methods are based on the Piecuch-Bartlett variant of the ECC theory. \hdots	26
2.3	(a) The overlaps of the normalized CCSD, BECCSD, QECCSD, and ECCSD wave functions, Eq. (2.62), with the normalized full CI wave function for the STO-3G N ₂ molecule, as functions of the N-N separa- tion R (in bohr). (b) The difference between the CCSD and ECCSD cluster operators T and the difference between the ECCSD cluster op- erators T and Σ , as defined by the quantity $d(Y, Z)$, Eq. (2.63), for the STO-3G N ₂ molecule, as functions of the N-N separation R (in bohr).	32
2.4	Ground-state potential energy curves of the N_2 molecule, as described by the STO-3G basis sets. The QECCSD and ECCSD methods are based on the Arponen-Bishop ECC theory.	38
2.5	Ground-state potential energy curves of the N_2 molecule, as described by the DZ basis sets. QECCSD represents the quadratic version of the ECCSD theory of Arponen and Bishop	40
3.1	Ground-state potential energy curves of the N_2 molecule as described by the STO-3G basis set. The QECCSD and ECCSD calculations were performed using the ECC theory of Arponen and Bishop	68
3.2	Ground-state potential energy curves of the N_2 molecule as described by the DZ basis set. The QECCSD calculations were performed using the ECC theory of Arponen and Bishop.	69

List of Tables

2.1	Ground-state energies of the MBS S4 model system as functions of the parameter α .	21
2.2	Ground-state energies of the N_2 molecule, as described by the STO-3G basis set	25
2.3	Ground-state energies of the N_2 molecule, as described by the STO-3G basis set.	29
2.4 2.5	The ground-state energies of the N_2 molecule obtained for several in- ternuclear separations R with the STO-3G basis set	37 39
3.1 3.2	The ground-state energies of the N_2 molecule obtained for several in- ternuclear separations R with the STO-3G basis set. All ECC-related calculations were performed with the Arponen-Bishop ECC theory The ground-state energies of the N_2 molecule obtained for several in- ternuclear separations R with the DZ basis set	66 67
4.1	A comparison of the ground-state energies of the MBS H8 system ob- tained with the $\exp(X)$ -like wave functions, where $X = X_2$ (a purely two-body operator) or $X_1 + X_2$ (a sum of one- and two-body opera- tors), with the exact, full CI, energies and energies obtained in various CI and CC calculations. Full CI energies are total energies in hartree and all other energies are errors relative to full CI (also in hartree). We also give the overlaps of the normalized ground-state wave functions obtained with the $\exp(X)$ ansatz and the full CI approach. The over- laps of the reference states $ \Phi_0\rangle$ and $ \tilde{\Phi}_0\rangle$ with the full CI ground-state wave function $ \Psi_0\rangle$ are given for comparison	87
4.2	Same as Table 4.1 for the first excited state of the ${}^{1}A_{g}$ symmetry	88

Chapter 1 Introduction

Great advances have been made in *ab initio* quantum chemistry. Highly accurate calculations for closed-shell and simple open-shell molecular systems involving a few atoms are nowadays routine. This, in particular, applies to coupled-cluster (CC) theory¹⁻⁵, which has become the *de facto* standard for high accuracy calculations for atomic and molecular systems $^{6-13}$. The basic single-reference CC methods, such as CCSD (CC approach with singles and doubles)¹⁴, and the noniterative $CCSD + T(CCSD) = CCSD[T]^{15}$ and $CCSD(T)^{16}$ approaches that account for the effect of triexcited clusters by using arguments based on the many-body perturbation theory (MBPT), in either the spin-orbital¹⁴⁻¹⁶ and spin-free¹⁷⁻¹⁹ or spin-adapted²⁰⁻²² forms, are nowadays routinely used in accurate electronic structure calculations. The idea of adding the *a posteriori* corrections due to higher-than-doubly excited clusters to CCSD energies, on which the CCSD[T] and CCSD(T) approaches and their more recent $CCSD(TQ_f)$ extension²³ are based, is particularly attractive, since it leads to methods that offer an excellent compromise between high accuracy and relatively low computer cost, as has been demonstrated over and over in numerous atomic and molecular applications $^{7-10,12,13}$.

There are, however, open problems in CC theory. First and foremost is the pervasive failing of the standard single-reference CC methods, such as CCSD, CCSD[T], CCSD(T), and CCSD(TQ_f), at larger internuclear separations, when the spin-adapted restricted Hartree-Fock (RHF) configuration is used as a reference, which limits the applicability of the standard CC approaches to molecules near their equilibrium geometries. Second is the large computer effort associated with the need for using higher-than-doubly excited clusters in calculations involving quasi-degenerate electronic states and bond breaking when larger many-electron systems are examined. Undoubtedly, it would be very useful to extend the applicability of the standard single-reference CC methods to entire molecular potential energy surfaces (PESs) involving bond breaking, and quasi-degenerate electronic states in general, without invoking complicated and time-consuming steps associated with the more traditional multi-reference approaches, in which one has to choose active orbitals and multidimensional reference spaces on an *ad hoc* molecule-by-molecule basis. Ideally, one would like to develop a straightforward theory which could provide a virtually exact description of many-electron wave functions with the exponential, CC-like, expansions involving one- and two-body operators only, since the molecular electronic Hamiltonian does not include higher-than-two-body interactions.

There are several specific challenges in all those areas. First of all, the RHF-based CCSD method, on which the noniterative CCSD[T], CCSD(T), and CCSD(TQ_f) approaches are based, is inadequate for the description of bond breaking and quasidegenerate states, since it neglects all higher-than-doubly excited clusters, including the important triply and quadruply excited T_3 and T_4 components. Second, the noniterative triples and quadruples corrections defining the CCSD[T], CCSD(T), and CCSD(TQ_f) methods aggravate the situation even further, since the usual arguments originating from MBPT, on which these standard noniterative CC approximations are based, fail due to the divergent behavior of the MBPT series at larger internuclear separations (or when the strong configurational quasi-degeneracy and large nondynamic correlation effects set in). In consequence, the ground-state PESs ob-

tained with the CCSD[T], CCSD(T), $CCSD(TQ_f)$, and other noniterative CC approaches are completely pathological when the RHF configuration is used as a reference (cf., e.g., Refs. 9, 11-13, 24-37 and references therein). The iterative extensions of the CCSD[T], CCSD(T), and $CCSD(TQ_f)$ methods, including, among many others, the CCSDT- $n^{21,38-41}$ and CCSDTQ- 1^{42} approaches, and the noniterative $CCSDT + Q(CCSDT) = CCSDT[Q]^{42}$ and $CCSDT(Q_f)^{23}$ methods, in which the a posteriori corrections due to T_4 cluster components are added to the full CCSDT $(CC singles, doubles, and triples)^{43,44}$ energies, improve the description of PES in the bond breaking region, but ultimately all of these approaches break down at larger internuclear distances (see, e.g., Refs. 28, 29, 34). One might try to resolve the failures of the standard single-reference CC approaches in the bond breaking region and for quasi-degenerate electronic states in a brute-force manner by including the triply excited, quadruply excited, pentuply excited, etc. clusters fully and in a completely iterative fashion, but, unfortunately, the resulting CCSDTQ (CC singles, doubles, triples, and quadruples)⁴⁵⁻⁴⁸, CCSDTQP (CC singles, doubles, triples, quadruples, and pentuples)⁴⁹, etc. approaches are far too expensive for routine molecular applications. For example, the full CCSDTQ method requires iterative steps that scale as $n_o^4 n_u^6$ (n_o (n_u) is the number of occupied (unoccupied) orbitals in the molecular orbital basis). This \mathcal{N}^{10} scaling with the system size (\mathcal{N}) restricts the applicability of the CCSDTQ approach to very small systems, consisting of $\sim 2-3$ light atoms described by small basis sets. For comparison, CCSD(T) is an $n_o^2 n_u^4$ (or \mathcal{N}^6) procedure in the iterative CCSD steps and an $n_o^3 n_u^4$ (or \mathcal{N}^7) procedure in the noniterative part related to the calculation of the triples (T) correction. In consequence, it is nowadays possible to perform the CCSD(T) calculations for systems with 10-20 light atoms or a few heavier (transition metal) atoms. This indicates that in searching for new

methods that would help to overcome the failures of the standard CC approaches in the bond breaking region, one should focus on the idea of improving the results of the low-order CC calculations, such as CCSD, with the noniterative corrections of the CCSD(T) type, since only such methods have a chance to be applied to larger molecular systems in the not-too-distant future.

In view of the above discussion, the question: Can one improve the quality of standard CC wave functions in the bond breaking region at the basic CCSD level of the single-reference CC theory? seems to be particularly important. In this thesis, we show that this can be accomplished by exploring the so-called extended coupledcluster (ECC) theory. The basic ECCSD results, particularly when multiple bonds are stretched or broken, are qualitatively much better than the corresponding standard CCSD results. However, they are not yet fully quantitative. This prompts another question: Can one improve the quality of the ECCSD results by adding a simple a posteriori correction to the ECCSD energy which is obtained by using the singly and doubly excited cluster amplitudes obtained with the ECCSD approach? In this thesis, we show that the answer to this question is affirmative if we use the generalized method of moments of coupled-cluster equations (GMMCC). Eventually, of course, one would prefer to use only one- and two-body clusters to obtain an exact or virtually exact description of many-electron systems, since, as we have already mentioned, the Hamiltonians used in quantum chemistry do not contain higher-than-pairwise interactions. This prompts the third and the final question of this thesis research: Can one obtain the exact or virtually exact many-electron wave functions by using two-body exponential cluster expansions?

Chapter 2 Practical Ways of Improving Standard Coupled-Cluster Methods Employing Singly and Doubly Excited Clusters via Extended Coupled-Cluster Theory

2.1 Extended Coupled-Cluster Theory: General Formalism

The extended coupled-cluster (ECC) theory is based on the asymmetric, doubly connected energy functional⁵⁰⁻⁶⁰,

$$E_0^{(\text{ECC})} = \langle \Phi | \bar{\bar{H}} | \Phi \rangle, \qquad (2.1)$$

where $|\Phi\rangle$ is the independent particle model reference configuration (e.g., the Hartree-Fock determinant) and

$$\bar{\bar{H}} = e^{\Sigma^{\dagger}} (e^{-T} H e^{T}) e^{-\Sigma^{\dagger}} = e^{\Sigma^{\dagger}} \bar{H} e^{-\Sigma^{\dagger}} = (e^{\Sigma^{\dagger}} \bar{H})_{C} = \left[e^{\Sigma^{\dagger}} (H e^{T})_{C} \right]_{C}$$
(2.2)

is the doubly transformed Hamiltonian, obtained by transforming the similarity transformed Hamiltonian \bar{H} used in the standard CC theory,

$$\bar{H} = e^{-T} H e^{T} = (H e^{T})_{C},$$
 (2.3)

where H is the Hamiltonian and C stands for the connected part of the corresponding operator expression, with the exponential operator $e^{-\Sigma^{\dagger}}$. T is the usual cluster operator, which is a particle-hole excitation operator generating the connected components of the many-electron ground-state wave function

$$|\Psi_0\rangle = e^T |\Phi\rangle, \tag{2.4}$$

and Σ^{\dagger} is the auxiliary hole-particle deexcitation operator. In the exact theory, T is a sum of all many-body components T_n with n = 1, ..., N,

$$T = \sum_{n=1}^{N} T_n, \qquad (2.5)$$

where N is the number of electrons and T_n is defined as

$$T_n = \sum_{i_1 < \dots < i_n, \ a_1 < \dots < a_n} t_{a_1 \dots a_n}^{i_1 \dots i_n} E_{i_1 \dots i_n}^{a_1 \dots a_n}, \qquad (2.6)$$

with

$$E_{i_1\cdots i_n}^{a_1\cdots a_n} = \prod_{\kappa=1}^n c^{a_\kappa} c_{i_\kappa}$$
(2.7)

representing the excitation operators and $t_{a_1}^{i_1 \cdots i_n}$ designating the corresponding cluster amplitudes. We use a notation in which c^p (c_p) are the usual creation (annihilation) operators $(c^p = c_p^{\dagger})$ associated with a given orthonormal spin-orbital basis set $\{p\}$. Letters i_1, \ldots, i_n represent the occupied spin-orbitals in $|\Phi\rangle$ and a_1, \ldots, a_n designate the unoccupied spin-orbitals. The auxiliary operator Σ^{\dagger} is defined as

$$\Sigma^{\dagger} = \sum_{n=1}^{N} \Sigma_{n}^{\dagger}, \qquad (2.8)$$

where

$$\Sigma_n^{\dagger} = \sum_{i_1 < \cdots < i_n, \ a_1 < \cdots < a_n} \sigma_{i_1 \cdots i_n}^{a_1 \cdots a_n} E_{a_1 \cdots a_n}^{i_1 \cdots i_n}.$$
(2.9)

The operators

$$E_{a_1\cdots a_n}^{i_1\cdots i_n} = \prod_{\kappa=1}^n c^{i_\kappa} c_{a_\kappa}$$
(2.10)

and the coefficients $\sigma_{i_1\cdots i_n}^{a_1\cdots a_n}$ are the corresponding deexcitation operators and amplitudes, respectively.

The operators T and Σ^{\dagger} (or the corresponding amplitudes $t_{a_1}^{i_1 \cdots i_n}$ and $\sigma_{i_1}^{a_1 \cdots a_n}$) defining the wave function $|\Psi_0\rangle$ through Eq. (2.4) and the energy $E_0^{(\text{ECC})}$ through Eq.

(2.1) can be determined in various ways. In the ECC theory of Piecuch and Bartlett⁶⁰, which can be applied to both ground and excited states, the T and Σ^{\dagger} operators are determined by considering the doubly transformed form of the electronic Schrödinger equation, i.e.,

$$\bar{H}|\Phi\rangle = E_0|\Phi\rangle,$$
 (2.11)

and its left-hand analog,

$$\langle \tilde{\Phi} | \bar{\bar{H}} = E_0 \langle \tilde{\Phi} |, \qquad (2.12)$$

where $\langle \tilde{\Phi} |$ is the left eigenstate of \overline{H} corresponding to the right reference eigenstate $|\Phi\rangle$. We obtain Eq. (2.11) by inserting the formula for the CC wave function, Eq. (2.4), into the Schrödinger equation,

$$H|\Psi_0\rangle = E_0|\Psi_0\rangle, \tag{2.13}$$

and by premultiplying the resulting equation by $e^{\Sigma^{\dagger}}e^{-T}$, while utilizing the fact that $|\Psi_0\rangle = e^T |\Phi\rangle = e^T e^{-\Sigma^{\dagger}} |\Phi\rangle$ (Σ^{\dagger} is a deexcitation operator, so that (Σ^{\dagger})^k $|\Phi\rangle = 0$ for k > 0). In general, the $\langle \tilde{\Phi} |$ dual state entering Eq. (2.12) depends on the values of T and Σ^{\dagger} defining \overline{H} . In the ECC theory of Piecuch and Bartlett, we simply require that T and Σ^{\dagger} are such that $\langle \tilde{\Phi} | = \langle \Phi |$. Thus, the final system of equations used to determine the two different cluster operators T and Σ^{\dagger} consists of Eq. (2.11) and

$$\langle \Phi | \bar{\bar{H}} = \langle \Phi | E_0, \qquad (2.14)$$

which is the left-hand counterpart of Eq. (2.11). It is worth mentioning that Eq. (2.14) can also be obtained by considering the bra counterpart of the connected cluster form of the electronic Schrödinger equation,

$$\bar{H}|\Phi\rangle = E_0|\Phi\rangle,$$
 (2.15)

where \bar{H} is defined by Eq. (2.3), i.e.

$$\langle \Phi | (1+\Lambda)\bar{H} = E_0 \langle \Phi | (1+\Lambda), \qquad (2.16)$$

where Λ is the well-known "lambda" deexcitation operator of the analytic gradient CC theory^{61,62}, and by identifying the left-hand ground eigenstate of \overline{H} , $\langle \Phi | (1 + \Lambda)$, with $\langle \Phi | e^{\Sigma^{\dagger}}$. An alternative reasoning that leads to Eq. (2.14) is based on considering the adjoint form of the electronic Schrödinger equation, i.e.

$$\langle \tilde{\Psi}_0 | H = E_0 \langle \tilde{\Psi}_0 |, \qquad (2.17)$$

where the dual state $\langle \tilde{\Psi}_0 |$, satisfying the condition $\langle \tilde{\Psi}_0 | \Psi_0 \rangle = 1$, is the CC bra ground state^{60,63,64},

$$\langle \tilde{\Psi}_0 | = \langle \Phi | (1+\Lambda) e^{-T}.$$
(2.18)

Clearly, in the exact, full CI or full CC, limit, there exists a deexcitation operator Σ^{\dagger} , such that $(1 + \Lambda) = e^{\Sigma^{\dagger}}$, so that one can always give the dual CC state $\langle \tilde{\Psi}_0 |$ a completely bi-exponential form,

$$\langle \tilde{\Psi}_0 | = \langle \Phi | e^{\Sigma^{\dagger}} e^{-T}.$$
(2.19)

By inserting Eq. (2.19) into Eq. (2.17) and by multiplying the resulting equation on the right by $e^T e^{-\Sigma^{\dagger}}$, we obtain the desired Eq. (2.14).

In the original work by Arponen and Bishop⁵⁰⁻⁵⁹, the operators T and Σ^{\dagger} of the ECC formalism are determined by imposing the stationary conditions on the energy functional $E_0^{(\text{ECC})}$, Eq. (2.1), with respect to operators T and Σ^{\dagger} ,

$$\frac{\delta E_0^{(\text{ECC})}}{\delta T} = 0, \quad \frac{\delta E_0^{(\text{ECC})}}{\delta \Sigma^{\dagger}} = 0.$$
(2.20)

The bi-variational character of the ECC theory of Arponen and Bishop is particularly useful in calculations of molecular properties other than energy, since one can apply the Hellmann-Feynmann theorem in such calculations⁶⁵⁻⁷³. The question addressed in this thesis is how the bi-variational ECC theory of Arponen and Bishop and the ECC theory of Piecuch and Bartlett, which uses Eqs. (2.11) and (2.14) rather than Eq. (2.20) to determine T and Σ^{\dagger} , work when molecular PESs along bond breaking coordinates are examined.

2.2 Extended Coupled-Cluster Methods with Singles and Doubles

2.2.1 The Piecuch-Bartlett ECCSD Approach

The approximate ECC methods, such as the ECCSD approaches tested in this work and developed in Refs. 74–77, are obtained as follows: First, as in all standard CC approximations, we truncate the many-body expansions of T and Σ^{\dagger} , Eqs. (2.5) and (2.8), respectively, at some excitation level $m_A < N$, so that T is replaced by $T^{(A)}$,

$$T^{(A)} = \sum_{n=1}^{m_A} T_n, \qquad (2.21)$$

and Σ^{\dagger} is replaced by

$$\Sigma^{(A)^{\dagger}} = \sum_{n=1}^{m_A} \Sigma_n^{\dagger}. \tag{2.22}$$

Next, we use either the Piecuch-Bartlett approach (Eqs. (2.11) and (2.14)) or the Arponen-Bishop approach (Eq. (2.20)) to obtain a system of equations for the unknown cluster amplitudes $t_{a_1...a_n}^{i_1...i_n}$ and $\sigma_{i_1...i_n}^{a_1...a_n}$, $n = 1, ..., m_A$. Once $t_{a_1...a_n}^{i_1...i_n}$ and $\sigma_{i_1...i_n}^{a_1...a_n}$ are determined, we use the approximate form of the energy functional (2.1),

$$E_0^{(A)} = \langle \Phi | \bar{\bar{H}}^{(A)} | \Phi \rangle, \qquad (2.23)$$

where

$$\bar{\bar{H}}^{(A)} = e^{\Sigma^{(A)}^{\dagger}} (e^{-T^{(A)}} H e^{T^{(A)}}) e^{-\Sigma^{(A)}^{\dagger}}, \qquad (2.24)$$

-

to calculate the ground-state energy.

In the specific case of the ECCSD approach, T is approximated by the sum of one- and two-body components, T_1 and T_2 , respectively,

$$T \approx T^{(\text{ECCSD})} = T_1 + T_2, \tag{2.25}$$

where

$$T_1 = \sum_{i,a} t^i_a E^a_i \tag{2.26}$$

and

$$T_2 = \sum_{i < j, a < b} t^{ij}_{ab} E^{ab}_{ij}, \qquad (2.27)$$

with

$$E_i^a = c^a c_i \tag{2.28}$$

and

$$E_{ij}^{ab} = c^a c_i c^b c_j \tag{2.29}$$

representing the elementary single and double excitation operators, and t_a^i and t_{ab}^{ij} designating the corresponding singly and doubly excited cluster amplitudes. A similar truncation scheme is applied to operator Σ^{\dagger} , i.e.,

$$\Sigma^{\dagger} \approx (\Sigma^{(\text{ECCSD})})^{\dagger} = \Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger}, \qquad (2.30)$$

where

$$\Sigma_1^{\dagger} = \sum_{i,a} \sigma_i^a E_a^i \tag{2.31}$$

and

$$\Sigma_2^{\dagger} = \sum_{i < j, a < b} \sigma_{ij}^{ab} E_{ab}^{ij}, \qquad (2.32)$$

with E_a^i and E_{ab}^{ij} designating the one- and two-body elementary deexcitation operators, $E_a^i = (E_i^a)^{\dagger}$ and $E_{ab}^{ij} = (E_{ij}^{ab})^{\dagger}$, respectively, and σ_i^a and σ_{ij}^{ab} representing the corresponding deexcitation amplitudes.

In the ECCSD method of Piecuch and Bartlett⁶⁰, we obtain the equations for the cluster amplitudes t_a^i , t_{ab}^{ij} , σ_i^a , and σ_{ij}^{ab} , defining T_1 , T_2 , Σ_1^{\dagger} , and Σ_2^{\dagger} , respectively, by left- and right-projecting Eqs. (2.11) and (2.14), where T is given by Eq. (2.25) and Σ^{\dagger} is given by Eq. (2.30), on the singly and doubly excited determinants $|\Phi_i^a\rangle = E_i^a |\Phi\rangle$ and $|\Phi_{ij}^{ab}\rangle = E_{ij}^{ab} |\Phi\rangle$. We obtain,

$$\langle \Phi_i^a | \bar{H}^{(\text{ECCSD})} | \Phi \rangle = 0,$$
 (2.33)

$$\langle \Phi_{ij}^{ab} | \bar{\bar{H}}^{(\text{ECCSD})} | \Phi \rangle = 0, \quad i < j, \ a < b,$$
(2.34)

$$\langle \Phi | \bar{\bar{H}}^{(\text{ECCSD})} | \Phi_i^a \rangle = 0, \qquad (2.35)$$

$$\langle \Phi | \bar{\bar{H}}^{(\text{ECCSD})} | \Phi_{ij}^{ab} \rangle = 0, \quad i < j, \ a < b,$$
(2.36)

where

$$\bar{\bar{H}}^{(\text{ECCSD})} = e^{\sum_{1}^{\dagger} + \sum_{2}^{\dagger}} (e^{-T_{1} - T_{2}} H e^{T_{1} + T_{2}}) e^{-\sum_{1}^{\dagger} - \sum_{2}^{\dagger}}$$
$$= e^{\sum_{1}^{\dagger} + \sum_{2}^{\dagger}} \bar{H}^{(\text{CCSD})} e^{-\sum_{1}^{\dagger} - \sum_{2}^{\dagger}}$$
(2.37)

is the doubly transformed Hamiltonian of the ECCSD method, with

$$\bar{H}^{(\text{CCSD})} = e^{-T_1 - T_2} H e^{T_1 + T_2} = (H e^{T_1 + T_2})_C$$
(2.38)

representing the similarly transformed Hamiltonian of the CCSD approximation. Once Eqs. (2.33)–(2.36) are solved for operators T_1 , T_2 , Σ_1^{\dagger} and Σ_2^{\dagger} , the ground-state energy $E_0^{(\text{ECCSD})}$ is calculated as follows:

$$E_{0}^{(\text{ECCSD})} = \langle \Phi | \bar{\bar{H}}^{(\text{ECCSD})} | \Phi \rangle$$

= $\langle \Phi | e^{\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger}} (e^{-T_{1} - T_{2}} H e^{T_{1} + T_{2}}) e^{-\Sigma_{1}^{\dagger} - \Sigma_{2}^{\dagger}} | \Phi \rangle$
= $\langle \Phi | e^{\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger}} (e^{-T_{1} - T_{2}} H e^{T_{1} + T_{2}}) | \Phi \rangle,$ (2.39)

where we used the fact that $e^{-\Sigma^{\dagger}} |\Phi\rangle = |\Phi\rangle$.

Since the full ECCSD formalism defined by Eqs. (2.33)-(2.39) is rather complex, in this thesis, in addition to the full ECCSD method, we consider the systematic sequence of the linear (LECCSD), bi-linear (BECCSD), and quadratic (QECCSD) approaches, which represent approximations to full ECCSD. The LECCSD, BECCSD, and QECCSD methods are obtained by replacing the doubly transformed Hamiltonian $\bar{H}^{(ECCSD)}$, Eq. (2.37), in Eqs. (2.33)-(2.36) and (2.39) by

$$\bar{\bar{H}}^{(\text{LECCSD})} = (1 + \Sigma_1^{\dagger} + \Sigma_2^{\dagger})\bar{H}^{(\text{CCSD})}(1 - \Sigma_1^{\dagger} - \Sigma_2^{\dagger}), \qquad (2.40)$$

in the LECCSD case,

$$\bar{\bar{H}}^{(\text{BECCSD})} = (1 + \Sigma_1^{\dagger} + \Sigma_2^{\dagger})\bar{H}^{(\text{CCSD})}(1 - \Sigma_1^{\dagger} - \Sigma_2^{\dagger}) + \frac{1}{2}(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})^2\bar{H}^{(\text{CCSD})} + \frac{1}{2}\bar{H}^{(\text{CCSD})}(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})^2, \qquad (2.41)$$

in the BECCSD case, and

$$\bar{\bar{H}}^{(\text{QECCSD})} = \left[1 + \Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger} + \frac{1}{2}(\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger})^{2}\right]\bar{H}^{(\text{CCSD})}$$
$$\times \left[1 - \Sigma_{1}^{\dagger} - \Sigma_{2}^{\dagger} + \frac{1}{2}(\Sigma_{1}^{\dagger} + \Sigma_{1}^{\dagger})^{2}\right], \qquad (2.42)$$

in the QECCSD case.

2.2.2 The Arponen-Bishop ECCSD Approach

In the case of the ECCSD method of Arponen and Bishop⁵⁰⁻⁵⁹, we obtain the system of equations for the singly and doubly excited cluster amplitudes t_a^i and t_{ab}^{ij} defining T_1 and T_2 and the deexcitation amplitudes σ_i^a and σ_{ij}^{ab} defining Σ_1^{\dagger} and Σ_2^{\dagger} from the stationary condition represented by Eq. (2.20). The resulting equations can be given in the following form:

$$\frac{\partial E_0^{(\text{ECCSD})}}{\partial \sigma_i^a} = \langle \Phi_i^a | e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}} \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \qquad (2.43)$$

$$\frac{\partial E_0^{(\text{ECCSD})}}{\partial \sigma_{ij}^{ab}} = \langle \Phi_{ij}^{ab} | e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}} \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \quad i < j, \ a < b, \tag{2.44}$$

$$\frac{\partial E_0^{(\text{ECCSD})}}{\partial t_a^i} = \langle \Phi | e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}} [\bar{H}^{(\text{CCSD})}, E_i^a] | \Phi \rangle$$
$$= \langle \Phi | e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}} (\bar{H}^{(\text{CCSD})} E_i^a)_C | \Phi \rangle = 0, \qquad (2.45)$$

$$\frac{\partial E_0^{(\text{ECCSD})}}{\partial t_{ab}^{ij}} = \langle \Phi | e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}} [\bar{H}^{(\text{CCSD})}, E_{ij}^{ab}] | \Phi \rangle$$
$$= \langle \Phi | e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}} (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_C | \Phi \rangle = 0, \quad i < j, \ a < b, \quad (2.46)$$

where $E_0^{(\text{ECCSD})}$ is the ECCSD energy functional, Eq. (2.39), and $\bar{H}^{(\text{CCSD})}$ is the similarity-transformed Hamiltonian defined by Eq. (2.38).

It can be shown that the full ECCSD method, defined by Eqs. (2.43)-(2.46), leads to computational steps that scale as \mathcal{N}^{10} with the system size. This is a lot more than the \mathcal{N}^6 steps of the standard CCSD approach. In order to reduce the computer costs of the full ECCSD calculations, we must truncate the power series expansion for $e^{\Sigma_1^{\dagger} + \Sigma_2^{\dagger}}$ in the ECCSD energy functional $E_0^{(\text{ECCSD})}$, Eq. (2.39), at the low power of $(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})$. The simplest approximation of this type that one might suggest is the linearized ECCSD (LECCSD) formalism, in which we replace the energy functional $E_0^{(\text{ECCSD})}$, Eq. (2.39), by the expression linear in Σ_1^{\dagger} and Σ_2^{\dagger} ,

$$E_{0}^{(\text{LECCSD})} = \langle \Phi | (1 + \Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger}) \bar{H}^{(\text{CCSD})} | \Phi \rangle$$

$$= \langle \Phi | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \langle \Phi | \Sigma_{1}^{\dagger} \bar{H}^{(\text{CCSD})} | \Phi \rangle + \langle \Phi | \Sigma_{2}^{\dagger} \bar{H}^{(\text{CCSD})} | \Phi \rangle$$

$$= \langle \Phi | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \sum_{i,a} \sigma_{i}^{a} \langle \Phi_{i}^{a} | \bar{H}^{(\text{CCSD})} | \Phi \rangle$$

$$+ \sum_{i < j,a < b} \sigma_{ij}^{ab} \langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle. \qquad (2.47)$$

The most expensive steps of the LECCSD method, as defined by Eq. (2.47), scale as $n_o^2 n_u^4$ (or \mathcal{N}^6 with the system size), but, unfortunately, the LECCSD method based

on the Arponen-Bishop theory does not improve the CCSD results at all, since once we impose the stationary conditions on the energy functional $E_0^{(\text{LECCSD})}$ with respect to operators T_1 , T_2 , Σ_1^{\dagger} , and Σ_2^{\dagger} , we obtain the following equations:

$$\frac{\partial E_0^{(\text{LECCSD})}}{\partial \sigma_i^a} = \langle \Phi_i^a | \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \qquad (2.48)$$

$$\frac{\partial E_0^{(\text{LECCSD})}}{\partial \sigma_{ij}^{ab}} = \langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \quad i < j, \ a < b, \tag{2.49}$$

which are the usual CCSD equations for T_1 and T_2 , and

$$\frac{\partial E_0^{(\text{LECCSD})}}{\partial t_a^i} = \langle \Phi | (1 + \Sigma_1^{\dagger} + \Sigma_2^{\dagger}) \left(\bar{H}^{(\text{CCSD})} E_i^a \right)_C | \Phi \rangle = 0, \qquad (2.50)$$

$$\frac{\partial E_0^{(\text{LECCSD})}}{\partial t_{ab}^{ij}} = \langle \Phi | (1 + \Sigma_1^{\dagger} + \Sigma_2^{\dagger}) (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_C | \Phi \rangle = 0,$$

$$i < j, \ a < b, \qquad (2.51)$$

which are the Σ^{\dagger} equations that are completely decoupled from the CCSD system, Eqs. (2.48) and (2.49), and solved only after the T_1 and T_2 cluster are determined. In other words, the LECCSD approximation does not provide for the coupling between the T and Σ^{\dagger} equations, which is necessary for improving the quality of T_1 and T_2 clusters in the bond breaking region.

The above analysis indicates that in order to obtain the T_1 and T_2 clusters, which are better than those provided by the LECCSD = CCSD approximation, we must introduce nonlinear terms in Σ_1^{\dagger} and Σ_2^{\dagger} into the LECCSD energy functional that couple the T and Σ^{\dagger} equations. The lowest-order approximation of this type is obtained by truncating the ECCSD energy functional, Eq. (2.39), at terms quadratic in $(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})$. We call it the QECCSD approximation. The formula for the QECCSD energy functional used in the Arponen-Bishop version of the QECCSD theory is

$$E_0^{(\text{QECCSD})} = \langle \Phi | [1 + (\Sigma_1^{\dagger} + \Sigma_2^{\dagger}) + \frac{1}{2} (\Sigma_1^{\dagger} + \Sigma_2^{\dagger})^2] \bar{H}^{(\text{CCSD})} | \Phi \rangle$$

$$= \langle \Phi | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \langle \Phi | \Sigma_{1}^{\dagger} \bar{H}^{(\text{CCSD})} | \Phi \rangle$$
$$+ \langle \Phi | [\Sigma_{2}^{\dagger} + \frac{1}{2} (\Sigma_{1}^{\dagger})^{2}] \bar{H}^{(\text{CCSD})} | \Phi \rangle$$
$$+ \langle \Phi | \Sigma_{1}^{\dagger} \Sigma_{2}^{\dagger} \bar{H}^{(\text{CCSD})} | \Phi \rangle + \frac{1}{2} \langle \Phi | (\Sigma_{2}^{\dagger})^{2} \bar{H}^{(\text{CCSD})} | \Phi \rangle$$
(2.52)

or, somewhat more explicitly,

$$\begin{split} E_{0}^{(\text{QECCSD})} &= \langle \Phi | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \sum_{i, a} \sigma_{i}^{a} \langle \Phi_{i}^{a} | \bar{H}^{(\text{CCSD})} | \Phi \rangle \\ &+ \sum_{i < j} (\sigma_{ij}^{ab} + \mathcal{A}_{ij} \sigma_{i}^{a} \sigma_{j}^{b}) \langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle \\ &+ \sum_{i < j < k} \mathcal{A}_{i/jk} \mathcal{A}_{a/bc} \sigma_{i}^{a} \sigma_{jk}^{bc} \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} | \Phi \rangle \\ &+ \sum_{i < j < k} \mathcal{A}_{i/jk} \mathcal{A}_{a/bc} \sigma_{ij}^{a} \sigma_{jk}^{bc} \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSD})} | \Phi \rangle \\ &+ \sum_{i < j < k < l} \mathcal{A}_{ij/kl} \mathcal{A}_{b/cd} \sigma_{ij}^{ab} \sigma_{kl}^{cd} \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSD})} | \Phi \rangle, (2.53) \end{split}$$

where $\mathcal{A}_{ij} = 1 - (ij)$, $\mathcal{A}_{i/jk} = 1 - (ij) - (ik)$, $\mathcal{A}_{a/bc} = 1 - (ab) - (ac)$, and $\mathcal{A}_{ij/kl} = 1 - (ik) - (il) - (jk) - (jl) + (ik)(jl)$ are the suitable index antisymmetrizers ((ij) and (ab) are index interchanges). The stationary conditions of the type of Eqs. (2.43)-(2.46), written for the QECCSD energy functional $E_0^{(QECCSD)}$, Eq. (2.52) or (2.53), are

$$\frac{\partial E_{0}^{(\text{QECCSD})}}{\partial \sigma_{i}^{a}} = \langle \Phi_{i}^{a} | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \sum_{j, b} \sigma_{j}^{b} \langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \sum_{j < k, b < c} \sigma_{jk}^{bc} \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \qquad (2.54)$$

$$\frac{\partial E_{0}^{(\text{QECCSD})}}{\partial \sigma_{ij}^{ab}} = \langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle + \sum_{k,c} \sigma_{k}^{c} \langle \Phi_{ijk}^{abc} | \bar{H}^{(\text{CCSD})} | \Phi \rangle
+ \sum_{k < l, c < d} \sigma_{kl}^{cd} \langle \Phi_{ijkl}^{abcd} | \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0,$$
(2.55)

$$\frac{\partial E_{0}^{(\text{QECCSD})}}{\partial t_{a}^{i}} = \langle \Phi | [1 + (\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger}) + \frac{1}{2} (\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger})^{2}] (\bar{H}^{(\text{CCSD})} E_{i}^{a})_{C} | \Phi \rangle$$

$$= \langle \Phi | (\bar{H}^{(\text{CCSD})} E_{i}^{a})_{C} | \Phi \rangle + \sum_{j,b} \sigma_{j}^{b} \langle \Phi_{j}^{b} | (\bar{H}^{(\text{CCSD})} E_{i}^{a})_{C} | \Phi \rangle$$

$$+ \sum_{j < k, b < c} (\sigma_{jk}^{bc} + \mathcal{A}_{jk} \sigma_{j}^{b} \sigma_{k}^{c}) \langle \Phi_{jk}^{bc} | (\bar{H}^{(\text{CCSD})} E_{i}^{a})_{C} | \Phi \rangle$$

$$+ \sum_{j < k < l} \mathcal{A}_{j/kl} \mathcal{A}_{b/cd} \sigma_{j}^{b} \sigma_{kl}^{cd} \langle \Phi_{jkl}^{bcd} | (\bar{H}^{(\text{CCSD})} E_{i}^{a})_{C} | \Phi \rangle$$

$$+ \sum_{j < k < l} \mathcal{A}_{jk/lm} \mathcal{A}_{c/de} \sigma_{jk}^{bc} \sigma_{lm}^{de} \langle \Phi_{jklm}^{bcde} | (\bar{H}^{(\text{CCSD})} E_{i}^{a})_{C} | \Phi \rangle$$

$$= 0, \qquad (2.56)$$

$$\frac{\partial E_{0}^{(\text{QECCSD})}}{\partial t_{ab}^{ij}} = \langle \Phi | [1 + (\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger}) + \frac{1}{2} (\Sigma_{1}^{\dagger} + \Sigma_{2}^{\dagger})^{2}] (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_{C} | \Phi \rangle$$

$$= \langle \Phi | (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_{C} | \Phi \rangle + \sum_{k,c} \sigma_{k}^{c} \langle \Phi_{k}^{c} | (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_{C} | \Phi \rangle$$

$$+ \sum_{k < l, c < d} (\sigma_{kl}^{cd} + \mathcal{A}_{kl} \sigma_{k}^{c} \sigma_{l}^{d}) \langle \Phi_{kl}^{cd} | (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_{C} | \Phi \rangle$$

$$+ \sum_{k < l < m} \mathcal{A}_{k/lm} \mathcal{A}_{c/de} \sigma_{k}^{c} \sigma_{lm}^{de} \langle \Phi_{klm}^{cde} | (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_{C} | \Phi \rangle$$

$$+ \sum_{k < l < m} \mathcal{A}_{kl/mn} \mathcal{A}_{d/ef} \sigma_{kl}^{cd} \sigma_{mn}^{ef} \langle \Phi_{klmn}^{cdef} | (\bar{H}^{(\text{CCSD})} E_{ij}^{ab})_{C} | \Phi \rangle$$

$$= 0. \qquad (2.57)$$

As one can see, the QECCSD equations for the t_a^i and t_{ab}^{ij} amplitudes, Eqs. (2.54) and (2.55), respectively, are strongly and nontrivially coupled with the equations for the σ_i^a and σ_{ij}^{ab} amplitudes, Eqs. (2.56) and (2.57), respectively, so that unlike in the standard CCSD approach, the T_1 and T_2 clusters are no longer calculated independent of the Σ_1^{\dagger} and Σ_2^{\dagger} clusters. Similarly strong coupling is present in the QECCSD, BECCSD, and ECCSD methods based on the Piecuch-Bartlett ECC theory described in Section 2.2.1. As shown below, a strong coupling between T and Σ^{\dagger} equations has a positive effect on improving the quality of the T_1 and T_2 clusters resulting from the QECCSD and similar calculations. At the same time, the most expensive steps of the QECCSD approach based on Eqs. (2.54)-(2.57) scale as n_u^6 (\mathcal{N}^6 with the system size). Similar remarks apply to the QECCSD and BECCSD methods based on the Piecuch-Bartlett approach discussed in Section 2.2.1.

2.3 Numerical Results for Multiple Bond Breaking

2.3.1 The Piecuch-Bartlett ECCSD Approach

In this section, we discuss the results of the full and approximate ECCSD calculations for the minimum basis set (MBS) S4 model⁷⁸⁻⁸⁰ and the STO-3G⁸¹ model of N₂ employing the ECC theory of Piecuch and Bartlett. All ECCSD, LECCSD, BECCSD, and QECCSD calculations reported here were performed using the original computer codes developed in this work, in which the relevant cluster amplitudes t_a^i , t_{ab}^{ij} , σ_i^a , and σ_{ij}^{ab} are determined by minimizing the sum of the squares of the equations constituting the ECCSD system, Eqs. (2.33)–(2.36), with the downhill simplex method⁸². This algorithm is based on an obvious fact that the global minimum being sought is zero. For smaller values of the parameter α describing the S4 model ($\alpha \approx 2.0$ bohr) and in the vicinity of the equilibrium geometry of N₂ (the N–N separation $R \approx 2.0$ bohr), where the standard CCSD approach provides a reasonable description of the groundstate wave function, the initial guesses for cluster amplitudes t_a^i , t_{ab}^{ij} , σ_i^a , and σ_{ij}^{ab} were obtained by using the CCSD values for t_a^i and t_{ab}^{ij} and by assuming that the initial values of σ_i^a and σ_{ij}^{ab} satisfy the conditions $\sigma_i^a = t_a^i$ and $\sigma_{ij}^{ab} = t_{ab}^{ij}$ (one can show that the first-order MBPT estimates of the cluster operator T and the auxiliary operator Σ are identical^{83,84}). For larger values of α and R, where T and Σ significantly differ, we used the previously converged t_a^i , t_{ab}^{ij} , σ_i^a , and σ_{ij}^{ab} amplitudes, obtained for smaller values of α and R, to initiate the numerical procedures used to determine T and Σ .

We begin the discussion of our ECCSD test calculations with the MBS S4 model, which consists of four hydrogen atoms arranged in a square configuration, described by a small basis set consisting of only one s orbital centered on each hydrogen atom $^{78-80}$. The geometry of the S4 model is determined by a single parameter α , which is defined as a distance between the nearest-neighbor hydrogen atoms. The small number of electrons and the fact that the molecular orbital basis set consists in this case of only four orbitals that are fully determined by the high spatial symmetry of the S4 model cause that there are only eight spin- and symmetry-adapted configurations that are relevant to the ground-state full CI problem. In addition to the RHF determinant, which we use as a reference, we only need six doubly excited configurations and one quadruply excited configuration to describe the exact, full CI, ground state of the MBS S4 model. There are no single and triple excitations in the full CI expansion of the ground-state wave function, so that $T_1 = T_3 = 0$ and the CCD, CCSD and CCSDT approximations become completely equivalent. Similarly, the one- and three-body components of the deexcitation operator Σ^{\dagger} vanish, so that the ECCD and ECCSD methods give identical results. Because of the presence of only four electrons in the MBS S4 model, there is no difference between the BECCSD = BECCD, QECCSD = QECCD, and full ECCSD = ECCD approximations in this case. Thus, we only report the results of the full ECCD calculations, which we compare with the CCD and exact full CI results.



Figure 2.1: Potential energy curves for the MBS S4 system. ECCD represents the Piecuch-Bartlett variant of the ECCD approach (in this case, the ECCD and ECCSD approaches are equivalent).

The MBS S4 models with larger values of α create a serious challenge for the standard single-reference CC methods (even the genuine MRCC approaches may have a difficulty in describing these models^{78,80}). This is related to the fact that larger α values correspond to a dissociation of the S4 model into four hydrogen atoms. This process is very difficult to describe by the RHF-based single-reference methods, particularly in the region of the intermediate α values where the ground-state wave function of the S4 system undergoes a significant rearrangement of its structure. As shown in Table 2.1 and Figure 2.1, the CCD = CCSD approximation breaks down at larger values of α . For $\alpha \leq 2.0$ bohr, the unsigned errors in the CCD results, relative to full CI, do not exceed 2.834 millihartree. However, for 3.5 bohr $\leq \alpha \leq$ 4.5 bohr, the unsigned errors in the CCD results increase to 13-14 millihartree and the CCD potential energy curve corresponding to the dissociation of the S4 model into four hydrogen atoms goes significantly below the exact, full CI, curve (see Figure 2.1). For $\alpha \leq 2.0$ bohr, the ground-state wave function of the MBS S4 model is dominated by two configurations: the RHF determinant and the doubly excited determinant corresponding to the HOMO \rightarrow LUMO biexcitation⁷⁸⁻⁸⁰. For larger values of α , essentially all electron configurations present in the full CI expansion become very important, creating a strongly quasi-degenerate situation, and the role of the T_4 cluster component becomes significant in the intermediate 2.5 bohr $\leq \alpha \leq 6.0$ bohr region. This leads to the failure of the CCD approximation observed in Table 2.1 and Figure 2.1.

α	Full CI	CCD P	ECCD ^c	$\langle CCD \rangle^d$	(ECCD) ^d
0.5	3.952114	-0.093	0.015	0.032	0.015
1.0	-0.668783	-0.424	0.110	0.236	0.110
1.5	-1.694327	-1.235	0.376	0.862	0.376
2 .0	-1.975862	-2.834	0.871	2.188	0.871
2.5	-2.043797	-5.674	1.609	4.447	1.609
3.0	-2.044850	-9.568	2.373	6.895	2.373
3.5	-2.028186	-13.083	2.754	7.698	2.754
4.0	-2.011713	-14.385	2.496	6.218	2.496
4.5	-2.000438	-13.094	1.812	3.896	1.812
5.0	-1.994021	-10.292	1.107	2.054	1.107
6.0	-1.989199	-4.748	0.297	0.431	0.296
7.0	-1.988164	-1.743	0.062	0.076	0.061

Table 2.1.: Ground-state energies of the MBS S4 model system as functions of the parameter α .^a

^aThe full CI total energies are in hartree. The CCD and ECCD energies are in millihartree relative to the corresponding full CI energy values. The parameter α is in bohr.

^bFor the MBS S4 model, CCD = CCSD.

^cThe ECCD method of Piecuch and Bartlett. For the MBS S4 model, ECCD = ECCSD = BECCSD = QECCSD.

 ${}^{d}\langle X \rangle$ (X = CCD, ECCD) is the expectation value of the Hamiltonian with the $e^{T_2} |\Phi\rangle$ wave function, where T_2 is obtained with method X (see Eq. (2.58)).

The results in Table 2.1 and Figure 2.1 show that the ECCD theory provides substantial improvements in the poor description of the dissociation of the S4 model into four hydrogen atoms by the CCD method. The 13–14 millihartree unsigned errors in the CCD energies in the 3.5 bohr $\leq \alpha \leq 4.5$ bohr region reduce to 2–3 millihartree, when the ECCD approach is employed. The considerable reduction of errors in the CCD results is observed at all values of α , even in the $\alpha \leq 2.0$ bohr region, where the maximum unsigned error in the ECCD results is 0.871 millihartree, as opposed to 2.834 millihartree obtained with the CCD approach (see Table 2.1). Unlike the CCD potential energy curve shown in Figure 2.1, which is located significantly below the full CI curve, the ECCD potential energy curve describing the dissociation of the S4 system into four H atoms is located slightly above the full CI curve. Thus, in spite of its formally nonvariational character, the ECCD approach based on the ECC theory of Piecuch and Bartlett⁶⁰ provides a highly accurate and variational description of the breaking of all four H–H bonds in the S4 system.

The substantial improvement in the description of the S4 model system offered by the ECCD approach suggests that the T_2 clusters resulting from the ECCD calculations are considerably better than the CCD T_2 values. This can be seen by calculating the expectation values of the Hamiltonian, designated by $\langle X \rangle$, where X = CCD or ECCD, with the normalized CCD-like wave functions

$$|\Psi_0^{(X)}\rangle = N^{(X)} e^{T_2^{(X)}} |\Phi\rangle, \quad (X = \text{CCD}, \text{ ECCD}),$$
 (2.58)

where $T_2^{(X)}$ (X = CCD, ECCD) are the T_2 cluster components obtained with the CCD and ECCD methods, respectively, and $N^{(X)} = \langle \Phi | e^{(T_2^{(X)})^{\dagger}} e^{T_2^{(X)}} | \Phi \rangle^{-1/2}$ are the appropriate normalization factors. Clearly, the $\langle \text{CCD} \rangle$ and $\langle \text{ECCD} \rangle$ values provide the upper bounds to the exact, full CI, energies. However, as demonstrated in Table 2.1, the $\langle \text{CCD} \rangle$ energies remain poor in the 3.5 bohr $\leq \alpha \leq 4.5$ bohr region, whereas

the expectation values of the Hamiltonian calculated with the ECCD wave function $|\Psi_0^{(\text{ECCD})}\rangle$, Eq. (2.58), are very close to the corresponding full CI energies.

The MBS S4 model is so simple that we can clearly understand the reasons of the excellent performance of the ECCSD = ECCD theory at larger α values. The MBS S4 system has only four electrons and the T_1 and T_3 components vanish due to the high symmetry of the Hamiltonian. Thus, in order to obtain a highly accurate description of the ground electronic state of the MBS S4 system, we must use a method which is capable of providing an accurate description of the effects due to connected quadruply excited (T_4) cluster components, which are missing in CCSD. It turns out that at least some of the T_4 effects are brought into the ECC formalism as products of the low-order many-body components of Σ^{\dagger} and T. Indeed, as shown, for example, in Ref. 23, the leading, fifth-order, contribution to the energy due to T_4 clusters can be estimated by adding the $E_{QQ}^{[5]}$ term, defined as

$$E_{\rm QQ}^{[5]} = \frac{1}{4} \langle \Phi | (T_2^{\dagger})^2 (V_N T_2^2)_C | \Phi \rangle, \qquad (2.59)$$

to

$$E_{\rm QT}^{[5]} = \frac{1}{2} \langle \Phi | (T_2^{\dagger})^2 (V_N T_3^{[2]})_C | \Phi \rangle, \qquad (2.60)$$

where

$$T_3^{[2]} = R_0^{(3)} (V_N T_2)_C. (2.61)$$

In the above expression, $R_0^{(3)}$ designates the three-body part of the MBPT reduced resolvent and V_N represents the two-body part of the Hamiltonian in the normal-ordered form, $H_N = H - \langle \Phi | H | \Phi \rangle$. For the MBS S4 model, the $E_{QT}^{[5]}$ energy component, Eq. (2.60), vanishes, since $T_3^{[2]} = 0$. Thus, the entire fifth-order effect due to T_4 can be estimated in this case by the $E_{QQ}^{[5]}$ contribution, Eq. (2.59), which appears in the ECCD energy as the $\frac{1}{4} \langle \Phi | (\Sigma_2^{\dagger})^2 (V_N T_2^2)_C | \Phi \rangle$ term, since T_2 and Σ_2 are similar when MBPT converges. This means that the ECCD energy for the MBS S4 model contains a great deal of information about the leading effects due to T_4 and these are sufficient to provide the excellent results at all values of the parameter α observed in Figure 2.1 and Table 2.1.

The MBS S4 model allows us to obtain useful insights into the performance of the ECC approximations in the calculations for quasi-degenerate electronic states, but we cannot use it to test all important aspects of the ECC theory. For example, the MBS S4 system is too simple to analyze the importance of the T_1 and Σ_1^{\dagger} components and it does not allow us to understand the significance of terms that distinguish the LECCSD, BECCSD, QECCSD, and full ECCSD approximations defined in Section 2.2.1. Moreover, multiple bond breaking in real molecules can be considerably more complicated than the dissociation of the H₄ cluster represented by the S4 system into four hydrogen atoms. A good example of the very challenging situation, which is considerably more complex than the situation created by the S4 model, is provided by the triple bond breaking in N₂, where the standard CCSD approach displays a colossal failure (see Table 2.2 and Figure 2.2). We tested the ECCSD, LECCSD, BECCSD, and QECCSD methods, based on Eqs. (2.33)–(2.42), using the minimum basis set STO-3G⁸¹ model of N₂. In all correlated calculations for N₂ discussed below, the lowest two core orbitals were kept frozen.

R ^b	Full CI	CCSD	BECCSD ^c	QECCSD ^c	ECCSD
1.0	-101.791600	0.319	0.298	0.298	0.298
1.5	-106.720117	1.102	0.885	0.885	0.885
2.0	-107.623240	3.295	1.897	1.897	1.897
2.5	-107.651880	9.220	3.442	3.442	3.428
3.0	-107.546614	13.176	3.919	3.908	3.758
3.5	-107.473442	-38.645	5.280	5.322	4.746
4.0	-107.447822	-140.376	15.580	15.968	14.122
4.5	-107.441504	-184.984	26.7 95	27.769	24.039
5.0	-107.439549	-200.958	34.134	35.732	30.390
5.5	-107.438665	-206.974	38.368	40.491	33.867
6.0	-107.438265	-209.538	40.730	43.227	35.746
7.0	-107.438054	-211.915	42.754	45.595	37.306
8.0	-107.438029	-213.431	43.405	46.320	37.799

Table 2.2.: Ground-state energies of the N_2 molecule, as described by the STO-3G basis set.^a

^aThe full CI total energies are in hartree. The CC and ECC energies are in millihartree relative to the corresponding full CI energy values. The lowest two occupied orbitals were frozen in the correlated calculations.

^bThe N-N separation in bohr. The equilibrium value of R is 2.068 bohr.

^cThe BECCSD, QECCSD, and full ECCSD methods are based on the Piecuch-Bartlett variant of the ECC theory.



Figure 2.2: Potential energy curves for the N_2 molecule, as described by the STO-3G basis set. The LECCSD, BECCSD, QECCSD, and ECCSD methods are based on the Piecuch-Bartlett variant of the ECC theory.
The results in Table 2.2 and Figure 2.2 clearly demonstrate that the complete ECCSD formalism of Piecuch and Bartlett⁶⁰, in which all nonlinear terms in $(\Sigma_1^{\dagger} +$ Σ_2^{\dagger}) and $(T_1 + T_2)$ are included, and its bilinear and quadratic variants, BECCSD and QECCSD, respectively, defined by the truncated Hamiltonians $\bar{H}^{(\text{BECCSD})}$ and $\bar{H}^{(QECCSD)}$, Eqs. (2.41) and (2.42), respectively, provide remarkable improvements in the very poor description of the potential energy curve of N_2 by the standard CCSD method. The huge negative errors in the CCSD results at larger N-N separations, $R \ge 4.5$ bohr, of about -200 millihartree, reduce to much smaller positive errors when the ECC methods are employed (24-38 millihartree in the full ECCSD case, 27-43 millihartree in the BECCSD case, and 28–46 millihartree in the QECCSD case). We also observe a considerable reduction of errors for smaller values of R, including the equilibrium, $R \approx 2.0$ bohr, region (see Table 2.2). As shown in Table 2.2 and Figure 2.2, the BECCSD, QECCSD, and full ECCSD approaches eliminate the pathological behavior of the CCSD method at larger N-N distances. As in the MBS S4 case, the BECCSD, QECCSD, and full ECCSD approaches restore the variational description of the potential energy curve of N_2 at all internuclear separations.

Our results for N₂ show that it is not necessary to insist on the bi-variational determination of the Σ^{\dagger} and T operators, exploited in the Arponen-Bishop ECC formalism, in order to obtain great improvements in the description of multiple bond breaking by the ECC theory. The fact that the ECC theory of Piecuch and Bartlett is not rigorously bi-variational seems to be of little significance, since our BECCSD, QECCSD, and ECCSD results obtained with this theory are of the same quality as the strictly bi-variational QECCSD results for N₂ reported in subsection 2.3.2. The presence of quadratic terms in Σ^{\dagger} in $\overline{H}^{(\text{ECCSD})}$, $\overline{H}^{(\text{BECCSD})}$, and $\overline{H}^{(\text{QECCSD})}$, Eqs. (2.37), (2.41), and (2.42), respectively, and the use of two independent cluster operators T and Σ in the ECC formalism, which are optimized by solving a coupled system of equations, are more important for improving the results in the bond breaking region than the particular way of obtaining the t_a^i , t_{ab}^{ij} , σ_i^a , and σ_{ij}^{ab} amplitudes that defines the ECCSD formalism of Piecuch and Bartlett.

The importance of the quadratic terms in $(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})$, such as $\frac{1}{2}(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})^2 \bar{H}^{(\text{CCSD})}$ and $\frac{1}{2}\bar{H}^{(\text{CCSD})}(\Sigma_1^{\dagger}+\Sigma_2^{\dagger})^2$, in the ECCSD equations becomes apparent when we compare the BECCSD or QECCSD results with the results of the LECCSD calculations. These quadratic terms are ignored in the LECCSD method (see Eq. (2.40)) and, in consequence, the LECCSD potential energy curve for $N_{\mathbf{2}}$ has the same type of hump for the intermediate values of R as the CCSD curve (see Figure 2.2). It is interesting to observe, though, a substantial reduction of errors in the CCSD results at larger N-N separations, when the LECCSD approach is employed. This corroborates our earlier statements that the use of two independent cluster operators, T and Σ^{\dagger} , in the ECC theory is more important for improving the poor CCSD results at larger R values than the specific procedure used to determine Tand Σ^{\dagger} . It is also worth noticing that we can safely neglect the higher-order nonlinear terms, such as $\frac{1}{2} (\Sigma_1^{\dagger} + \Sigma_2^{\dagger})^2 \bar{H}^{(\text{CCSD})} (\Sigma_1^{\dagger} + \Sigma_2^{\dagger}), \frac{1}{2} (\Sigma_1^{\dagger} + \Sigma_2^{\dagger}) \bar{H}^{(\text{CCSD})} (\Sigma_1^{\dagger} + \Sigma_2^{\dagger})^2$, and $\frac{1}{4}(\Sigma_1^{\dagger}+\Sigma_2^{\dagger})^2 \bar{H}^{(\text{CCSD})}(\Sigma_1^{\dagger}+\Sigma_2^{\dagger})^2$, which are present in the QECCSD and full ECCSD methods (cf. Eqs. (2.42) and (2.37), respectively) and absent in the BECCSD approach (see Eq. (2.41)). The BECCSD results are of the same quality as the QECCSD and full ECCSD results (see Figure 2.2 and Table 2.2). The BECCSD theory represents the lowest-order ECC approach among various ECC methods examined in this work capable of providing the qualitatively correct description of triple bond breaking in N_2 .

Rb	Full CI	(CCSD) ^c	⟨BECCSD⟩ ^c	(QECCSD) ^c	(ECCSD) ^c
1.0	-101.791600	0.298	0.298	0.298	0.298
1.5	-106.720117	0.890	0.888	0.888	0.888
2.0	-107.623240	2.004	1.946	1.946	1.946
2.5	-107.651880	4.316	3.775	3.775	3.775
3.0	-107.546614	5.288	4.160	4.160	4.161
3.5	-107.473442	16.755	3.378	3.387	3.388
4.0	-107.447822	80.696	6.922	7.206	7.145
4.5	-107.441504	95.003	10.603	11.506	11.277
5.0	-107.439549	91.561	12.391	13.877	13.327
5.5	-107.438665	86.652	13.056	14.931	14.224
6.0	-107.438265	83.037	13.276	15.356	14.553
7.0	-107.438054	79.607	13.393	15.571	14.724
8.0	-107.438029	78.563	13.451	15.578	14.758

Table 2.3.: Ground-state energies of the N_2 molecule, as described by the STO-3G basis set.^a

^aThe full CI total energies are in hartree. The remaining energies are in millihartree relative to the corresponding full CI energy values. The lowest two occupied orbitals were frozen in the correlated calculations.

^bThe N-N separation in bohr. The equilibrium value of R is 2.068 bohr.

 ${}^{c}\langle X \rangle$ (X = CCSD, BECCSD, QECCSD, ECCSD) is the expectation value of the Hamiltonian with the $e^{T_1+T_2}|\Phi\rangle$ wave function, where T_1 and T_2 are obtained with method X (cf. Eq. (2.62)).

Finally, before describing the calculations employing the Arponen-Bishop ECC theory, let us discuss the quality of the T_1 and T_2 clusters resulting from various types of the ECCSD calculations for N₂. The remarkable improvements in the description of triple bond breaking in N₂, offered by the BECCSD, QECCSD, and full ECCSD methods, imply that the T_1 and T_2 clusters resulting from the bilinear, quadratic, and full ECCSD calculations are much more accurate than the T_1 and T_2 operators obtained with the standard CCSD approach. As in the case of the MBS S4 model, we examined the quality of the T_1 and T_2 clusters obtained in the CCSD and various ECC calculations for N₂ by computing the expectation values of the Hamiltonian, designated by $\langle X \rangle$, where X = CCSD, BECCSD, QECCSD, and ECCSD, with the normalized CCSD-like wave functions

$$\begin{split} |\Psi_0^{(X)}\rangle &= N^{(X)} e^{T_1^{(X)} + T_2^{(X)}} |\Phi\rangle, \ (X = \text{CCSD}, \text{BECCSD}, \text{QECCSD}, \text{ECCSD}), \ (2.62) \\ \text{where } T_1^{(X)} \text{ and } T_2^{(X)} \ (X = \text{CCSD}, \text{BECCSD}, \text{QECCSD}, \text{ECCSD}) \text{ are the } T_1 \text{ and } \\ T_2 \text{ cluster components obtained with the CCSD, BECCSD, QECCSD, and complete } \\ \text{ECCSD methods, respectively, and } N^{(X)} &= \langle \Phi | e^{(T_1^{(X)})^{\dagger} + (T_2^{(X)})^{\dagger}} e^{T_1^{(X)} + T_2^{(X)}} |\Phi\rangle^{-1/2} \text{ are the } \\ \text{corresponding normalization factors (see Table 2.3). As demonstrated in Table 2.3, } \\ \text{the expectation values of the Hamiltonian obtained with the CCSD wave function } \\ \text{are extremely poor at larger N-N separations, whereas the expectation values of } \\ \text{functions, } |\Psi_0^{(\text{BECCSD})}\rangle, |\Psi_0^{(\text{QECCSD})}\rangle, \text{ and } |\Psi_0^{(\text{ECCSD})}\rangle, \text{ respectively, are very close to } \\ \text{the corresponding full CI energies at all values of } R. The fact that we observe a fairly \\ \text{substantial } (2-3-\text{fold}) \text{ reduction of unsigned errors, when the standard CCSD energy \\ expression, E_0^{(\text{CCSD})} = \langle \Phi | \bar{H}^{(\text{CCSD})} | \Phi \rangle, \text{ is replaced by the expectation value of the \\ \text{Hamiltonian with the CCSD wave function, does not help the CCSD theory too much, \\ \text{since the } \langle \text{CCSD} \rangle \text{ energies are characterized by the very large, 79-95 millihartree, } \end{split}$$

errors in the $R \ge 4.0$ bohr region. When the BECCSD, QECCSD, and ECCSD energy expressions, based on Eq. (2.39), are replaced by the expectation values of the Hamiltonian with the BECCSD, QECCSD, and ECCSD wave functions, the 24–46 millihartree errors in the BECCSD, QECCSD, and ECCSD energies in the $R \ge 4.5$ bohr region reduce to 11–13 millihartree in the BECCSD case, 12–16 millihartree in the QECCSD case, and 11–15 millihartree in the full ECCSD case (see Table 2.3). The fact that we can improve the description of the potential energy curve of N₂ by simply replacing the BECCSD, QECCSD, and ECCSD energy expressions, based on Eq. (2.39), by the expectation values of the Hamiltonian calculated with the BECCSD, QECCSD, and ECCSD wave functions, $|\Psi_0^{(BECCSD)}\rangle$, $|\Psi_0^{(QECCSD)}\rangle$, and $|\Psi_0^{(ECCSD)}\rangle$, respectively, is interesting and worth further exploration.

The high quality of the T_1 and T_2 clusters resulting from the BECCSD, QECCSD, and full ECCSD calculations and the poor quality of the T_1 and T_2 clusters obtained with the standard CCSD approach can also be seen by examining the overlaps of the normalized CCSD, BECCSD, QECCSD, and ECCSD wave functions, $|\Psi_0^{(CCSD)}\rangle$, $|\Psi_0^{(BECCSD)}\rangle$, $|\Psi_0^{(QECCSD)}\rangle$, and $|\Psi_0^{(ECCSD)}\rangle$, respectively, Eq. (2.62), with the normalized full CI wave function $|\Psi_0^{(Full CI)}\rangle$, as functions of the N-N separation R (see Figure 2.3 (a)). As demonstrated in Figure 2.3 (a), the overlaps between the normalized BECCSD, QECCSD, or ECCSD wave functions and $|\Psi_0^{(Full CI)}\rangle$ vary between 0.98 and 1.0 in the entire R region. For comparison, the overlap of the normalized CCSD and full CI wave functions, which is close to 1.0 in the vicinity of the equilibrium geometry ($R \approx 2.0$ bohr), decreases to ~ 0.93 for larger N-N separations.



the normalized full CI wave function for the STO-3G N_2 molecule, as functions of the N-N separation R (in bohr). (b) The Figure 2.3: (a) The overlaps of the normalized CCSD, BECCSD, QECCSD, and ECCSD wave functions, Eq. (2.62), with difference between the CCSD and ECCSD cluster operators T and the difference between the ECCSD cluster operators T and Σ , as defined by the quantity d(Y, Z), Eq. (2.63), for the STO-3G N₂ molecule, as functions of the N-N separation R (in bohr).

The substantial differences between the T_1 and T_2 cluster components resulting from the standard CCSD calculations and their analogs obtained in the BECCSD, QECCSD, and ECCSD calculations at larger N-N distances that lead to the big differences between the behavior of the CCSD and BECCSD/QECCSD/ECCSD approaches in the bond breaking region are shown in Figure 2.3 (b). Suppose Y and Z are two excitation operators, defined by the amplitudes y_J and z_J , respectively (t_a^i and t_{ab}^{ij} for the cluster operator T and $\sigma_a^i = \sigma_i^a$ and $\sigma_{ab}^{ij} = \sigma_{ij}^{ab}$ for the cluster operator Σ). The following quantity provides us with an accurate measure of how different or how similar the two operators Y and Z are:

$$d(Y,Z) = \sqrt{\sum_{J} (y_{J} - z_{J})^{2}}.$$
 (2.63)

When Y = Z, we obtain d(Y, Z) = 0. When the cluster amplitudes y_J and z_J defining operators Y and Z, respectively, have similar values, d(Y, Z) is close to 0. Otherwise, the value of d(Y, Z) is significantly greater than 0. As shown in Figure 2.3 (b), the cluster operators T resulting from the CCSD and BECCSD/QECCSD/ECCSD calculations are almost identical when R is small or when R is close to the equilibrium bond length in N₂ ($R \approx 2.0$ bohr). This is why the CCSD and BECCSD/QECCSD/ECCSD wave functions and the corresponding energies (particularly, the expectation values of the Hamiltonian calculated with the CCSD and BECCSD/QECCSD/ECCSD wave functions) are virtually identical for the equilibrium and smaller values of R. The situation drastically changes when the triple bond in N₂ is stretched or broken. For R >3.0 bohr, the differences between the CCSD and BECCSD/QECCSD/ECCSD operators T increase so much (cf. Figure 2.3 (b)) that the behavior of the CCSD method and the behavior of the BECCSD/QECCSD/ECCSD approaches for stretched nuclear geometries are totally different. As we have already discussed, the standard CCSD method completely fails at larger R values, whereas the BECCSD/QECCSD/ECCSD/ECCSD 22 t approaches provide very good results. Figure 2.3 (b) illustrates another important feature of the ECCSD theory, namely, the similarity of the T and Σ operators in the equilibrium region and the significant difference between the T and Σ operators obtained in the ECCSD calculations in the region of larger N-N distances. As mentioned earlier, the lowest-order MBPT estimates of the operators Σ and T are identical (cf. Refs. 60-62, 83, 84). In consequence, in the equilibrium region, where the MBPT series rapidly converges, we have $\Sigma \approx T$ (see Figure 2.3 (b)). The situation changes, when the convergence of the MBPT series is slow or when the MBPT series diverges, as is the case when the N-N bond is stretched or broken. For larger N-N separations, the operators Σ and T become completely different. Figure 2.3 (b) provides us with a direct evidence that this is indeed what happens at larger R values.

2.3.2 The Arponen-Bishop ECCSD Approach

So far, we have tested the ECCSD methods based on the ECC formalism of Piecuch and Bartlett. We have demonstrated considerable improvements offered by the ECCSD approximations when multiple bonds are broken. The question is if similar improvements can be obtained when the alternative formulation of the ECC theory, proposed by Arponen and Bishop, is exploited. This question is addressed in this subsection.

The usefulness of the bi-variational ECCSD theory of Arponen and Bishop based on Eqs. (2.39) and (2.43)–(2.46) in improving the results for multiple bond breaking becomes apparent when we examine the results for the STO-3G model of N₂ shown in Table 2.4 and Figure 2.4. As one can see, the ECCSD method of Arponen and Bishop employing the ground-state RHF determinant as a reference provides remarkable improvements in the very poor description of the potential energy curve of N₂ by the standard CCSD method. The results are as good as those obtained with the ECCSD

approach of Piecuch and Bartlett. Indeed, the huge negative errors in the CCSD results at larger N-N separations, which exceed -200 millihartree in the R > 4.5 bohr region, reduce to much smaller positive errors (on the order of 31–38 millihartree when R > 4.5 bohr) in the ECCSD/Arponen-Bishop case. As shown in Figure 2.4, the ECCSD approach of Arponen and Bishop completely eliminates the pathological behavior of the standard CCSD method at larger N-N distances, restoring the variational description of the potential energy curve of N_2 at all internuclear separations. As in the case of the Piecuch-Bartlett theory, the ECCSD approach of Arponen and Bishop is capable of capturing the most essential nondynamic correlation effects (which the small STO-3G basis set used in these calculations already describes) in spite of the use of the single, spin- and symmetry-adapted, RHF determinant as a reference. The comparison of the CCSD and ECCSD potential energy curves for N_2 shown in Table 2.4 and Figure 2.4 confirms once again that the T_1 and T_2 cluster components obtained in the ECCSD calculations are of much higher quality in the bond breaking region than the T_1 and T_2 clusters resulting from the standard CCSD calculations.

As shown in Table 2.4 and Figure 2.4, the QECCSD/Arponen-Bishop results for the STO-3G model of the N₂ molecule, obtained by truncating the ECCSD energy functional at terms quadratic in $(\Sigma_1^{\dagger} + \Sigma_2^{\dagger})$ (see Eqs. (2.52)-(2.57)), are much better than the corresponding CCSD results and almost as good as the results of the full ECCSD calculations. The huge negative errors in the CCSD results in the R > 4.5bohr region, which exceed -200 millihartree, reduce to much smaller positive errors on the order of 35-46 millihartree, when the QECCSD method is employed. As in the case of the full ECCSD approach, the QECCSD approximation employing the Arponen-Bishop ECC formalism eliminates the pathological behavior of the CCSD method at larger N–N distances, restoring the variational description of the potential energy curve of N₂ at all internuclear separations. The variational and qualitatively correct behavior of the QECCSD method based on the Arponen-Bishop ECC theory is independent of the basis set. Indeed, as shown in Table 2.5 and Figure 2.5, the QECCSD potential energy curve for the double zeta (DZ)⁸⁵ model of N₂ is much better than the corresponding CCSD curve. The QECCSD potential energy curve for the N₂ molecule obtained with the DZ basis set, shown in Figure 2.5, is located above the full CI curve. The large negative errors in the CCSD results in the $R \ge 2R_e$ region of (-70) – (-121) millihartree ($R_e = 2.068$ bohr is the equilibrium value of R) are replaced by the considerably smaller positive errors of 40–50 millihartree when the QECCSD method is employed. As in the case of the STO-3G basis set, the QECCSD approach based on the Arponen-Bishop ECC theory eliminates the well pronounced hump on the CCSD potential energy curve, when the DZ basis set is employed.

Thus, the QECCSD approach exploiting the Arponen-Bishop ECC theory provides a practical method of capturing the large nondynamic correlation effects in N₂, in spite of the single-reference nature of the ECC formalism, in spite of the use of the RHF determinant in the QECCSD calculations, and, what is probably most remarkable, in spite of the two-body character of the QECCSD (or ECCSD) theory, which uses one- and two-body cluster T_1 , T_2 , Σ_1 , and Σ_2 only. As a matter of fact, the QECCSD results at larger internuclear separations of N₂ are much better than those obtained with the standard CC methods with singly, doubly, triply, and even quadruply excited clusters (cf. the QECCSD results in Table 2.5 and Figure 2.5 with the corresponding CCSD(T), CCSDT, CCSD(TQ_f), and CCSDT(Q_f) results).

R ^b	Full CI	CCSD	QECCSD ^c	ECCSD
1.5	-106.720117	1.102	0.886	0.885
2.0	-107.623240	3.295	1.898	1.897
2.5	-107.651880	9.220	3.443	3.427
3.0	-107.546614	13.176	3.909	3.757
3.5	-107.473442	-38.645	5.294	4.746
4.0	-107.447822	-140.376	15.815	14.148
4.5	-107.441504	-184.984	27.792	24.198
5.0	-107.439549	-200.857	35.335	30.590
5.5	-107.438665	-206.974	39.983	34.158
6 .0	-107.438265	-209.538	42.609	36 .082
7.0	-107.438054	-211.915	44.839	37.671
8.0	-107.438029	-213.431	45.508	38.161

Table 2.4.: The ground-state energies of the N_2 molecule obtained for several internuclear separations R with the STO-3G basis set.^a

^aThe full CI energies are in hartree. The CCSD, QECCSD, and ECCSD energies are in millihartree relative to the corresponding full CI energy values. The lowest two occupied orbitals were kept frozen.

^bThe N-N separation in bohr. The equilibrium value of R is 2.068 bohr.

^cThe QECCSD and full ECCSD methods are based on the Arponen-Bishop ECC theory.



Figure 2.4: Ground-state potential energy curves of the N_2 molecule, as described by the STO-3G basis sets. The QECCSD and ECCSD methods are based on the Arponen-Bishop ECC theory.

Method	0.75 <i>R</i> e	R_e^{b}	$1.25R_{e}$	1.5 <i>R</i> _e	1.75R _e	2R _e	$2.25R_e$
CCSD	3.132	8.289	19.061	33.545	17.714	-69.917	-120.836
CCSD1 CCSD(T) ^c	0.380	2.107 2.156	4.971	4.880	-51.869	-246.405	-387.448
$CCSD(TQ_f)^c$ $CCSDT(Q_f)^d$	0.226 0.047	0.323 -0.010	0.221 -0.715	-2.279 -4.584	-14.243 3.612	92.981 177.641	334.985 426.175
QECCSD ^e	2.506	6.236	13.609	23.485	31.060	40.085	49.741

Table 2.5.: The ground-state energies of the N_2 molecule obtained for several internuclear separations R with the DZ basis set.^a

^aAll energies are in millihartree relative to the corresponding full CI energy values, which are -108.549027, -109.105115, -109.054626, -108.950728, -108.889906, -108.868239, and -108.862125 hartree at $R = 0.75R_e$, R_e , $1.25R_e$, $1.5R_e$, $1.75R_e$, $2R_e$, and $2.25R_e$, respectively. The lowest two occupied and the highest two unoccupied orbitals were frozen in correlated calculations.

^bThe equilibrium value of R, $R_e = 2.068$ bohr.

^cFrom Ref. 31.

^dFrom Ref. 34.

^eThe quadratic approximation to the full ECCSD theory of Arponen and Bishop.



Figure 2.5: Ground-state potential energy curves of the N_2 molecule, as described by the DZ basis sets. QECCSD represents the quadratic version of the ECCSD theory of Arponen and Bishop.

2.4 Conclusion

We can summarize this chapter by stating that the ECCSD approach of Piecuch and Bartlett, and its approximate BECCSD and QECCSD variants⁷⁴⁻⁷⁶, and the ECCSD method of Arponen and Bishop and its QECCSD variant⁷⁷ represent interesting new alternatives for accurate electronic structure calculations of quasi-degenerate electronic states and bond breaking. The BECCSD, QECCSD, and full ECCSD methods remove the pervasive failing of the standard CCSD approach at larger internuclear separations and provide very good values of the T_1 and T_2 cluster amplitudes in the bond breaking region, in spite of using the RHF configuration as a reference and in spite of the two-body character of all ECCSD approximations. The BECCSD, QECCSD, and full ECCSD approaches improve the quality of the T_1 and T_2 cluster components so much that we can start thinking about using the BECCSD, QECCSD, or full ECCSD theories to design new single-reference *ab initio* methods for quantitative, high-accuracy calculations for bond breaking. For example, by having an access to very good T_1 and T_2 cluster amplitudes, resulting from the BECCSD, QECCSD, or ECCSD calculations, which are much better than the T_1 and T_2 clusters resulting from the standard CCSD calculations, we should be able to propose simple noniterative corrections to the BECCSD, QECCSD, or ECCSD energies, which may provide further improvements in the ECC results in the bond breaking region (see Chapter 3).

Based on a comparison of the ECCSD and QECCSD results for N_2 obtained with the Piecuch-Bartlett and Arponen-Bishop theories, we do not expect the differences between the ECC methods of Arponen and Bishop⁵⁰⁻⁵⁹ and Piecuch and Bartlett⁶⁰ to be large in the context of bond breaking. It seems to us that the rigorously bivariational character of the ECC formalism of Arponen and Bishop is of the secondary importance in the calculations of PESs involving bond breaking (cf. Section 2.3). Our experiences to date indicate that the most important factor that contributes to significant improvements in the quality of the T_1 and T_2 cluster components in the bond breaking region is the flexibility of the ECC theories, which rely on two independent sets of cluster amplitudes that are optimized by solving coupled systems of equations. The standard CC theory uses only one set of cluster amplitudes and this is not sufficient to obtain a correct description of multiple bond breaking by the standard CCSD method.

On the other hand, costs of the ECCSD calculations employing the ECC theory of Arponen and Bishop are somewhat smaller than those charactering the ECCSD approach of Piecuch and Bartlett. This, in particular, applies to the QECCSD approximation, which is simpler when the Arponen-Bishop ECC theory is employed. The QECCSD method based on Arponen's and Bishop's formulation of the ECC theory is an \mathcal{N}^6 procedure. The relatively low cost of the QECCSD approximation, the single reference ("black-box") character of the QECCSD calculations, and the qualitatively correct description of multiple bond breaking by the QECCSD approach make the QECCSD method an attractive theory for the design of noniterative corrections to CC energies. These corrections are discussed next.

Chapter 3 Noniterative Corrections to Extended Coupled-Cluster Energies for High Accuracy Electronic Structure Calculations: Generalized Method of Moments of Coupled-Cluster Equations

In Chapter 2, we showed that we can provide substantial improvements in the quality of the calculated potential energy curves and electronic quasi-degeneracies if we switch from the standard CCSD theory to its extended ECCSD counterpart. The question arises if we can improve the ECCSD (or QECCSD) results even further and obtain the quantitative description of bond breaking by adding the *a posteriori* corrections to ECCSD energies that would be reminiscent of the popular triples and quadruples corrections of the CCSD(T) and CCSD(TQ_f) methods and that would eliminate failures of these methods at larger internuclear separations. In this chapter, we show that such corrections to ECCSD energies can be developed if we use the generalized version of the method of moments of CC equations (the MMCC theory) of Piecuch and Kowalski^{11-13,30-32,74,75}. Before describing the generalized MMCC approach, we discuss the standard MMCC formalism.

3.1 The Standard Method of Moments of Coupled-Cluster Equations

The main idea of the standard MMCC theory $^{11-13,30-32,74,75}$ is that of the noniterative energy correction

$$\delta_0^{(A)} \equiv E_0 - E_0^{(A)}, \tag{3.1}$$

which, when added to the energy $E_0^{(A)}$ obtained in some standard approximate CC calculation A, such as CCSD, recovers the exact, full CI, energy E_0 . The objective of the approximate MMCC methods is to approximate corrections $\delta_0^{(A)}$, such that the resulting MMCC energies, defined as

$$E_0^{(\text{MMCC})} = E_0^{(A)} + \delta_0^{(A)}, \qquad (3.2)$$

are close to the corresponding full CI energies E_0 . The ground-state MMCC formalism can be extended to excited states^{12,13,86-88} and genuine multi-reference CC theories^{32,89,90}. In this thesis, we focus on the ground-state problem and merging the single-reference MMCC formalism with the non-standard CC theories, such as the ECC method of Arponen and Bishop⁵⁰⁻⁵⁹ (see Refs. 75 and 77).

In the standard formulation of the ground-state MMCC theory, we use the noniterative corrections $\delta_0^{(A)}$ to improve the results of the standard CC calculations. By the standard CC calculation, we mean any single-reference CC calculation in which the many-body expansion for the cluster operator T, defining the CC ground state $|\Psi_0\rangle$, Eq. (2.4), is truncated at some excitation level $m_A < N$ (recall that N is the number of electrons in a system). The general form of the truncated cluster operator $T^{(A)}$, defining the standard CC approximation A characterized by the excitation level m_A , is given by Eq. (2.20). An example of the standard CC approximation is the CCSD method. In this case, $m_A = 2$ and the cluster operator T is approximated by

$$T \approx T^{(\text{CCSD})} = T_1 + T_2, \tag{3.3}$$

where T_1 and T_2 are defined by Eqs. (2.25) and (2.26), respectively. Other examples of the standard CC approximations are the full CCSDT, CCSDTQ, and CCSDTQP approaches mentioned in the Introduction, in which $m_A = 3$, 4, and 5, respectively.

The standard CC system of equations for the cluster amplitudes $t_{a_1 \dots a_n}^{i_1 \dots i_n}$ defining the T_n components of $T^{(A)}$ has the following general form:

$$\langle \Phi_{i_1 \dots i_n}^{a_1 \dots a_n} | \bar{H}^{(A)} | \Phi \rangle = 0, \quad i_1 < \dots < i_n, \quad a_1 < \dots < a_n,$$
 (3.4)

where $n = 1, \ldots, m_A$,

$$\bar{H}^{(A)} = e^{-T^{(A)}} H e^{T^{(A)}} = (H e^{T^{(A)}})_C$$
(3.5)

is the similarity-transformed Hamiltonian of the CC theory, and $|\Phi_{i_1 \dots i_n}^{a_1 \dots a_n}\rangle = E_{i_1 \dots i_n}^{a_1 \dots a_n} |\Phi\rangle$ are the *n*-tuply excited determinants. In particular, the standard CCSD equations for the singly and doubly excited cluster amplitudes t_a^i and t_{ab}^{ij} defining operators T_1 and T_2 , respectively, are

$$\langle \Phi_i^a | \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \qquad (3.6)$$

$$\langle \Phi_{ij}^{ab} | \bar{H}^{(\text{CCSD})} | \Phi \rangle = 0, \quad i < j, \ a < b, \tag{3.7}$$

where $\bar{H}^{(\text{CCSD})}$ is the similarity-transformed Hamiltonian of the CCSD approach defined by Eq. (2.37). Once the system of equations, Eq. (3.4), is solved for $T^{(A)}$ (or Eqs. (3.6) and (3.7) are solved for T_1 and T_2), the CC energy corresponding to the standard approximation A is calculated as

$$E_0^{(A)} = \langle \Phi | \bar{H}^{(A)} | \Phi \rangle.$$
(3.8)

The fundamental formula of the ground-state MMCC formalism introduced in Refs. 11 and 30 (cf., also, Refs. 12, 30–32, 74, 75), which expresses the energy difference $\delta_0^{(A)}$, Eq. (3.1), in terms of the generalized moments of the single-reference CC equations of method A, has the following form:

$$\delta_{0}^{(A)} \equiv E_{0} - E_{0}^{(A)} = \sum_{n=m_{A}+1}^{N} \sum_{k=m_{A}+1}^{n} \langle \Psi_{0} | C_{n-k}(m_{A}) M_{k}(m_{A}) | \Phi \rangle / \langle \Psi_{0} | e^{T^{(A)}} | \Phi \rangle.$$
(3.9)

Here,

$$C_{n-k}(m_A) = (e^{T^{(A)}})_{n-k}$$
 (3.10)

is the (n-k)-body component of the wave operator $e^{T^{(A)}}$ defining the CC method A, $|\Psi_0\rangle$ is the full CI ground state, and

$$M_{k}(m_{A})|\Phi\rangle = \sum_{\substack{i_{1} < \cdots < i_{k} \\ a_{1} < \cdots < a_{k}}} \mathcal{M}_{a_{1} \cdots a_{k}}^{i_{1} \cdots i_{k}}(m_{A}) |\Phi_{i_{1} \cdots i_{k}}^{a_{1} \cdots a_{k}}\rangle.$$
(3.11)

The coefficients $\mathcal{M}_{a_1 \dots a_k}^{i_1 \dots i_k}(m_A)$ entering Eq. (3.11) represent the general moments of the CC equations of method A,

$$\mathcal{M}_{a_1\ldots a_k}^{i_1\ldots i_k}(m_A) = \langle \Phi_{i_1\ldots i_k}^{a_1\ldots a_k} | \bar{H}^{(A)} | \Phi \rangle.$$
(3.12)

By comparing Eqs. (3.12) and (3.4), we can state that moments $\mathcal{M}_{a_1}^{i_1\cdots i_k}(m_A)$ entering $\delta_0^{(A)}$ represent the projections of the CC equations of method A on excited determinants $|\Phi_{i_1\cdots i_k}^{a_1\cdots a_k}\rangle$ with $k > m_A$.

Equation (3.9) states that by calculating quantities $C_{n-k}(m_A)$, Eq. (3.10), and moments $\mathcal{M}_{a_1}^{i_1 \dots i_k}(m_A)$, Eq. (3.12), with $k > m_A$, we can determine the noniterative energy correction $\delta_0^{(A)}$ to the CC energy $E_0^{(A)}$ that we recovers the full CI energy E_0 . The determination of moments $\mathcal{M}_{a_1}^{i_1 \dots i_k}(m_A)$ for the low-order CC methods, such as CCSD, is relatively straightforward (cf. Eqs. (3.15)-(3.18) below). The $C_{n-k}(m_A)$ terms entering Eq. (3.9) are easy to calculate too. The zero-body term, $C_0(m_A)$, equals 1; the one-body term, $C_1(m_A)$, equals T_1 ; the two-body term, $C_2(m_A)$, equals $T_2 + \frac{1}{2}T_1^2$ if $m_A \ge 2$; the three-body term, $C_3(m_A)$, equals $T_1T_2 + \frac{1}{6}T_1^3$ if $m_A = 2$ and $T_3 + T_1T_2 + \frac{1}{6}T_1^3$ if $m_A \ge 3$, etc. Thus, the above formula for $\delta_0^{(A)}$, Eq. (3.9), is an excellent starting point for developing noniterative CC approaches, in which the ground-state energies are calculated by adding corrections $\delta_0^{(A)}$ to the standard CC energy $E_0^{(A)}$. For example, we can develop a hierarchy of approximations, in which the energy corrections based on Eq. (3.9) are added to the CCSD energies. In this case, the formula for the correction $\delta_0^{(CCSD)}$, which must be added to the CCSD energy $E_0^{(CCSD)}$ to recover the full CI energy E_0 , is

$$\delta_0^{(\text{CCSD})} \equiv E_0 - E_0^{(\text{CCSD})} = \sum_{n=3}^N \sum_{k=3}^{\min(n,6)} \langle \Psi_0 | C_{n-k}(2) M_k(2) | \Phi \rangle / \langle \Psi_0 | e^{T_1 + T_2} | \Phi \rangle, \quad (3.13)$$

where

$$M_{k}(2)|\Phi\rangle = \sum_{\substack{i_{1} < \cdots < i_{k} \\ a_{1} < \cdots < a_{k}}} \mathcal{M}_{a_{1} \cdots a_{k}}^{i_{1} \cdots i_{k}}(2) |\Phi_{i_{1} \cdots i_{k}}^{a_{1} \cdots a_{k}}\rangle, \qquad (3.14)$$

with moments $\mathcal{M}_{a_1 \dots a_k}^{i_1 \dots i_k}(2)$, k = 3-6, defined as projections of the CCSD equations on triply, quadruply, pentuply, and hextuply excited determinants,

$$\mathcal{M}_{abc}^{ijk}(2) = \langle \Phi_{ijk}^{abc} | [H_N(T_2 + T_1T_2 + \frac{1}{2}T_2^2 + \frac{1}{2}T_1^2T_2 + \frac{1}{2}T_1T_2^2 + \frac{1}{6}T_1^3T_2)]_C | \Phi \rangle, \quad (3.15)$$

$$\mathcal{M}_{abcd}^{ijkl}(2) = \langle \Phi_{ijkl}^{abcd} | [H_N(\frac{1}{2}T_2^2 + \frac{1}{2}T_1T_2^2 + \frac{1}{6}T_2^3 + \frac{1}{4}T_1^2T_2^2)]_C | \Phi \rangle, \qquad (3.16)$$

$$\mathcal{M}_{abcde}^{ijklm}(2) = \frac{1}{6} \langle \Phi_{ijklm}^{abcde} | [H_N(T_2^3 + T_1T_2^3)]_C | \Phi \rangle, \qquad (3.17)$$

$$\mathcal{M}_{abcdef}^{ijklmn}(2) = \frac{1}{24} \langle \Phi_{ijklmn}^{abcdef} | (H_N T_2^4)_C | \Phi \rangle.$$
(3.18)

The exact MMCC corrections $\delta_0^{(A)}$, Eq. (3.9), or $\delta_0^{(\text{CCSD})}$, Eq. (3.13), have the form of the complete many-body expansions involving all *n*-tuply excited configurations with $n = m_A + 1, \ldots, N$, where N is the number of electrons in a system (cf. the summations over n in Eqs. (3.9) and (3.13)). Thus, in order to develop practical methods based on the MMCC theory, we must first truncate the many-body expansions for corrections $\delta_0^{(A)}$ or $\delta_0^{(\text{CCSD})}$ at some, preferably low, excitation level m_B satisfying $m_A < m_B < N$. This leads to the MMCC(m_A, m_B) schemes, in which we calculate the energy as follows^{11-13,30-32,74,75}:

$$E_0^{(\text{MMCC})}(m_A, m_B) = E_0^{(A)} + \delta_0(m_A, m_B), \qquad (3.19)$$

where $E_0^{(A)}$ is the energy obtained with the CC method A and

$$\delta_0(m_A, m_B) = \sum_{n=m_A+1}^{m_B} \sum_{k=m_A+1}^n \langle \Psi_0 | C_{n-k}(m_A) M_k(m_A) | \Phi \rangle / \langle \Psi_0 | e^{T^{(A)}} | \Phi \rangle$$
(3.20)

is the relevant MMCC correction. Examples of the $MMCC(m_A, m_B)$ schemes are the basic MMCC(2,3) and MMCC(2,4) approximations, in which energies are calculated as follows^{11-13,30-32,74,75}:

$$E_0^{(\text{MMCC})}(2,3) = E_0^{(\text{CCSD})} + \langle \Psi_0 | M_3(2) | \Phi \rangle / \langle \Psi_0 | e^{T_1 + T_2} | \Phi \rangle, \qquad (3.21)$$

$$E_0^{(\text{MMCC})}(2,4) = E_0^{(\text{CCSD})} + \langle \Psi_0 | \{ M_3(2) + [M_4(2) + T_1 M_3(2)] \} | \Phi \rangle / \langle \Psi_0 | e^{T_1 + T_2} | \Phi \rangle ,$$
(3.22)

where $E_0^{(\text{CCSD})}$ is the CCSD ground-state energy. The MMCC(2,3) approach requires that we only determine the triply excited moments $\mathcal{M}_{abc}^{ijk}(2)$, Eq. (3.15). The MMCC(2,4) method requires that we determine the triply excited moments $\mathcal{M}_{abc}^{ijk}(2)$, Eq. (3.15), and the quadruply excited moments $\mathcal{M}_{abcd}^{ijkl}(2)$, Eq. (3.16). The second issue that has to be addressed before the MMCC(m_A, m_B) methods can be used in practice is the fact that in the exact MMCC theory the wave function $|\Psi_0\rangle$ that enters Eqs. (3.9) and (3.13) is the exact, full CI, ground state. Thus, in order to propose the computationally tractable approaches based on the MMCC theory, we must approximate $|\Psi_0\rangle$ in some, preferably inexpensive, manner. Various ways of approximating $|\Psi_0\rangle$ in Eqs. (3.20)-(3.22), leading to the completely renormalized (CR) CC methods^{11-13,30-35,37,74,75}, the quasi-variational and quadratic MMCC methods^{74-76,91}, the CI-corrected MMCC approaches^{11-13,75,86,87,92}, and the multireference MBPT-corrected MMCC approach⁹³ – all employing the CCSD values of T_1 and T_2 to construct the relevant corrections $\delta_0^{(CCSD)}$ – have been suggested. For example, the CR-CCSD(T) method^{11-13,30,31,33,35,37,74,75,94-96} is obtained by replacing the wave function $|\Psi_0\rangle$ in the MMCC(2,3) formula, Eq. (3.21), by the MBPT(2)[SDT]like expression

$$|\Psi_0^{\text{CCSD}(T)}\rangle = (1 + T_1 + T_2 + T_3^{[2]} + Z_3)|\Phi\rangle,$$
 (3.23)

where T_1 and T_2 are the singly and doubly excited clusters obtained in the CCSD calculations, $T_3^{[2]}$ is an approximation of the connected triply excited clusters T_3 , defined by Eq. (2.61), and

$$Z_3|\Phi\rangle = R_0^{(3)} V_N T_1 |\Phi\rangle \tag{3.24}$$

is the disconnected triples correction, which is responsible for the difference between the [T] and (T) triples corrections of the standard CCSD[T] and CCSD(T) approaches. The higher-order CR-CCSD(TQ) methods^{11-13,30,31,33-35,74,75} are obtained in a similar manner, by inserting the MBPT(2)[SDTQ]-like expressions for $|\Psi_0\rangle$ into the MMCC(2,4) formula, Eq. (3.22). For example, the wave function $|\Psi_0\rangle$ defining variant "b" of the CR-CCSD(TQ) approach (the CR-CCSD(TQ),b method) is defined as follows:

$$|\Psi_0^{\text{CCSD(TQ)},b}\rangle = |\Psi_0^{\text{CCSD(T)}}\rangle + \frac{1}{2}T_2^2|\Phi\rangle, \qquad (3.25)$$

where $|\Psi_0^{\text{CCSD(T)}}\rangle$ is given by Eq. (3.23).

As shown in Refs. 11–13, 30–35, 37, 74, 75, and 94–96, the CR-CCSD(T) and CR-CCSD(TQ) approaches eliminate or considerably reduce the failures of the standard CCSD(T) and CCSD(TQ_f) methods at larger internuclear separations and for

diradicals without making the calculations substantially more difficult or expensive. In particular, the CR-CCSD(T) and CR-CCSD(TQ) methods provide a very good description of single bond breaking. Unfortunately, performance of these methods for multiply bonded systems is often much less impressive, partly due to the very poor quality of the singly and doubly excited cluster amplitudes resulting from the CCSD calculations, on which the CR-CCSD(T) and CR-CCSD(TQ) methods are based, in calculations involving multiple bond breaking. The purpose of this work is to examine an alternative approach, in which the MMCC corrections $\delta_0^{(A)}$ or $\delta_0^{(CCSD)}$ are constructed using the cluster components obtained in the ECC or ECCSD calculations. As shown in Chapter 2, the ECCSD method provides much better values of T_1 and T_2 clusters than the CCSD approach when multiple bonds are stretched or broken, so that we may be able to improve the CR-CCSD(T) and CR-CCSD(TQ)results in cases where the standard CCSD values of T_1 and T_2 are of very poor quality. The use of the ECC or ECCSD values of cluster amplitudes in MMCC calculations requires the generalization of the MMCC theory to non-standard CC methods, which is described in the next section.

3.2 The Generalized MMCC Formalism

Interestingly, Eqs. (3.9) and (3.13) can be generalized to a situation where the truncated cluster operator $T^{(A)}$ is not determined by solving the standard CC equations, Eq. (3.4). Here is how this works: When the cluster operator $T^{(A)}$ is obtained in a non-standard way, we can no longer assume that the generalized moments $\mathcal{M}_{a_1 \dots a_k}^{i_1 \dots i_k}(m_A)$ with $k = 1, \dots, m_A$ vanish; they only vanish in the standard CC case (cf. Eqs. (3.4) and (3.12)). It has been shown in Refs. 13 and 30 that when moments $\mathcal{M}_{a_1\cdots a_k}^{i_1\cdots i_k}(m_A)$ with $k=1,\ldots m_A$ do not vanish, we have to use the following expression for the exact, full CI energy E_0 instead of Eq. (3.9):

$$E_{0} = \sum_{n=0}^{N} \sum_{k=0}^{n} \langle \Psi_{0} | C_{n-k}(m_{A}) M_{k}(m_{A}) | \Phi \rangle / \langle \Psi_{0} | e^{T^{(A)}} | \Phi \rangle$$

$$= M_{0}(m_{A}) + \sum_{n=1}^{N} \sum_{k=1}^{n} \langle \Psi_{0} | C_{n-k}(m_{A}) M_{k}(m_{A}) | \Phi \rangle / \langle \Psi_{0} | e^{T^{(A)}} | \Phi \rangle.$$
(3.26)

Here, $M_0(m_A)$ designates the zero-body moment, which is calculated in exactly the same way as the CC energy $E_0^{(A)}$, i.e. (cf. Eq. (3.8))

$$M_0(m_A) = \langle \Phi | \bar{H}^{(A)} | \Phi \rangle. \tag{3.27}$$

Although the formulas for $M_0(m_A)$ and $E_0^{(A)}$ are identical, there is a fundamental difference between $M_0(m_A)$ and $E_0^{(A)}$. The energy $E_0^{(A)}$ is determined using the cluster amplitudes originating from the standard CC equations, Eq. (3.4), and $E_0^{(A)}$ is the energy expression used in the CC theory. The zero-body moment $M_0(m_A)$ can be computed with any cluster operator $T^{(A)}$, obtained, for example, by performing some nonstandard calculations, such as the ECC calculations, and $M_0(m_A)$ does not have to represent the energy expression of the nonstandard CC theory used to generate $T^{(A)}$ (e.g., $M_0(m_A)$ does not represent the energy expression of the ECC formalism).

As in the case of the standard MMCC formalism defined by Eq. (3.9), the $C_{n-k}(m_A)$ and $M_k(m_A)|\Phi\rangle$ quantities entering Eq. (3.26) are defined by Eqs. (3.10) and (3.11), respectively, although we must consider now all quantities $M_k(m_A)|\Phi\rangle$ with $k \geq 1$, not just those with $k > m_A$. In other words, since we no longer assume that the generalized moments $\mathcal{M}_{a_1}^{i_1 \dots i_k}(m_A)$ with $k = 1, \dots, m_A$ vanish, since Eq. (3.4) is no longer satisfied, we must consider all generalized moments $\mathcal{M}_{a_1}^{i_1 \dots i_k}(m_A)$ with a given cluster operator $T^{(A)}$ produces. In particular, if we want to use the T_1 and T_2 clusters to construct the full CI energy E_0 , which are no longer determined

by solving the standard CCSD equations, we must use the following formula for the exact energy E_0 :

$$E_0 = M_0(2) + \sum_{n=1}^{N} \sum_{k=1}^{\min(n,6)} \langle \Psi_0 | C_{n-k}(2) M_k(2) | \Phi \rangle / \langle \Psi_0 | e^{T_1 + T_2} | \Phi \rangle, \qquad (3.28)$$

where $M_k(2)|\Phi\rangle$ is defined by Eq. (3.14). As one can see, Eq. (3.28) is very similar to Eq. (3.13). In particular, we do not have to consider moments $\mathcal{M}_{a_1}^{i_1 \dots i_k}(2)$ with k > 6, since for Hamiltonians containing up to two-body interactions the generalized moments $\mathcal{M}_{a_1}^{i_1 \dots i_k}(2)$ with k > 6 vanish, independent of the source of T_1 and T_2 clusters. There is, however, a difference between Eqs. (3.13) and (3.28): in Eq. (3.28) we consider the singly and doubly excited moments, $\mathcal{M}_a^i(2) = \langle \Phi_i^a|\bar{H}^{(\text{CCSD})}|\Phi\rangle$ and $\mathcal{M}_{ab}^{ij}(2) = \langle \Phi_{ij}^{ab}|\bar{H}^{(\text{CCSD})}|\Phi\rangle$, respectively, which are no longer zeroed, along with moments $\mathcal{M}_{a_1 \dots a_k}^{i_1 \dots i_k}(2)$ with k = 3 - 6 considered in the standard CCSD case; Eq. (3.13) uses moments $\mathcal{M}_{a_1 \dots a_k}^{i_1 \dots i_k}(2)$ with k = 3 - 6 only, since $\mathcal{M}_a^i(2) = \mathcal{M}_{ab}^{ij}(2) = 0$ in the standard CCSD case. The formulas for the $\mathcal{M}_a^i(2)$ and $\mathcal{M}_{ab}^{ij}(2)$ moments, entering Eq. (3.28), in terms of the T_1 and T_2 clusters, are identical to the left-hand sides of the standard CCSD equations. Thus, we obtain (cf. Eqs. (3.6) and (3.7)),

$$\mathcal{M}_{a}^{i}(2) = \langle \Phi_{i}^{a} | [H_{N}(1+T_{1}+T_{2}+\frac{1}{2}T_{1}^{2}+T_{1}T_{2}+\frac{1}{6}T_{1}^{3})]_{C} | \Phi \rangle, \qquad (3.29)$$

$$\mathcal{M}_{ab}^{ij}(2) = \langle \Phi_{ij}^{ab} | [H_N(1+T_1+T_2+\frac{1}{2}T_1^2+T_1T_2+\frac{1}{6}T_1^3+\frac{1}{2}T_2^2+\frac{1}{2}T_1^2T_2+\frac{1}{24}T_1^4)]_C | \Phi \rangle.$$
(3.30)

Equations (3.26) and (3.28) define the generalized version of the ground-state MMCC theory, designated as GMMCC. Clearly, the GMMCC theory reduces to the standard MMCC formalism if cluster components T_n defining operator $T^{(A)}$ are determined by solving the standard CC equations, Eq. (3.4). In this case, the $\mathcal{M}_{a_1 \dots a_k}^{i_1 \dots i_k}(m_A)$ moments with $k = 1, \dots, m_A$ vanish and the summations $\sum_{n=1}^N \sum_{k=1}^n$ entering Eq. (3.26) reduce to $\sum_{n=m_A+1}^N \sum_{k=m_A+1}^n$, giving Eq. (3.9). The obvious advantage of Eqs. (3.26) and (3.28) is that they are much more general than Eqs. (3.9) and (3.13), enabling us to use the non-standard values of $T^{(A)}$ clusters. Otherwise, we use Eqs. (3.26) and (3.28) in exactly the same way as Eqs. (3.9) and (3.13). Thus, once the cluster components defining the truncated cluster operator $T^{(A)}$ are determined, we calculate the relevant moments $\mathcal{M}_0(m_A)$ and $\mathcal{M}_{a_1\cdots a_k}^{i_1\cdots i_k}(m_A)$ and use these moments to determine the ground-state energy E_0 .

In analogy to the standard MMCC theory, a few issues have to be addressed before using the GMMCC formalism in practical calculations. First of all, the exact GMMCC expressions, Eqs. (3.26) or (3.28), represent the complete many-body expansions involving all *n*-tuply excited configurations with n = 1, ..., N, where Nis the number of electrons in a system (see the summations over n in Eqs. (3.26) and (3.28)). Thus, in order to develop the computationally tractable GMMCC methods, we must first truncate the many-body energy expansions, Eqs. (3.26) or (3.28), at some, preferably low, excitation level m_B , where $m_A < m_B < N$. This leads to the GMMCC(m_A, m_B) schemes, which are the non-standard analogs of the MMCC(m_A, m_B) approximations discussed in Section 3.1. By limiting ourselves to the wave functions $|\Psi_0\rangle$ that do not contain higher-than- m_B -tuply excited components relative to reference $|\Phi\rangle$ and by restricting the summation over n in Eq. (3.26) accordingly, we obtain the following energy expression for the GMMCC(m_A, m_B) methods:

$$E_0^{(\text{GMMCC})}(m_A, m_B) = M_0(m_A) + \sum_{n=1}^{m_B} \sum_{k=1}^n \langle \Psi_0 | C_{n-k}(m_A) M_k(m_A) | \Phi \rangle / \langle \Psi_0 | e^{T^{(A)}} | \Phi \rangle,$$
(3.31)

where $M_0(m_A)$ is defined by Eq. (3.27). The exact GMMCC formalism, equivalent to calculating the full CI energies, is obtained when $|\Psi_0\rangle$ in Eq. (3.31) is the full CI ground-state wave function and $m_B = N$. In this case, it is irrelevant what is the value of m_A and where is the cluster operator $T^{(A)}$ taken from. When $|\Psi_0\rangle$ is exact and $m_B = N$, Eq. (3.31) produces the exact energy, independent of the excitation level m_A and the source of cluster amplitudes defining $T^{(A)}$.

In this thesis, we focus on the GMMCC(m_A, m_B) schemes with $m_A = 2$, which can be used to correct the results of the non-standard CCSD-like (e.g., ECCSD or QECCSD) calculations. In analogy to the standard MMCC theory, two GMMCC(m_A, m_B) approximations are expected to be particularly useful: the GMMCC(2,3) method and the GMMCC(2,4) approach. According to Eq. (3.31), the GMMCC(2,3) and GMMCC(2,4) energies are calculated as follows:

$$E_0^{(\text{GMMCC})}(2,3) = M_0(2) + \langle \Psi_0 | \{ M_1(2) + [M_2(2) + T_1 M_1(2)] + [M_3(2) + T_1 M_2(2) + (T_2 + \frac{1}{2}T_1^2)M_1(2)] \} | \Phi \rangle / \langle \Psi_0 | e^{T_1 + T_2} | \Phi \rangle , \quad (3.32)$$

$$E_{0}^{(\text{GMMCC})}(2,4) = M_{0}(2) + \langle \Psi_{0} | \{ M_{1}(2) + [M_{2}(2) + T_{1}M_{1}(2)] + [M_{3}(2) + T_{1}M_{2}(2) + (T_{2} + \frac{1}{2}T_{1}^{2})M_{1}(2)] + [M_{4}(2) + T_{1}M_{3}(2) + (T_{2} + \frac{1}{2}T_{1}^{2})M_{2}(2) + (T_{1}T_{2} + \frac{1}{6}T_{1}^{3})M_{1}(2)] \} | \Phi \rangle / \langle \Psi_{0} | e^{T_{1}+T_{2}} | \Phi \rangle .$$
(3.33)

As one can see, in the case of the GMMCC(2,3) approximation, we must calculate moments $\mathcal{M}_{a}^{i}(2)$, $\mathcal{M}_{ab}^{ij}(2)$, and $\mathcal{M}_{abc}^{ijk}(2)$, using Eqs. (3.29), (3.30), and (3.15), respectively, in addition to moment $\mathcal{M}_{0}(2)$, Eq. (3.27). In the case of the GMMCC(2,4) approach, we must determine moments $\mathcal{M}_{abc}^{ijk}(2)$ and $\mathcal{M}_{abcd}^{ijkl}(2)$, Eqs. (3.15) and (3.16), respectively, along with moments $\mathcal{M}_{0}(2)$, $\mathcal{M}_{a}^{i}(2)$ and $\mathcal{M}_{ab}^{ij}(2)$, Eqs. (3.27), (3.29) and (3.30), respectively. Clearly, when the singly and doubly excited moments $\mathcal{M}_{a}^{i}(2)$ and $\mathcal{M}_{ab}^{ij}(2)$ vanish, the GMMCC(2,3) and GMMCC(2,4) energy expressions, Eqs. (3.32) and (3.33), respectively, reduce to the MMCC(2,3) and MMCC(2,4) formulas given by Eqs. (3.21) and (3.22). This can only happen when the T_1 and T_2 clusters are obtained by solving the standard CCSD equations, Eqs. (3.6) and (3.7).

The second issue that needs to be addressed before the GMMCC(2,3) and GMMCC(2,4) methods and other GMMCC(m_A, m_B) approximations are used in practice is the issue of the wave function $|\Psi_0\rangle$ that enters Eqs. (3.31)–(3.33), which in the exact theory is a full CI ground state. Clearly, in order to make the GMMCC(2,3), GMMCC(2,4), and other GMMCC(m_A, m_B) schemes usable in practical applications, we must suggest some approximate forms of $|\Psi_0\rangle$ that can be easily generated with one of the inexpensive ab initio approaches. Finally, the third issue is the source of the T_1 and T_2 cluster amplitudes which are needed to construct the GMMCC(2,3) and GMMCC(2,4) energy expressions, Eqs. (3.32) and (3.33), respectively. As mentioned earlier, in calculating the GMMCC(2,3) and GMMCC(2,4) energies we would like to use the T_1 and T_2 cluster components which are more accurate in cases involving multiple bond breaking than those obtained in the standard CCSD calculations. As shown in Chapter 2, the T_1 and T_2 cluster components resulting from various types of ECCSD calculations are much better than their standard CCSD counterparts, when multiply bonded systems are examined. Thus, it is worth examining the possibility of combining the GMMCC(2,3) and GMMCC(2,4) schemes with the ECCSD methods. The resulting ECCSD(T), ECCSD(TQ), QECCSD(T) and QECCSD(TQ)approaches are discussed next.

3.3 The ECCSD(T), ECCSD(TQ), QECCSD(T), and QECCSD(TQ) Methods and their Performance in Calculations for Triple Bond Breaking in N_2

The ECCSD results for triple bond breaking in N₂ are so much better than their standard CCSD analogs that it is very important to analyze the effect of replacing the CCSD values of the T_1 and T_2 cluster components in the MMCC calculations by the ECCSD values of these clusters. Since the ECCSD and QECCSD methods are no longer the standard CC theories, so that we can no longer assume that the singly and doubly excited moments, $\mathcal{M}_a^i(2)$ and $\mathcal{M}_{ab}^{ij}(2)$, Eqs. (3.29) and (3.30), respectively, vanish, we must use the GMMCC formalism discussed in Section 3.2 rather than the standard MMCC formalism discussed in Section 3.1 in such considerations. In the following, we test an idea of using the QECCSD and ECCSD values of T_1 and T_2 , resulting from the application of the Arponen-Bishop variant of the ECC theory, in the GMMCC calculations.

As explained in Section 3.2, in practice we are interested in a truncated form of the GMMCC theory that leads to relatively low costs of calculating the final energy. All of our tests to date, including the calculations for N₂ discussed below, indicate that the lowest-order GMMCC scheme, employing the ECCSD or QECCSD values of T_1 and T_2 , which provides substantial improvements in the results for multiple bond breaking, is the GMMCC(2,4) approach defined by Eq. (3.33). In this approach, we only consider the generalized moments $\mathcal{M}_{a_1\cdots a_k}^{i_1\cdots i_k}(2)$ with k = 1 - 4, i.e. moments corresponding to the projections of $\bar{H}^{(\text{CCSD})}|\Phi\rangle$ on singly, doubly, triply, and quadruply excited determinants. The lower-order GMMCC(2,3) approach, defined by Eq. (3.32), employing the ECCSD or QECCSD values of T_1 and T_2 and ignoring the quadruply excited moments $\mathcal{M}_{abcd}^{ijkl}(2)$, provides improvements too, but the neglect of the quadruply excited moments $\mathcal{M}_{abcd}^{ijkl}(2)$ in the GMMCC calculations has a negative impact on the results for multiply bonded systems, such as N₂.

As in all approximate MMCC calculations, we must decide what to do with the wave function $|\Psi_0\rangle$ that enters the GMMCC(2,3) and GMMCC(2,4) energy formulas, Eqs. (3.32) and (3.33). Since we are mainly interested in the "black-box" GMMCC approaches of the CCSD(T) or CCSD(TQ) type, in this study of the performance of the GMMCC theory employing the ECCSD and QECCSD values of T_1 and T_2 , we use the same types of the wave functions $|\Psi_0\rangle$ in Eqs. (3.32) and (3.33) as the wave functions $|\Psi_0^{\text{CCSD}(\text{T})}\rangle$ and $|\Psi_0^{\text{CCSD}(\text{TQ}),b}\rangle$ used in the CR-CCSD(T) and CR-CCSD(TQ),b calculations (cf. Eqs. (3.23) and (3.25)). Depending on the choice of $|\Psi_0\rangle$ and the precise source of T_1 and T_2 clusters for the GMMCC(2,3) and GMMCC(2,4) calculations, we introduce the following four approximations:

- ECCSD(T): The ECCSD(T) approach is defined as the GMMCC(2,3) method, in which $|\Psi_0\rangle$ is defined by Eq. (3.23) and in which T_1 and T_2 clusters originate from the full ECCSD calculations of the Arponen-Bishop type.
- QECCSD(T): The QECCSD(T) approach is defined as the GMMCC(2,3) method, in which |Ψ₀⟩ is defined by Eq. (3.23) and in which T₁ and T₂ clusters originate from the QECCSD calculations of the Arponen-Bishop type.
- ECCSD(TQ): The ECCSD(TQ) approach is defined as the GMMCC(2,4) method, in which |Ψ₀⟩ is defined by Eq. (3.25) and in which T₁ and T₂ clusters originate from the full ECCSD calculations of the Arponen-Bishop type.
- QECCSD(TQ): The QECCSD(TQ) approach is defined as the GMMCC(2,4) method, in which $|\Psi_0\rangle$ is defined by Eq. (3.25) and in which T_1 and T_2 clusters originate from the QECCSD calculations of the Arponen-Bishop type.

Based on the above definitions, it is easy to verify that the ECCSD(T)/QECCSD(T) energies can be calculated as follows:

$$E_0^{((Q) \text{ECCSD(T)})} = M_0(2) + N^{(T)}/D^{(T)}, \qquad (3.34)$$

where

$$N^{(\mathrm{T})} = \langle \Phi | T_1^{\dagger} M_1(2) | \Phi \rangle + \langle \Phi | T_2^{\dagger} [M_2(2) + T_1 M_1(2)] | \Phi \rangle + \langle \Phi | (T_3^{[2]} + Z_3)^{\dagger} [M_3(2) + T_1 M_2(2) + (T_2 + \frac{1}{2} T_1^2) M_1(2)] | \Phi \rangle$$
(3.35)

and

$$D^{(\mathrm{T})} = 1 + \langle \Phi | T_1^{\dagger} T_1 | \Phi \rangle + \langle \Phi | T_2^{\dagger} \left(T_2 + \frac{1}{2} T_1^2 \right) | \Phi \rangle + \langle \Phi | (T_3^{[2]} + Z_3)^{\dagger} (T_1 T_2 + \frac{1}{6} T_1^3) | \Phi \rangle, \quad (3.36)$$

with $T_3^{[2]}$ and Z_3 defined by Eqs. (2.61) and (3.24), respectively, and T_1 and T_2 obtained in the ECCSD/QECCSD calculations. The ECCSD(TQ)/QECCSD(TQ) energies are calculated as

$$E_0^{((Q) \in CCSD(TQ))} = M_0(2) + N^{(TQ)} / D^{(TQ)}, \qquad (3.37)$$

where

$$N^{(\mathrm{TQ})} = N^{(T)} + \frac{1}{2} \langle \Phi | (T_2^{\dagger})^2 [M_4(2) + T_1 M_3(2) + (T_2 + \frac{1}{2} T_1^2) M_2(2) + (T_1 T_2 + \frac{1}{6} T_1^3) M_1(2)] | \Phi \rangle$$
(3.38)

and

$$D^{(\mathrm{TQ})} = D^{(\mathrm{T})} + \frac{1}{2} \langle \Phi | (T_2^{\dagger})^2 (\frac{1}{2}T_2^2 + \frac{1}{2}T_1^2 T_2 + \frac{1}{24}T_1^4) | \Phi \rangle, \qquad (3.39)$$

with $N^{(T)}$ and $D^{(T)}$ defined by Eqs. (3.35) and (3.36), respectively, and T_1 and T_2 obtained in the ECCSD/QECCSD calculations.

The ECCSD(T)/QECCSD(T) methods reduce to the CR-CCSD(T) approach of Refs. 11 and 30 (cf. Section 3.1) if the T_1 and T_2 clusters originating from the

ECCSD/QECCSD calculations are replaced in Eqs. (3.34)-(3.36) by their standard CCSD values. Similarly, the ECCSD(TQ)/QECCSD(TQ) methods reduce to the CR-CCSD(T),b approach of Ref. 31 (cf., also, Section 3.1) if the T_1 and T_2 clusters originating from the ECCSD/QECCSD calculations are replaced in Eqs. (3.37)-(3.39) by their CCSD analogs. These straightforward relationships between the ECCSD(T)/QECCSD(T) and ECCSD(TQ)/QECCSD(TQ) methods on the one hand and the CR-CCSD(T) and CR-CCSD(T),b approaches on the other hand immediately imply that once the T_1 and T_2 clusters are determined by solving the ECCSD/QECCSD equations, the costs of calculating the ECCSD(T)/QECCSD(T) and

ECCSD(TQ)/QECCSD(TQ) energies are essentially the same as the costs of the corresponding CR-CCSD(T) and CR-CCSD(T), b calculations or their standard CCSD(T) and CCSD(TQ_f) counterparts. In particular, the most expensive steps of the ECCSD/QECCSD-based ECCSD(T)/QECCSD(T) calculations (if we ignore the costs of the ECCSD/QECCSD calculations) scale as $n_o^3 n_u^4$. The most expensive steps of the ECCSD(TQ)/QECCSD(TQ) calculations (again, ignoring the costs of the ECCSD/QECCSD calculations) scale as either $n_o^2 n_u^5$ or n_u^6 . The ECCSD(T) and ECCSD/QECCSD calculations) scale as either $n_o^2 n_u^5$ or n_u^6 . The ECCSD(T) and ECCSD(TQ) methods are much less practical, since the underlying ECCSD calculations that provide T_1 and T_2 clusters have steps that scale as \mathcal{N}^{10} with the system size. However, the QECCSD(T) and QECCSD(TQ) methods are very promising in this regard, since both the underlying QECCSD calculations that provide T_1 and T_2 clusters and the calculations of the final QECCSD(T) and QECCSD(TQ) energies, Eqs. (3.34) and (3.37), respectively, have steps that scale, at worst, as $n_o^2 n_u^6$ or n_u^6 (\mathcal{N}^7 or \mathcal{N}^6 with the system size). We must keep in mind, however, that the QECCSD(T) and QECCSD(TQ) methods are approximations to the more complete ECCSD(T) and ECCSD(TQ) approaches. The questions, therefore, are:

- (i) Do the QECCSD(T) and QECCSD(TQ) methods provide the results of the full ECCSD(T) and ECCSD(TQ) quality?
- (ii) Are the QECCSD(T) and QECCSD(TQ) methods sufficiently accurate to eliminate the problems observed in the standard and completely renormalized CCSD(T) and CCSD(TQ) calculations for triple bond breaking in N₂?

The answers to both questions can be provided if we examine the results of the benchmark ECCSD(T), ECCSD(TQ), QECCSD(T), and QECCSD(TQ) calculations for N₂ shown in Tables 3.1, 3.2 and Figures 3.1, 3.2, which we performed with the computer codes developed in this thesis work⁷⁷.

As shown in Table 3.1 and Figure 3.1, the ECCSD(TQ) approach employing the T_1 and T_2 clusters obtained in the full ECCSD (Arponen-Bishop) calculations is capable of providing spectacular improvements in the description of triple bond breaking in the N₂ molecule, as described by the STO-3G basis set, reducing the large unsigned errors in the CCSD and CR-CCSD(TQ),b results in the $R \ge 5.0$ bohr region, on the order of 201–213 and 39–54 millihartree, respectively, and the 31–38 millihartree errors in the ECCSD results to less than 4 millihartree. Remarkably enough, the ECCSD(TQ) results for the N₂ molecule, as described by the STO-3G basis set, in the entire $R \le 8.0$ bohr ($\approx 4R_e$) region do not exceed ~ 4 millihartree, being much smaller in the $R \approx R_e$ region. As shown in Table 3.1 and Figure 3.1, the ECCSD(TQ) potential energy curve is located only slightly above the full CI curve and there is no unphysical hump on it. Interestingly enough, the zero-body moment $M_0(2)$, Eq. (3.27), calculated with the ECCSD values of T_1 and T_2 , which is corrected in the ECCSD(TQ) energy formula, Eq. (3.37), by adding terms expressed via moments $\mathcal{M}_{a_1 \cdots a_k}^{i_1 \cdots i_k}(2)$ with k = 1 - 4, is a poor approximation to the exact, full CI, energy in the region of large N-N separations (see Table 3.1). This demonstrates the remarkable ability of the MMCC (or GMMCC) formalism to restore high accuracies in the bond breaking region even when the CC energy that we are trying to correct (in this case, the $\mathcal{M}_0(2)$ /ECCSD energy) is itself very poor.

The ECCSD(TQ) results for the STO-3G model of N_2 shown in Table 3.1 and Figure 3.1 are very encouraging, but, as mentioned earlier, the ECCSD(TQ) approach is not too practical due to the expensive \mathcal{N}^{10} steps of the underlying ECCSD calculations. It is, therefore, important to examine if the much more manageable QECCSD(TQ) approximation, which relies on the T_1 and T_2 clusters obtained in the QECCSD calculations and which is the $\mathcal{N}^6 - \mathcal{N}^7$ procedure, provides the results of the ECCSD(TQ) quality. The results in Table 3.1 indicate that the QECCSD(TQ) and ECCSD(TQ) energies are virtually indistinguishable when the internuclear separation R does not exceed 3.5 bohr ($\approx 1.75R_e$). Only when the N-N separations exceed 3.5 bohr, the differences between the QECCSD(TQ) and ECCSD(TQ) results become larger. Although the errors in the QECCSD(TQ) results in the R > 3.5 bohr are greater than the corresponding errors obtained with the ECCSD(TQ) method, the QECCSD(TQ) method provides an excellent description of the potential energy curve of the STO-3G N_2 molecule, reducing the 35–46 millihartree errors in the QECCSD results and much larger errors in the CCSD and CR-CCSD(TQ),b results in the R > 4.5 bohr region to 5–6 millihartree (see Table 3.1). As in the ECCSD(TQ) case, the QECCSD(TQ) potential energy curve of the STO-3G N₂ molecule is located very close to and above the exact, full CI curve (see Figure 3.1). Thus, the QECCSD(TQ) approach provides a highly accurate description of the large nondynamic correlation effects characterizing the N₂ molecule at larger N-N separations.
The question is if the above observations obtained for the very small STO-3G basis apply to larger basis sets. In order to answer this question, we performed the QECCSD(TQ) calculations for the DZ model of N_2 (see Table 3.2 and Figure 3.2). As shown in Table 3.2, the errors in the QECCSD(TQ) results are somewhat greater than in the case of the STO-3G basis set, but the overall patterns are the same. Thus, the QECCSD(TQ) method provides considerable improvements in the CCSD and QECCSD results, reducing the large negative, (-70) - (-121) millihartree, errors in the CCSD results in the $R = 2R_e - 2.25R_e$ region and the relatively large positive, 40-50 millihartree, errors in the QECCSD results in the same region to 16–20 millihartree. For smaller values of R, the improvements offered by the QECCSD(TQ) approach are even greater. For example, the QECCSD(TQ) method reduces the 6.236 and 8.289 millihartree errors in the QECCSD and CCSD energies at $R = R_e$ and the 23.485 and 33.545 millihartree errors in the QECCSD and CCSD energies at $R = 1.5R_e$ to 1.002 and 6.011 millihartree, respectively. As shown in Table 3.2 and Figure 3.2, the QECCSD(TQ) potential energy curve for the DZ N₂ molecule is located above the full CI curve and the QECCSD(TQ) approach completely eliminates the unphysical humps on the CR-CCSD(T) and CR-CCSD(TQ) curves, obtained with the CCSD values of T_1 and T_2 . This shows that the use of the QECCSD rather than CCSD values of the T_1 and T_2 clusters improves the results of the MMCC calculations. All of this is very encouraging, since the QECCSD(TQ) method is a relatively inexpensive single-reference approach employing the spin-adapted RHF reference. As shown in Table 3.2 and Figure 3.2, all standard CC methods using the RHF determinant as a reference, including the very expensive CCSDT and $CCSDT(Q_f)$ approaches, which require iterative steps that scale as \mathcal{N}^8 with the system size, completely break down at larger N-N separations. The QECCSD(TQ) method provides a much smoother and more accurate description of triple bond breaking in N_2 , with the relatively small errors which monotonically increase with R, while eliminating all of the pathologies observed in the standard CC calculations.

Interestingly enough, even the simplest QECCSD(T) method works reasonably well, when the DZ model of N_2 is examined (see Table 3.2), although one has to keep in mind that the errors in the QECCSD(T) energies in the $R \leq 1.5R_e$ region are 2-3 times larger than in the QECCSD(TQ) case (see Table 3.2). Moreover, the QECCSD(T) energies do not vary with R as smoothly as the QECCSD(TQ) energies (cf. the nonmonotonic changes in errors in the QECCSD(T) results shown in Table 3.2). The same behavior is observed when the STO-3G basis set is employed. The calculations for the STO-3G basis set show that the QECCSD(T) and ECCSD(T) methods are incapable of providing significant improvements in the corresponding QECCSD and ECCSD results when the internuclear separation R becomes large (see Table 3.1). This is related to the absence of the quadruply excited moments $\mathcal{M}_{abcd}^{ijkl}(2)$ in the QECCSD(T) and ECCSD(T) energy expressions. It is interesting to observe, though, that the QECCSD(T) results for the DZ model of N₂ are much better than the results of the CR-CCSD(T) calculations (not to mention the CCSD(T) results), in which the quadruply excited moments $\mathcal{M}^{ijkl}_{abcd}(2)$ are neglected too. This shows once again that the T_1 and T_2 clusters resulting from QECCSD (and ECCSD) calculations are of much higher quality than the T_1 and T_2 clusters obtained with the standard CCSD approach, improving the results of the MMCC calculations.

3.4 Summary

We have demonstrated that the noniterative ECCSD(TQ) and QECCSD(TQ) methods, obtained by merging the ECC formalism of Arponen and Bishop with the generalized version of the MMCC theory that enables one to use the non-standard cluster components to design the noniterative corrections to CC energies, provide an accurate and variational description of potential energy surfaces involving multiple bond breaking with the ease of a single-reference "black-box" calculation. In particular, the GMMCC-based ECCSD(TQ) and QECCSD(TQ) approximations employing T_1 and T_2 clusters obtained in ECCSD and QECCSD calculations do not suffer from the non-variational collapse or unphysical behavior observed in the standard CCSD, CCSD(T), CCSD(TQ_f), CCSDT, and CCSDT(Q_f) calculations. The use of the ECCSD and QECCSD values of the singly and doubly excited clusters has a very positive impact on improving the results of the MMCC calculations in the bond breaking region. In particular, the ECCSD(TQ) and QECCSD(TQ) methods employing the ECCSD and QECCSD values of T_1 and T_2 clusters improve the results of the CR-CCSD(T) and CR-CCSD(TQ) calculations for triple bond breaking in N_2 , which also use the MMCC theory but rely on the T_1 and T_2 clusters obtained with the standard CCSD approach.

The results obtained in this thesis show that the new GMMCC theory is a flexible formalism, which enables one to design the relatively inexpensive noniterative CC methods employing the non-standard values of cluster components. The generalized version of the MMCC theory provides us with precise information about the many-body structure of the exact energy and the corrections that must be added to standard or non-standard CC energies to recover full CI results. This is particularly valuable in situations involving multiple bond stretching or breaking, where the standard arguments based on MBPT fail due to the divergence of the MBPT series at larger internuclear separations. The fact that one can use the standard as well as non-standard cluster amplitudes in the GMMCC calculations is a very useful feature, which gives us an opportunity to improve the results by using non-traditional sources of cluster amplitudes that are more suitable for the applications of interest. The results obtained in this thesis, in which we used the ECCSD and QECCSD methods to generate the T_1 and T_2 clusters for the GMMCC calculations, are a clear demonstration of how useful the idea of using the non-standard values of cluster amplitudes might be in the most challenging cases involving multiple bond breaking where all standard single-reference CC methods (including high level methods such as CCSDT or CCSDT(Q_f)) fail.

We tested a few ECCSD- and QECCSD-based GMMCC approximations, including ECCSD(T), QECCSD(T), ECCSD(TQ), and QECCSD(TQ). Although the best results are obtained with the ECCSD(TQ) method, we cannot recommend this method at this time for practical calculations due to the large costs of the underlying ECCSD calculations. The QECCSD(TQ) approach employing the QECCSD values of T_1 and T_2 clusters is more promising in this regard, offering an accurate description of large nondynamic and substantial dynamic correlation effects with the computational steps that scale as $\mathcal{N}^6 - \mathcal{N}^7$ with the system size.

ן גר		rielateu calcula	automs we	ninoriad ar		fonsig-nanod v		·		
Pp Wp	CCSD (CR-CCSD(TQ),b	QECCSD	ECCSD Mo	(2)/QECCSD ^c	M ₀ (2)/ECCSD ^d	QECCSD(T)*	ECCSD(T) ^f	QECCSD(TQ) ^g	ECCSD(TQ) ^h
1.5	1.102	0.202	0.886	0.885	0.704	0.666	0.404	0.404	0.195	0.195
2.0	3.295	0.687	1.898	1.897	0.503	0.442	1.866	1.865	0.612	0.613
2.5	9.220	2.497	3.443	3.427	-2.752	-2.878	6.419	6.405	1.966	1.970
3.0	13.176	4.483	3.909	3.757	-6.337	-6.850	8.001	7.836	3.485	3.492
3.5	-38.645	6.168	5.294	4.746	16.028	14.867	-0.985	-1.661	3.252	3.119
4.0	-140.376	-5.820	15.815	14.148	65.713	62.848	2.822	1.148	3.597	2.950
4.5	-184.984	-26.557	27.792	24.198	111.844	106.097	12.915	10.336	4.653	3.397
5.0	-200.857	-39.221	35.335	30.590	143.639	134.790	22.336	17.811	5.212	3.589
5.5	-206.974	-46.162	39.983	34.158	163.434	152.116	28.225	22.704	5.574	3.654
6.0	-209.538	-49.962	42.609	36.082	175.122	161.804	31.847	25.497	5.806	3.624
7.0	-211.915	-53.158	44.839	37.671	185.091	169.847	35.046	27.951	6.031	3.648
8.0	-213.431	-54.336	45.508	38.161	187.896	171.935	35.988	28.618	6.119	3.672

Table 3.1.: The ground-state energies of the N_2 molecule obtained for several internuclear separations R with the STO-3G basis J with the A. Set a All F.C.C. releted nelvinition

-107.546614, -107.473442, -107.447822, -107.441504, -107.439549, -107.438665, -107.438265, -107.438054, and -107.438029 hartree at R = 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 7.0, and 8.0 bohr, respectively. The lowest two occupied orbitals were frozen in ^aAll energies are in millihartree relative to the corresponding full CI energy values, which are -106.720117, -107.623240, -107.651880, ^bThe N-N separation in bohr. The equilibrium value of R is 2.068 bohr. correlated calculations.

^cThe zero-body moment or the CCSD-like energy expression, Eq. (3.27), calculated using the QECCSD values of T_1 and T_2 . ^dThe zero-body moment or the CCSD-like energy expression, Eq. (3.27), calculated using the ECCSD values of T_1 and T_2 . ^gThe GMMCC(2,4) result obtained using the QECCSD values of T_1 and T_2 and $|\Psi_0\rangle$ defined by Eq. (3.25). ^eThe GMMCC(2,3) result obtained using the QECCSD values of T_1 and T_2 and $|\Psi_0\rangle$ defined by Eq. (3.23) ¹The GMMCC(2,4) result obtained using the ECCSD values of T_1 and T_2 and $|\Psi_0\rangle$ defined by Eq. (3.25). ^fThe GMMCC(2,3) result obtained using the ECCSD values of T_1 and T_2 and $|\Psi_0\rangle$ defined by Eq. (3.23).

Method	0.75R _e	$R_e{}^{b}$	$1.25R_{e}$	$1.5R_e$	$1.75R_{e}$	$2R_e$	$2.25R_e$
CCSD	3.132	8.289	19.061	33.545	17.714	-69.917	-120.836
CCSDT ^c	0.580	2.107	6.064	10.158	-22.468	-109.767	-155.656
CCSD(T) ^c	0.742	2.156	4.971	4.880	-51.869	-246.405	-387.448
$\rm CCSD(TQ_f)^c$	0.226	0.323	0.221	-2.279	-14.243	92.981	334.985
$\operatorname{CCSDT}(\operatorname{Q}_{f})^{d}$	0.047	-0.010	-0.715	-4.584	3.612	177.641	426.175
CR-CCSD(T) ^c	1.078	3.452	9.230	17.509	-2.347	-86.184	-133.313
$CR-CCSD(TQ),a^{c}$	0.448	1.106	2.474	5.34 1	1.498	-40.784	-69.259
CR-CCSD(TQ),b ^c	0.451	1.302	3.617	8.011	13.517	25.069	14.796
QECCSD	2.506	6.236	13.609	23.485	31.060	40.085	49.741
$M_0(2)/\mathrm{QECCSD}^{\mathrm{e}}$	1.908	4.164	7.935	13.623	33.109	70.120	106.399
QECCSD(T) ^f	1.000	2.941	7.121	11.661	8.454	10.330	17.977
QECCSD(TQ) ^g	0.412	1.002	2.443	6.011	11.393	16.103	19.958

Table 3.2.: The ground-state energies of the N_2 molecule obtained for several internuclear separations R with the DZ basis set.^a

^aAll energies are in millihartree relative to the corresponding full CI energy values, which are -108.549027, -109.105115, -109.054626, -108.950728, -108.889906, -108.868239, and -108.862125 hartree at $R = 0.75R_e$, R_e , $1.25R_e$, $1.5R_e$, $1.75R_e$, $2R_e$, and $2.25R_e$, respectively. The lowest two occupied and the highest two unoccupied orbitals were frozen in correlated calculations.

^bThe equilibrium value of R, $R_e = 2.068$ bohr.

^cFrom Ref. 31.

^dFrom Ref. 34.

^eThe zero-body moment or the CCSD-like energy expression, Eq. (3.27), calculated using the QECCSD values of T_1 and T_2 (obtained with the Arponen-Bishop ECC theory).

^fThe GMMCC(2,3) result obtained using the QECCSD values of T_1 and T_2 and $|\Psi_0\rangle$ defined by Eq. (3.23) (obtained with the Arponen-Bishop ECC theory).

^gThe GMMCC(2,4) result obtained using the QECCSD values of T_1 and T_2 and $|\Psi_0\rangle$ defined by Eq. (3.25) (obtained with the Arponen-Bishop ECC theory).



Figure 3.1: Ground-state potential energy curves of the N_2 molecule as described by the STO-3G basis set. The QECCSD and ECCSD calculations were performed using the ECC theory of Arponen and Bishop.



Figure 3.2: Ground-state potential energy curves of the N_2 molecule as described by the DZ basis set. The QECCSD calculations were performed using the ECC theory of Arponen and Bishop.

Chapter 4 Exactness of Two-Body Cluster Expansions in Many-Body Quantum Theory

As demonstrated in Chapters 2 and 3, we can considerably improve the description of chemical bond breaking and quasi-degenerate electronic states by performing ECCSD calculations and by adding new types of noniterative energy corrections, based on the GMMCC formalism, to the ECCSD energies. The ECCSD approaches and the ECCSD-based GMMCC approaches are capable of eliminating the failures of the standard CCSD, CCSD(T), and CCSD(TQ_f) methods, in spite of the fact that we only use one- and two-body cluster operators in the ECCSD and ECCSD-based GMMCC calculations. Our positive experiences with the ECCSD-based methods imply that there is a lot of unexplored flexibility in the CC theories using only one- and two-body clusters. The natural question arises if one can improve the quality of many-electron wave functions based on the cluster expansions involving one- and two-body (or only two-body) operators to a degree where the results become exact or virtually exact. After all, the many-electron Hamiltonians used in quantum chemistry and atomic and molecular physics do not contain higher-than-two-body terms. Thus, one may wonder if there is a way to obtain the exact or virtually exact description of manyelectron systems with the CC-like theories that does not use higher-than-two-body operators to construct the corresponding many-particle wave functions. An issue of the exactness of the exponential cluster expansions employing two-body operators is discussed in this chapter.

4.1 Theory

It is well known that one can always obtain the exact solution of the electronic Schrödinger equation within a given basis set by performing the full CI calculations. Unfortunately, the dimension of the full CI eigenvalue problem can easily run into astronomical figures, even for small many-electron systems. This can be easily verified by applying the well-known Weyl's formula⁹⁷ for the number f(n, N, S) of spinadapted electron configurations that enter the full CI expansion of an eigenstate of the electronic Hamiltonian H, i.e.,

$$f(n, N, S) = \frac{2S+1}{n+1} {n+1 \choose N/2 - S} {n+1 \choose N/2 + S + 1}, \qquad (4.1)$$

where *n* represent the number of correlated orbitals, *N* is the number of correlated electrons, *S* is the total spin of an eigenstate under consideration, and $\binom{m}{k} = m!/[k!(m-k)!]$ is the binomial coefficient. For example, a full CI calculation of a singlet electronic state for a system consisting of 10 electrons (e.g., the HF molecule or the Ne atom) and described by only 20 orbitals requires using f(20, 10, 0) = 52, 581, 816 configurations. A modest increase of the number of orbitals from 20 to 30 results in a steep increase in the number of configurations defining the full CI problem to 4.04×10^9 . Those numbers should be compared with the much smaller numbers of 22,155 and 108,345 two-electron integrals defining the electronic Hamiltonians for the n = 20 and 30 cases, respectively. The point-group symmetry and other symmetries of the Hamiltonian can, on occasion, reduce the dimension of the full CI problem, but savings resulting from the use of symmetry are minimal compared to the rapidly increasing numbers of configurations defining the full CI wave functions with *n* and *N*.

Indeed, the realistic calculation for a 10 electron system would require using ~ 100 orbitals. According to Eq. (4.1), the n = 100, N = 10, and S = 0 case leads to an astronomical number of 9.94×10^{14} configurations in the corresponding full CI expansion. This should be compared to a much smaller number of 12,753,775 two-electron integrals defining the electronic Hamiltonian for the n = 100 case. Clearly, the numbers of full CI configurations would be significantly larger for N > 10 and for larger n values due to the factorial scaling of the dimension of the full CI problem with the system size. This should be contrasted with the relatively slow, n^4 -like, increase of the number of two-electron integrals defining the Hamiltonian H with n. Thus, there seems to be a conflict between the huge dimensionality of the full CI problem, which prevents us from performing exact ab initio calculations for larger systems, and the fact that the Hamiltonians for many-electron systems involve only one- and two-body integrals, whose numbers are much smaller than the numbers of full CI coefficients defining the exact wave functions. In spite of the tremendous progress in the area of full CI calculations and computer technology (cf., e.g., Ref. 98 and references therein), the exact *ab initio* calculations employing the full CI method remain limited to a few electron systems described by small (usually, $n \sim 20 - 30$) basis sets.

4.1.1 The $\exp(X)$ Conjecture

It has recently been suggested that it may be possible to represent the exact or virtually exact ground-state wave function of an arbitrary many-fermion pairwise interacting system by an exponential cluster expansion involving a general two-body operator⁹⁹⁻¹⁰⁵. If these statements were true, completely new ways of performing *ab initio* quantum calculations for many-fermion (e.g., many-electron) systems might be suggested, which could provide enormous reductions in computational require-

ments for accurate quantum calculations for pairwise-interacting many-fermion systems, eliminating the astronomical costs of generating the exact many-particle wave functions by solving the full CI eigenvalue problem. Specifically, it has been proposed that the exact ground-state wave function $|\Psi_0\rangle$ of a given many-fermion system described by the Hamiltonian,

$$H = z_p^q c^p c_q + \frac{1}{2} v_{pq}^{rs} c^p c^q c_s c_r,$$
(4.2)

containing up to two-body terms, obtained in a finite spin-orbital basis set, has the following simple form⁹⁹:

$$|\Psi_0\rangle \equiv |\Psi_0(X)\rangle = e^X |\Phi_0\rangle, \qquad (4.3)$$

where X is a general two-body operator and $|\Phi_0\rangle$ is a normalized reference state, which in principle is an arbitrary wave function that has a nonzero overlap with $|\Psi_0\rangle$, but in practice should provide us with a reasonable approximation of $|\Psi_0\rangle$. In Eq. (4.2) and equations presented below, we use, whenever possible, the Einstein summation convention over repeated upper and lower indices. As in the earlier parts of this thesis, the c^p and c_p operators represent the usual creation and annihilation operators, respectively, associated with the one particle basis $\{p\}$. The $z_p^q = \langle p|\hat{z}|q\rangle$ and $v_{pq}^{rs} = \langle pq|\hat{v}|rs\rangle$ represent the usual one- and two-particle integrals defining the Hamiltonian. In the language of second quantization,

$$X = X_2 = \frac{1}{2} x_{pq}^{rs} c^p c^q c_s c_r, \qquad (4.4)$$

where x_{pq}^{rs} are some coefficients. According to Nooijen⁹⁹ (cf., also, Refs. 104, 105), the number of independent coefficients x_{pq}^{rs} should be identical to the number of twoparticle integrals v_{pq}^{rs} entering the Hamiltonian H, Eq. (4.2). One could redefine the operator X by considering the one- and two-body components in Eq. (4.4)^{99,101,102} and write

$$X = X_1 + X_2 = x_p^q c^p c_q + \frac{1}{2} x_{pq}^{rs} c^p c^q c_s c_r, \qquad (4.5)$$

but this is not really necessary, since for a fixed number of particles (N), one can always rewrite the Hamiltonian H, Eq. (4.2), in terms of two-body terms only. A straightforward manipulation shows that

$$H = \frac{1}{2} h_{pq}^{rs} c^p c^q c_s c_r, \qquad (4.6)$$

where

$$h_{pq}^{rs} = v_{pq}^{rs} + (z_p^r \delta_q^s + \delta_p^r z_q^s) / (N - 1),$$
(4.7)

with δ_p^q representing the usual Kronecker delta. On the other hand, it may be beneficial to use Eq. (4.5) rather than Eq. (4.4) in actual calculations, since the presence of one-body term $X_1 = x_p^q c^p c_q$ in the operator X may accelerate the convergence of the resulting wave functions and energies.

The above representation of the exact ground-state wave function, Eq. (4.3), is reminiscent of the exponential ansatz of the single-reference CC theory, Eq. (2.4). There is, however, a fundamental difference between Eqs. (4.3) and (2.4). The cluster operator T entering Eq. (2.4) is defined in terms of the particle-hole excitation operators $E_{i_1 \dots i_n}^{a_1 \dots a_n}$, Eq. (2.7), where i_1, \dots, i_n (a_1, \dots, a_n) are the spin-orbitals that are occupied (unoccupied) in the reference configuration $|\Phi_0\rangle$ (which is then a single Slater determinant), and it contains all many-body terms T_n with $n = 1, \dots, N$ in the exact case. The operator X, Eq. (4.4), or its analog defined by Eq. (4.5), entering Eq. (4.3), has at most two-body terms, but of the general type (excitations, deexcitations, and all other combinations of indices p, q, r, s). Moreover, the reference configuration $|\Phi_0\rangle$ can be a multi-determinantal state. One could, of course, truncate the manybody expansion for the cluster operator T in Eq. (2.4) at a two-body component T_2 , Eq. (2.27), but then the resulting wave functions $e^{T_2}|\Phi_0\rangle$ and $e^{T_1+T_2}|\Phi_0\rangle$, which are used in the CCD and CCSD methods, respectively, are only approximate wave functions. Thus, as we can see, there are significant differences between Eqs. (4.3) and (2.4). However, because of the formal similarity of the CCD wave function, $e^{T_2}|\Phi_0\rangle$, and Eq. (4.3), Nooijen and Lotrich¹⁰⁶ and Van Voorhis and Head-Gordon¹⁰⁴ call the wave function ansatz defined by Eqs. (4.3) and (4.4) generalized CCD (GCCD) (the wave function ansatz using Eqs. (4.3) and (4.5) is then called generalized CCSD or GCCSD). A similar terminology has been used in Refs. 100–103.

4.1.2 Formal Arguments in Favor of the exp(X) Conjecture (Ground States)

There are several facts that speak in favor of the correctness of Eq. (4.3). Nooijen based his reasoning⁹⁹ on the fact that the number of two-body coefficients x_{pq}^{rs} is identical to the number of components of the Nakatsuji two-particle density equation¹⁰⁷, which is, in turn, equivalent to the time-independent Schrödinger equation for Hamiltonians containing up to two-body terms. The problem that was left unsolved by Nooijen is the solubility of a rather complicated exponential variant of the Nakatsuji density equation, which forms an essential part of Nooijen's analysis (cf. Eq. (11) in Ref. 99). Van Voorhis and Head-Gordon based their reasoning¹⁰⁴ on the fact (exploited in Quantum Monte Carlo techniques) that one can always obtain the exact wave function by considering the expression

$$|\Psi_0\rangle = \lim_{t \to \infty} e^{Z_t} |\Phi_0\rangle, \tag{4.8}$$

where the two-body operator Z_t is defined as follows:

$$Z_t = -(H - E_0)t, (4.9)$$

with E_0 representing the exact energy. They used Eq. (4.8) to write the two-body operator X defining the exact wave function $|\Psi_0\rangle$ via Eq. (4.3) as

$$X = -\lim_{t \to \infty} Z_t. \tag{4.10}$$

Similar arguments and equations have been presented by Nakatsuji^{102,103}, who also considered the conditions for the wave functions parameterized by two-body operators to be exact¹⁰⁰. The problem with using Eq. (4.8) in this fashion is that the operator Z_t , Eq. (4.9), provides the exact wave function only in the $t \to \infty$ limit, whereas the operator X, Eq. (4.4) or (4.5), entering Eq. (4.3), is a finite operator. This immediately implies that the operator X defining the exact wave function $|\Psi_0\rangle$ through Eq. (4.3) cannot be constrained to be of the Hamiltonian form, Eq. (4.9) (cf. below for additional remarks). Eq. (4.8) does not open up the possibility of the existence of a finite two-body operator X, which is not necessarily defined through infinite coefficients x_{pq}^{rs} (which Eq. (4.10) produces). There also seem to exist some contradictions between statements made in Refs. 100 and 102, 103. The exactness of the GCCD or GCCSD wave functions was questioned in Ref. 100 and supported in Refs. 102, 103 (see Refs. 108–112 for further debate; see the discussion below for additional comments).

We have recently provided an evidence¹⁰⁵ that the exact ground state of a manyfermion system, described by the Hamiltonian containing one- and two-body terms, may indeed be represented by the exponential cluster expansion employing a general two-body operator, Eq. (4.3), by connecting the problem with the Horn-Weinstein formula for the exact energy¹¹³,

$$E_{0} = \lim_{t \to \infty} \langle \Phi_{0} | e^{-tH} H | \Phi_{0} \rangle / \langle \Phi_{0} | e^{-tH} | \Phi_{0} \rangle$$

=
$$\lim_{t \to \infty} E_{0,t} = \lim_{t \to \infty} E(X_{t}), \qquad (4.11)$$

where

$$E_{0,t} \equiv E(X_t) = \langle \Phi_0 | e^{X_t^{\dagger}} H e^{X_t} | \Phi_0 \rangle / \langle \Phi_0 | e^{X_t^{\dagger}} e^{X_t} | \Phi_0 \rangle$$
(4.12)

and

$$X_t = -\frac{1}{2}tH,\tag{4.13}$$

and by determining the operator X entering Eq. (4.3) through a direct minimization of the expectation value expression

$$E_0(\tilde{X}) = \langle \Phi_0 | e^{\tilde{X}^{\dagger}} H e^{\tilde{X}} | \Phi_0 \rangle / \langle \Phi_0 | e^{\tilde{X}^{\dagger}} e^{\tilde{X}} | \Phi_0 \rangle$$
(4.14)

over general two-body operators

$$\tilde{X} = \frac{1}{2} \tilde{x}_{pq}^{rs} c^p c^q c_s c_r \tag{4.15}$$

This analysis (taken from Ref. 105) is described below.

Let us consider the family \mathcal{M} of all two-body operators \tilde{X} , Eq. (4.15), that are defined by finite coefficients \tilde{x}_{Pq}^{rs} and that have a general structure of the Hamiltonian H, Eq. (4.6). This means that \mathcal{M} consists of all two-body operators that are, for example, Hermitian, since H is Hermitian; that satisfy relations, such as $\tilde{x}_{Pq}^{rs} = \tilde{x}_{qp}^{sr}$, since $h_{Pq}^{rs} = h_{qp}^{sr}$, etc. Obviously, the number of independent parameters \tilde{x}_{Pq}^{rs} is identical to the number of coefficients h_{Pq}^{rs} or v_{Pq}^{rs} defining the Hamiltonian. It should be noticed that all operators X_t , Eq. (4.13), and Z_t , Eq. (4.9), belong to \mathcal{M} , although \mathcal{M} is a much larger operator family, which contains infinitely many operators that are not multiples of H. This remark is important for the considerations discussed in this section, since one can always obtain the exact wave function and energy by applying Eqs. (4.8) and (4.11), with Z_t and X_t defined by Eqs. (4.9) and (4.13), respectively. As pointed out above, neither the operator Z_t nor its X_t analog can provide the exact description of a many-fermion system for a finite value of t. We should search for the operator X defining $|\Psi_0\rangle$ via Eq. (4.3) by minimizing the expectation value expression, $E_0(\tilde{X})$, Eq. (4.14), over all operators in \mathcal{M} .

Let us, therefore, examine what the direct minimization of $E_0(\tilde{X})$ in \mathcal{M} leads to. According to the Ritz variational principle, $E_0(\tilde{X})$ is bounded from below by the exact, full CI, energy, so that

$$E_0 \le E_0(\hat{X}) \tag{4.16}$$

for all operators $\tilde{X} \in \mathcal{M}$. This implies that there should exist a two-body operator $X \in \mathcal{M}$ that minimizes $E_0(\tilde{X})$. We can write

$$E_0(X) = \min_{\tilde{X} \in \mathcal{M}} E_0(\tilde{X}). \tag{4.17}$$

Obviously,

$$E_0 \le E_0(X). \tag{4.18}$$

Let us now consider the energy expression $E_{0,t}$, Eq. (4.12), for an arbitrary (fixed) value of t. We can immediately write,

$$E_0(X) < E_{0,t},\tag{4.19}$$

since $E_0(X)$ is a minimum value of $E_0(\tilde{X})$, Eq. (4.14), in a space of all two-body operators \tilde{X} , whereas $E_{0,t} = E_0(X_t)$ is the value of $E_0(\tilde{X})$ at $\tilde{X} = X_t$ (cf. Eqs. (4.14) and (4.12)). As a matter of fact, for a given value of t, one can always find a two-body operator Y from \mathcal{M} such that $E_0(Y) < E_{0,t}$. An example of such operator might be provided by $X_{t'}$ with t' > t, since, as shown in Ref. 113, $E_{0,t}$, Eq. (4.12), is a monotonically decreasing function of t. However, since the operator family \mathcal{M} is much larger than the "one-dimensional" manifold of operators X_t , which are multiples of H, there is a chance that there exist two-body operators $Y \in \mathcal{M}$ which satisfy $E_0(Y) <$ $E_{0,t}$ and which are not given by Eq. (4.13). This indicates that the operator X minimizing $E_0(\tilde{X})$ may very well be a finite operator (i.e., defined by finite coefficients x_{pq}^{rs} and not obtained by considering the limiting case of the $t \to \infty$ operators X_t), although we cannot provide a rigorous mathematical proof that this is indeed the case and the existence of finite operator X may depend on the actual form of the reference state $|\Phi_0\rangle$, which does not have to be a single Slater determinant (see, e.g. Refs. 108-110). The existence of a finite operator $X \in \mathcal{M}$ that minimizes $E_0(\tilde{X})$ according to Eq. (4.17) and that is not of the Hamiltonian form is supported by the numerical calculations for a few many-electron systems¹⁰⁵ (see the discussion below).

The inequalities (4.18) and (4.19) can be combined into the following result:

$$E_0 \le E_0(X) < E_{0,t},\tag{4.20}$$

true for any value of t. In view of the Horn-Weinstein energy expression, Eq. (4.11), by considering the $t \to \infty$ limit in Eq. (4.20), we obtain the identity

$$E_0 = E_0(X). (4.21)$$

This means that the two-body operator X, obtained by minimizing the expectation value expression $E_0(\tilde{X})$, Eq. (4.14), gives the exact energy E_0 and, by the virtue of the variational principle, the exact ground state $|\Psi_0\rangle$, as stated in Eq. (4.3).

The above analysis makes the exactness of Eq. (4.3) a real possibility, but one should not treat it as a complete mathematical proof of Eq. (4.3), since we cannot rigorously prove the existence of the finite coefficients x_{pq}^{rg} that would define the optimum operator X corresponding to a global minimum of $E_0(\tilde{X})$ for an arbitrary reference $|\Phi_0\rangle$. We can make, however, several useful observations. First, the reasoning presented above, which is based on combining the Horn-Weinstein energy formula, Eq. (4.11), with the minimization of $E_0(\tilde{X})$, has an advantage over the arguments given in Refs. 102–104 in that it frees us from necessarily assuming that the operator X can only be obtained by studying the $t \to \infty$ operators X_t , Eq. (4.13), or Z_t , Eq. (4.9). By minimizing $E_0(\tilde{X})$, Eq. (4.14), in a space of all two-body operators (or, equivalently, by minimizing $E_0(\tilde{X})$ in a finite-dimensional space of variables \tilde{x}_{pq}^{rs}), which is exactly how we obtained operator X in Ref. 105 (cf. the examples below), we may be able to find finite parameters x_{pq}^{rs} , defining the exact wave function $|\Psi_0\rangle$, precisely because the operator X is not constrained to be a multiple of the Hamiltonian. If the finite operator X, determined by some numerical procedure for minimizing $E_0(\tilde{X})$ in \mathcal{M} , is a local rather than a global minimum on the $E_0(\tilde{X})$ multi-parameter surface, then the resulting energy $E_0(X)$, calculated by substituting $\tilde{X} = X$ into Eq. (4.14), and the resulting wave function $|\Psi_0(X)\rangle$, Eq. (4.3), do not have to be exact. However, even in this case, the operator X may provide excellent results, opening up a possibility of using the exponential wave functions (4.3), with X defined by Eq. (4.4)(or (4.5)), in high-accuracy *ab initio* calculations. Second, the mathematical analysis described above does not tell us anything about the specific form of the reference state $|\Phi_0\rangle$ that should be used in the calculations exploiting the exp(X) ansatz. It is possible, for example, that the finite parameters x_{pq}^{rs} that give the exact state $|\Psi_0(X)
angle$ via Eq. (4.3) exist only for certain types of references $|\Phi_0\rangle$ that do not have to be represented by a single Slater determinant (based on the analyses presented in Refs. 108-110, it is rather unlikely that $|\Phi_0\rangle$ is a single determinant although there is no proof of this statement). It may happen, however, that highly accurate results are already obtained with the ordinary Hartree-Fock reference $|\Phi_0\rangle$ and the additional improvements are obtained by using a multi-determinantal reference state $|\Phi_0\rangle$ in the definition of $|\Psi_0(X)\rangle$, Eq. (4.3). All of these issues are explored here by preforming numerical calculations for a few small many-electron systems.

4.1.3 Extension of the $\exp(X)$ Conjecture to Excited States

The $\exp(X)$ conjecture was originally proposed for the ground states, but we can easily extend it to the excited states using the standard variational approach. For example, we can construct the trial wave function for the first excited state $|\Psi_1\rangle$ as

$$|\tilde{\Psi}_1(\tilde{X}^{(1)})\rangle = |\Psi_1(\tilde{X}^{(1)})\rangle - \langle \Psi_0 | \Psi_1(\tilde{X}^{(1)})\rangle | \Psi_0\rangle.$$
(4.22)

Here,

$$|\Psi_0\rangle = \mathcal{N}_0 \, e^{X^{(0)}} |\tilde{\Phi}_0\rangle \tag{4.23}$$

is the previously obtained normalized ground state (\mathcal{N}_0 is the normalization factor, $X^{(0)}$ is the optimum two-body operator representing $|\Psi_0\rangle$, and $|\tilde{\Phi}_0\rangle$ is the single- or multi-determinantal reference state used in the calculation of $|\Psi_0\rangle$) and

$$|\Psi_1(\tilde{X}^{(1)})\rangle = \mathcal{N}_1 e^{\tilde{X}^{(1)}} |\tilde{\Phi}_1\rangle \tag{4.24}$$

is the normalized $\exp(X)$ -like form of the first excited state, where $|\tilde{\Phi}_1\rangle$ represents the reference state for $|\Psi_1\rangle$. By minimizing the energy functional

$$E_1(\tilde{X}^{(1)}) = \frac{\langle \tilde{\Psi}_1(\tilde{X}^{(1)}) | H | \tilde{\Psi}_1(\tilde{X}^{(1)}) \rangle}{\langle \tilde{\Psi}_1(\tilde{X}^{(1)}) | \tilde{\Psi}_1(\tilde{X}^{(1)}) \rangle}$$
(4.25)

over all two-body operators $\tilde{X}^{(1)} \in \mathcal{M}$, we obtain the energy of the first excited state and the corresponding wave function $|\Psi_1\rangle \equiv |\Psi_1(X^{(1)})\rangle$, where $X^{(1)}$ is the optimum value of $\tilde{X}^{(1)}$ obtained by minimizing $E_1(\tilde{X}^{(1)})$, Eq. (4.25).

For the k-th excited state, we can define the trial wave function for variational calculations as

$$|\tilde{\Psi}_{k}(\tilde{X}^{(k)})\rangle = |\Psi_{k}(\tilde{X}^{(k)})\rangle - \sum_{m=0}^{k-1} \langle \Psi_{m} | \Psi_{k}(\tilde{X}^{(k)})\rangle | \Psi_{m}\rangle, \qquad (4.26)$$

where $|\Psi_m\rangle$, $m = 1, \ldots, k - 1$, are the previously computed states and

$$|\Psi_{k}(\tilde{X}^{(k)}) = \mathcal{N}_{k} e^{\tilde{X}^{(k)}} |\tilde{\Phi}_{k}\rangle$$
(4.27)

is the normalized $\exp(X)$ -like form of the k-th excited state $|\Psi_k\rangle$, with $|\tilde{\Phi}_k\rangle$ representing the reference state for the k-th excited state. By minimizing the energy functional

$$E_{k}(\tilde{X}^{(k)}) = \frac{\langle \tilde{\Psi}_{k}(\tilde{X}^{(k)}) | H | \tilde{\Psi}_{k}(\tilde{X}^{(k)}) \rangle}{\langle \tilde{\Psi}_{k}(\tilde{X}^{(k)}) | \tilde{\Psi}_{k}(\tilde{X}^{(k)}) \rangle}$$
(4.28)

over all two-body operators $\tilde{X}^{(k)} \in \mathcal{M}$, we obtain the energy of the k-th excited state and the corresponding wave function $|\Psi_k\rangle \equiv |\Psi_k(X^{(k)})\rangle$, where $X^{(k)}$ is the optimum value of $\tilde{X}^{(k)}$.

Since excited states of many-electron systems are almost always very multi-determinantal and since they often have significant singly excited components, we may improve accuracies and accelerate the convergence of the numerical calculations for excited states using Eqs. (4.22)-(4.28) by using Eq. (4.5) (rather than Eq. (4.4)) to represent operator $\tilde{X}^{(k)}$, while obtaining the reference states $|\tilde{\Phi}_k\rangle$ by diagonalizing the Hamiltonian H in a small space spanned by a few Slater determinants that we believe are important to describe the ground and excited states of interest. This is what we did in our excited-state calculations described below.

4.2 Numerical Results

In our calculations for a few many-electron systems, we first concentrated on the following two aspects of theory: (i) the existence of finite coefficients x_{pq}^{rs} defining the two-body operator X, Eq. (4.4), and (ii) the non-Hamiltonian nature of the operator X. Initially, we focused on the ground-state theory. We tested the exact groundstate theory, in which we used the unexpanded form of the exponential operator $e^{\tilde{X}}$ to define the $E_0(\tilde{X})$ energy rather than the truncated power series expansion in \tilde{X} . This was made possible by representing the operators H and \tilde{X} as matrices in the

finite-dimensional N-electron Hilbert spaces relevant to a molecular system under consideration (using all symmetry-adapted Slater determinants $|\Phi_0\rangle$ and $|\Phi_{i_1\cdots i_n}^{a_1\cdots a_n}\rangle$, n = 1, ..., N, defining the full CI problem, as basis states). In order to calculate the exact value of $E_0(\tilde{X})$, Eq. (4.14), in a given iteration of the numerical procedure used to minimize $E_0(\tilde{X})$, we first diagonalized the matrix representing \tilde{X} with some unitary matrix \tilde{U} to obtain the diagonal matrix $\tilde{D} = \tilde{U}\tilde{X}\tilde{U}^{-1}$. Next, we constructed $e^{\tilde{D}}$ by taking the exponentials of the diagonal elements of \tilde{D} . Finally, after constructing $e^{\tilde{D}}$, we calculated the matrix representing $e^{\tilde{X}}$ as $\tilde{U}^{-1}e^{\tilde{D}}\tilde{U}$ and applied it to a column vector representing $|\Phi_0\rangle$ to obtain $|\Psi_0(\tilde{X})\rangle$ according to Eq. (4.3). The value of $E_0(\tilde{X})$ was obtained by calculating $\langle \Psi_0(\tilde{X}) | H | \Psi_0(\tilde{X}) \rangle / \langle \Psi_0(\tilde{X}) | \Psi_0(\tilde{X}) \rangle$, using the matrices representing H and $|\Psi_0(\tilde{X})\rangle$, as described above. In order to determine the optimum operator X and the corresponding energy $E_0(X)$, we used the downhill simplex method⁸² to minimize the energy $E_0(\tilde{X})$ in \mathcal{M} (recall that \mathcal{M} is a family of all two-body operators). Typically, our calculations required ~ 100 iterations to obtain a reasonably converged result, although in some cases we had to iterate much longer, particularly if we wanted to determine high decimal places. We realize that the numerical procedures used to obtain the optimum operator X are not suitable for routine calculations or for larger many-electron systems. The main point of this study is addressing the question if the exact or virtually exact many-fermion wave functions can be represented by Eq. (4.3). In this context, the efficiency of the numerical procedures used to determine X is of the secondary importance.

Our test calculations were performed for a few many-electron systems, including, among others, the 8-electron/16-spin-orbital H8 model¹¹⁴, which consists of eight hydrogen atoms arranged in a distorted octagonal configuration and described by the minimum basis set (MBS) obtained by placing one s function on each hydrogen

An example of the H8 model is very important for the discussion of the atom. exactness of Eq. (4.3). This model provides us with a highly nontrivial situation, where the number of independent two-body parameters x_{pq}^{rs} , defining operator X, or the number of independent two-body integrals h_{pq}^{rs} , defining the Hamiltonian, is considerably smaller than the dimension of the corresponding N-electron Hilbert space. Indeed, the number of two-body parameters x_{pq}^{rs} defining operator X is in this case 186, whereas the number of all spin- and symmetry-adapted configurations defining the corresponding full CI problem is 468. The one-body parameter x_p^q that enter Eq. (4.5) do not change this situation. The total number of one- and two-body parameters defining X, Eq. (4.5), is 198, which is still a lot less than the number of full CI configurations describing the exact wave function. The H8 model is described by a single parameter α (in bohr), which describes the deviation of the geometry of the D_{2h} -symmetric H8 system from the regular octagon¹¹⁴. The following values of α were particularly important for testing: $\alpha = 1.0$ and $\alpha = 0.0001$. The $\alpha = 1.0$ H8 system is somewhat less demanding, since in this case the exact ground-state wave function is dominated by the RHF configuration $|\Phi_0\rangle$. The more demanding $\alpha = 0.0001$ H8 model is characterized by a strong configurational quasi-degeneracy of the ground electronic state involving the RHF reference $|\Phi_0\rangle$ and the doubly excited configuration $|\Phi_1\rangle$ of the HOMO-LUMO type¹¹⁴.

The results of our calculations for the ground-sate of the MBS H8 system are shown in Table 4.1. As we can see, the $\exp(X)$ ansatz gives remarkably accurate results for the ground-state of H8, even when the overlap of the reference state is as small as 0.668268, as is the case when $\alpha = 0.0001$ and we use the RHF determinant as a reference. Independent of the value of α , we obtain microhartree accuracies, which are a lot better than the accuracies resulting from various traditional CI and CC calculations. Our $\exp(X)$ calculations for the ground state employing two-body or one- plus two-body operators X are order of magnitude more accurate than the CISD and CCSD calculations, which also use one- and two-body operators only. In fact, the $\exp(X)$ calculations provide a significantly better description of the ground state than the CISDTQ calculations employing one-, two-, three-, and even fourbody excitations. The number of parameters used to describe the CISDTQ wave function is much bigger than the number of parameters used to define our $\exp(X)$ wave functions. The results of our $\exp(X)$ calculations employing at most two-body operators are as good as the CCSDTQ results. In fact, by using one- and two-body parameters in X and the two-determinantal reference $|\tilde{\Phi}_0\rangle$ (see Table 4.1), we obtain the results which are considerably better than those obtained with CCSDTQ, when the challenging case of the quasi-degenerate $\alpha = 0.0001$ H8 model is examined.

As we can see, the presence of one-body operator $X_1 = x_p^q c^p c_q$ in X improves the convergence toward the full CI results. The use of a two-determinantal reference state $|\tilde{\Phi}_0\rangle$, which reflects the predominantly two-determinantal character of the ground-state wave function at $\alpha = 0.0001$, has a positive impact on the accuracy of $\exp(X)$ calculations. Even if our numerical procedures do not produce the exact (in a mathematical sense) energies, the microhartree accuracies, obtained with the exponential cluster expansions involving up to two-body operators only, are truly intriguing. The parameters x_{pq}^{rs} or x_p^q and x_{pq}^{rs} defining the optimum operators X are finite. For example, if we optimize X assuming that $X = X_2 = \frac{1}{2} x_{pq}^{rs} c^p c^q c_s c_r$, the largest values if x_{pq}^{rs} obtained in our ground-state calculations are 0.316845 at $\alpha = 1.0$ and 0.596623 at $\alpha = 0.0001$. The corresponding operators X do not commute with the Hamiltonian, so that the optimum X operators producing the highly accurate results in Table 4.1 are not of the Hamiltonian form.

In Table 4.2 we show the results of the $\exp(X)$ calculations for the first excited state of the ${}^{1}A_{g}$ symmetry. This state is dominated by the RHF configuration $|\Phi_{0}\rangle$ and the doubly excited determinant of the HOMO-LUMO type $|\Phi_1\rangle$ in the quasidegenerate, $\alpha = 0.0001$, region, so that it is natural to choose a two-determinantal state $|\tilde{\Phi}_1\rangle = c_{01}|\Phi_0\rangle + c_{11}|\Phi_1\rangle$ as a reference in our $\exp(X)$ calculations (see footnote "a" in Table 4.2). The $\exp(X)$ calculations for the first excited ${}^{1}A_{g}$ state were performed using the numerical procedure described in Section 4.1.3. To facilitate our numerical effort, we considered the truncated form of the $\exp(X)$ expansion, where the $\exp(X)$ series is truncated at the X^{50} term. Since the optimum coefficients x_p^q and x_{pq}^{rs} are rather small and X^n enters the $\exp(X)$ expansion as $\frac{1}{n!}X^n$, the truncation of the exp(X) expansion at the $\frac{1}{50!}X^{50}$ term produces essentially no errors (errors that cannot be detected in numerical calculations with the double precision Fortran). As shown in Table 4.2, the $\exp(X)$ calculations for the first excited ${}^{1}A_{g}$ state of H8 at $\alpha = 0.0001$ produce microhartree accuracies. For $\alpha = 1.0$ we obtained a somewhat larger error on the order of 0.4 millihartree, but this can be understood if we realize that the reference state $|\tilde{\Phi}_1\rangle$ is not a very good representation of the exact first excited-state wave function in this region of α . We must realize, however, that this error is on the same order as the error resulting from the equation-of-motion (EOM) CCSDTQ calculations (for basic information about EOMCC methods, see, e.g., Refs. 9, 60, 63), which use a much larger number of parameters to define the corresponding wave function than the $\exp(X)$ expansion employing one- and two-body parameters only. For $\alpha = 0.0001$, the exp(X) calculations give a microhartree-level accuracy for the first excited ${}^{1}A_{g}$ state of H8. None of the conventional CI or CC methods, including CISDTQ or even EOMCCSDTQ can produce the results of this quality.

Table 4.1.: A comparison of the ground-state energies of the MBS H8 system obtained with the $\exp(X)$ -like wave functions, where $X = X_2$ (a purely two-body operator) or $X_1 + X_2$ (a sum of one- and two-body operators), with the exact, full CI, energies and energies obtained in various CI and CC calculations. Full CI energies are total energies in hartree and all other energies are errors relative to full CI (also in hartree). We also give the overlaps of the normalized ground-state wave functions obtained with the $\exp(X)$ ansatz and the full CI approach. The overlaps of the reference states $|\Phi_0\rangle$ and $|\tilde{\Phi}_0\rangle$ with the full CI ground-state wave function $|\Psi_0\rangle$ are given for comparison.

Method (wave function)	Number of parameters in the wave function	$\alpha = 1.0$	$\alpha = 0.0001$
Energies			
Full CI	467	-4.352990	-4.204803
$\mathcal{N}_0 \exp(X_2) \Phi_0 angle^{\mathbf{a}}$	186	0.000008	0.000052
$\mathcal{N}_0 \exp(X_1 + X_2) \Phi_0 angle^{\mathbf{a}}$	198	0.000002	0.000020
$ ilde{\mathcal{N}}_0 \exp(X_1 + X_2) ilde{\Phi}_0 angle^{\mathrm{b}}$	198	0.000005	0.000007
CISD	46	0.008396	0.022779
CISDT	146	0.007984	0.016064
CISDTQ	320	0.000254	0.000449
CCSD	46	0.000546	0.005034
CCSDT	146	0.000026	-0.008362
CCSDTQ	320	-0.000001	-0.000035
Overlaps with a full CI wave function:			
$\overline{\mathcal{N}_0 \exp(X_2) \Phi_0 angle^{\mathrm{a}}}$		0.999998	0.999987
$\mathcal{N}_0 \exp(X_1 + X_2) \Phi_0 angle^{\mathbf{a}}$		0.999999	0.999995
$ ilde{\mathcal{N}_0} \exp(X_1 + X_2) ilde{\Phi}_0 angle^{\mathrm{b}}$		0.999998	0.999998
$ \Phi_0 angle^{\mathbf{a}}$		0.939657	0.668268
$ ilde{\Phi}_0 angle^{ extbf{b}}$		0.942804	0.909461

 $|\Phi_0\rangle$ is the ground-state RHF determinant.

^b $|\tilde{\Phi}_0\rangle = c_{00}|\Phi_0\rangle + c_{10}|\Phi_1\rangle$, where $|\Phi_0\rangle$ is the RHF determinant and $|\Phi_1\rangle$ is the doubly excited determinant of the (HOMO)² \rightarrow (LUMO)² type. The coefficients c_{00} and c_{10} defining the reference $|\tilde{\Phi}_0\rangle$ were obtained by diagonalizing the Hamiltonian in a space spanned by $|\Phi_0\rangle$ and $|\Phi_1\rangle$ and by selecting the lower energy eigenstate of H in this two-dimensional subspace as $|\tilde{\Phi}_0\rangle$.

Method (wave function)	Number of parameters in the wave function	$\alpha = 1.0$	$\alpha = 0.0001$
Energies			
Full CI	467	-3.998978	-4.144027
$ ilde{\mathcal{N}}_1 \exp(X_1 + X_2) ilde{\Phi}_1 angle^{\mathrm{a}}$	198	0.000405	0.000020
CISD	46	0.059314	0.042374
CISDT	146	0.031697	0.009726
CISDTQ	320	0.001608	0.000435
EOMCCSD	46	0.019605	0.015011
EOMCCSDT	146	-0.003174	-0.010505
EOMCCSDTQ	320	-0.000140	-0.000384
Overlaps with a full CI wave function:			
$\overline{ ilde{\mathcal{N}}_1 \exp(X_1 + X_2)} ilde{\Phi}_1 angle^{\mathrm{a}}$		0.999998	0.999992
$ ilde{\Phi}_1 angle^{\mathbf{a}}$		0.791762	0.902255

Table 4.2.: Same as Table 4.1 for the first excited state of the ${}^{1}A_{g}$ symmetry.

^a $|\tilde{\Phi}_1\rangle = c_{01}|\Phi_0\rangle + c_{11}|\Phi_1\rangle$, where $|\Phi_0\rangle$ is the RHF determinant and $|\Phi_1\rangle$ is the doubly excited determinant of the (HOMO)² \rightarrow (LUMO)² type. The coefficients c_{01} and c_{11} defining the reference $|\tilde{\Phi}_1\rangle$ were obtained by diagonalizing the Hamiltonian in a space spanned by $|\Phi_0\rangle$ and $|\Phi_1\rangle$ and by selecting the higher energy eigenstate of H in this two-dimensional subspace as $|\tilde{\Phi}_1\rangle$.

4.3 Summary

Let us summarize the results discussed in this chapter. By combining the theoretical arguments based on the Horn-Weinstein energy expression with variational principle and numerical calculations, we demonstrated that one can obtain the virtually exact description of pairwise interacting many-fermion systems, including molecular systems, by representing their ground- and exited-state wave functions by exponential cluster expansions employing general two-body or one- and two-body operators. Based on the evidence reported in this chapter, we can conclude that the optimum two-body or one- and two-body operators defining these cluster expansions are finite and not of the Hamiltonian form. The results discussed in this chapter confirm that one can tremendously improve the description of many-electron wave functions without using higher-than-two-body cluster operators. All of this implies that there is a lot of flexibility in the exponential cluster expansions, which was not utilized in the past. Based on the formal arguments and the extraordinarily high accuracies obtained in the calculations based on the $\exp(X)$ expansions, where X is a sum of one- and two-body components or a purely two-body operator, we can conclude that it is quite likely that one can formulate the exact or virtually exact many-electron theories based on these kinds of expansions.

Bibliography

Bibliography

- 1. F. COESTER, Nucl. Phys. 7, 421 (1958).
- 2. F. COESTER and H. KÜMMEL, Nucl. Phys. 17, 477 (1960).
- 3. J. ČÍŽEK, J. Chem. Phys. 45, 4256 (1966).
- 4. J. Čížek, Adv. Chem. Phys. 14, 35 (1969).
- 5. J. ČÍŽEK and J. PALDUS, Int. J. Quantum Chem. 5, 359 (1971).
- 6. J. PALDUS, in Methods in Computational Molecular Physics, edited by S. WIL-SON and G. DIERCKSEN, volume 293 of NATO Advanced Study Institute, Series B: Physics, pp. 99–194, Plenum Press, New York, 1992.
- 7. R. J. BARTLETT, in *Modern Electronic Structure Theory, Part I*, edited by D. R. YARKONY, pp. 1047–1131, World Scientific, Singapore, 1995.
- 8. T. J. LEE and G. E. SCUSERIA, in *Quantum Mechanical Electronic Structure* Calculations with Chemical Accuracy, edited by S. R. LANGHOFF, pp. 47–108, Kluwer, Dordrecht, 1995.
- 9. J. PALDUS and X. LI, Adv. Chem. Phys. 110, 1 (1999).
- 10. T. D. CRAWFORD and H. F. SCHAEFER III, *Rev. Comp. Chem.* 14, 33 (2000).
- 11. P. PIECUCH and K. KOWALSKI, in Computational Chemistry: Reviews of Current Trends, edited by J. LESZCZYŃSKI, volume 5, pp. 1–104, World Scientific, Singapore, 2000.
- P. PIECUCH, K. KOWALSKI, I. S. O. PIMIENTA, and S. A. KUCHARSKI, in Low-Lying Potential Energy Surfaces, edited by M. R. HOFFMANN and K. G. DYALL, volume 828, pp. 31-64, American Chemical Society, Washington, D.C., 2002.
- 13. P. PIECUCH, K. KOWALSKI, I. S. O. PIMIENTA, and M. J. MCGUIRE, Int. Rev. Phys. Chem. 21, 527 (2002).
- 14. G. D. PURVIS III and R. J. BARTLETT, J. Chem. Phys. 76, 1910 (1982).

- 15. M. URBAN, J. NOGA, S. J. COLE, and R. J. BARTLETT, J. Chem. Phys. 83, 4041 (1985).
- 16. K. RAGHAVACHARI, G. W. TRUCKS, J. A. POPLE, and M. HEAD-GORDON, Chem. Phys. Lett. 157, 479 (1989).
- 17. G. E. SCUSERIA, A. E. SCHEINER, T. J. LEE, J. E. RICE, and H. F. SCHAEFER III, J. Chem. Phys. 86, 2881 (1987).
- 18. G. E. SCUSERIA, C. L. JANSSEN, and H. F. SCHAEFER III, J. Chem. Phys. 89, 7382 (1988).
- 19. T. J. LEE and J. E. RICE, Chem. Phys. Lett. 150, 406 (1988).
- 20. P. PIECUCH and J. PALDUS, Int. J. Quantum Chem. 36, 429 (1989).
- 21. P. PIECUCH and J. PALDUS, Theor. Chim. Acta 78, 65 (1990).
- 22. P. PIECUCH, R. TOBOLA, and J. PALDUS, Int. J. Quantum Chem. 55, 133 (1995).
- 23. S. A. KUCHARSKI and R. J. BARTLETT, J. Chem. Phys. 108, 9221 (1998).
- 24. W. D. LAIDIG, P. SAXE, and R. J. BARTLETT, J. Chem. Phys. 86, 887 (1987).
- 25. K. B. GHOSE, P. PIECUCH, and L. ADAMOWICZ, J. Chem. Phys. 103, 9331 (1995).
- 26. P. PIECUCH, V. ŠPIRKO, A. E. KONDO, and J. PALDUS, J. Chem. Phys. 104, 4699 (1996).
- 27. L. ADAMOWICZ, P. PIECUCH, and K. B. GHOSE, Mol. Phys. 94, 225 (1998).
- 28. P. PIECUCH, S. A. KUCHARSKI, and R. J. BARTLETT, J. Chem. Phys. 110, 6103 (1999).
- 29. P. PIECUCH, S. A. KUCHARSKI, and V. ŠPIRKO, J. Chem. Phys. 111, 6679 (1999).
- 30. K. KOWALSKI and P. PIECUCH, J. Chem. Phys. 113, 18 (2000).
- 31. K. KOWALSKI and P. PIECUCH, J. Chem. Phys. 113, 5644 (2000).
- 32. K. KOWALSKI and P. PIECUCH, J. Mol. Struct.: THEOCHEM 547, 191 (2001).
- 33. K. KOWALSKI and P. PIECUCH, Chem. Phys. Lett. 344, 165 (2001).
- 34. P. PIECUCH, S. A. KUCHARSKI, and K. KOWALSKI, Chem. Phys. Lett. 344, 176 (2001).

- 35. P. PIECUCH, S. A. KUCHARSKI, V. ŠPIRKO, and K. KOWALSKI, J. Chem. Phys. 115, 5796 (2001).
- 36. P. PIECUCH, K. KOWALSKI, and I. S. O. PIMIENTA, Int. J. Mol. Sci. 3, 475 (2002).
- 37. M. J. MCGUIRE, K. KOWALSKI, and P. PIECUCH, J. Chem. Phys. 117, 3617 (2002).
- 38. Y. S. LEE and R. J. BARTLETT, J. Chem. Phys. 80, 4371 (1984).
- 39. Y. S. LEE, S. A. KUCHARSKI, and R. J. BARTLETT, J. Chem. Phys. 81, 5906 (1984), *ibid.* 82, 5761 (1985) (Erratum).
- 40. J. NOGA, R. J. BARTLETT, and M. URBAN, Chem. Phys. Lett. 134, 126 (1987).
- 41. G. W. TRUCKS, J. NOGA, and R. J. BARTLETT, Chem. Phys. Lett. 145, 548 (1988).
- 42. S. A. KUCHARSKI and R. J. BARTLETT, Chem. Phys. Lett. 158, 550 (1989).
- 43. J. NOGA and R. J. BARTLETT, J. Chem. Phys. 86, 7041 (1987), *ibid.* 89, 3401 (1988) (Erratum).
- 44. G. E. SCUSERIA and H. F. SCHAEFER III, Chem. Phys. Lett. 152, 382 (1988).
- 45. S. A. KUCHARSKI and R. J. BARTLETT, Theor. Chim. Acta 80, 387 (1991).
- 46. S. A. KUCHARSKI and R. J. BARTLETT, J. Chem. Phys. 97, 4282 (1992).
- 47. N. OLIPHANT and L. ADAMOWICZ, J. Chem. Phys. 95, 6645 (1991).
- 48. P. PIECUCH and L. ADAMOWICZ, J. Chem. Phys. 100, 5792 (1994).
- 49. M. MUSIAŁ, S. A. KUCHARSKI, and R. J. BARTLETT, J. Chem. Phys. 116, 4382 (2002).
- 50. J. S. ARPONEN, Ann. Phys. 151, 311 (1983).
- 51. J. S. ARPONEN, R. F. BISHOP, and E. PAJANNE, *Phys. Rev. A* 36, 2519 (1987).
- 52. J. S. ARPONEN, R. F. BISHOP, and E. PAJANNE, Condensed Matter Theor. 2, 357 (1987).
- 53. R. F. BISHOP, J. S. ARPONEN, and E. PAJANNE, in Aspects of Many-Body Effects in Molecules and Extended Systems, edited by D. MUKHERJEE, volume 50 of Lecture Notes in Chemistry, pp. 79–100, Springer, Berlin, 1989.
- 54. R. F. BISHOP and J. S. ARPONEN, Int. J. Quantum Chem. Symp. 24, 197 (1990).

- 55. J. S. ARPONEN and R. F. BISHOP, Ann. Phys. 207, 171 (1991).
- 56. R. F. BISHOP, N. I. ROBINSON, and J. S. ARPONEN, Condensed Matter Theor. 5, 37 (1990).
- 57. R. F. BISHOP, Theor. Chim. Acta 80, 95 (1991).
- 58. J. S. ARPONEN, Phys. Rev. A 55, 2686 (1997).
- 59. J. S. ARPONEN, R. F. BISHOP, and E. PAJANNE, *Phys. Rev. A* 36, 2539 (1987).
- 60. P. PIECUCH and R. J. BARTLETT, Adv. Quantum Chem. 34, 295 (1999).
- 61. E. A. SALTER, G. W. TRUCKS, and R. J. BARTLETT, J. Chem. Phys. 90, 1752 (1989).
- 62. E. A. SALTER and R. J. BARTLETT, J. Chem. Phys. 90, 1767 (1989).
- 63. J. F. STANTON and R. J. BARTLETT, J. Chem. Phys. 98, 7029 (1993).
- 64. J. F. STANTON and R. J. BARTLETT, J. Chem. Phys. 99, 5178 (1993).
- 65. S. PAL, Theor. Chim. Acta 66, 151 (1984).
- 66. S. PAL, Phys. Rev. A 33, 2240 (1986).
- 67. S. PAL, Phys. Rev. A 34, 2682 (1986).
- 68. K. B. GHOSE and S. PAL, Phys. Rev. A 36, 1539 (1987).
- 69. N. VAVAL, K. B. GHOSE, and S. PAL, J. Chem. Phys. 101, 4914 (1994).
- 70. N. VAVAL and S. PAL, Phys. Rev. A 54, 250 (1996).
- 71. A. B. KUMAR, N. VAVAL, and S. PAL, Chem. Phys. Lett. 295, 189 (1998).
- 72. N. VAVAL, A. B. KUMAR, and S. PAL, Int. J. Mol. Sci. 2, 89 (2001).
- 73. N. VAVAL, Chem. Phys. Lett. 318, 168 (2000).
- 74. P. PIECUCH, K. KOWALSKI, P.-D. FAN, and I. S. O. PIMIENTA, in Advanced Topics in Theoretical Chemical Physics, edited by J. MARUANI, R. LEFEBVRE, and E. BRÄNDAS, volume 12 of Progress in Theoretical Chemistry and Physics, pp. 119-206, Kluwer, Dordrecht, 2003.
- 75. P. PIECUCH, K. KOWALSKI, I. S. O. PIMIENTA, P.-D. FAN, M. LODRIGU-ITO, M. J. MCGUIRE, S. A. KUCHARSKI, T. KUŚ, and M. MUSIAŁ, *Theor. Chem. Acc.* **112**, 349 (2004).
- 76. P. PIECUCH, I. S. O. PIMIENTA, P.-D. FAN, and K. KOWALSKI, in Recent Progress in Electron Correlation Methodology, edited by A. WILSON and K. PETERSON, volume XXX of ACS Symposium Series, pp. XXX-XXXX, American Chemical Society, Washington, D.C., 2005, in press.

- 77. P.-D. FAN, K. KOWALSKI, and P. PIECUCH, Mol. Phys. (2005), in press.
- 78. J. PALDUS, P. PIECUCH, L. PYLYPOW, and B. JEZIORSKI, *Phys. Rev. A* 47, 2738 (1993).
- 79. P. PIECUCH and J. PALDUS, Phys. Rev. A 49, 3479 (1994).
- 80. P. PIECUCH, R. TOBOŁA, and J. PALDUS, Chem. Phys. Lett. 210, 243 (1993).
- 81. W. J. HEHRE, R. F. STEWART, and J. A. POPLE, J. Chem. Phys. 51, 2657 (1969).
- 82. Numerical Recipes in Fortran: The Art of Scientific Computing, Cambridge University Press, Cambridge, U.K., 1992.
- 83. S. A. KUCHARSKI and R. J. BARTLETT, J. Chem. Phys. 108, 5243 (1998).
- 84. S. A. KUCHARSKI and R. J. BARTLETT, J. Chem. Phys. 108, 5255 (1998).
- 85. T. H. DUNNING, J. Chem. Phys. 53, 2823 (1970).
- 86. K. KOWALSKI and P. PIECUCH, J. Chem. Phys. 115, 2966 (2001).
- 87. K. KOWALSKI and P. PIECUCH, J. Chem. Phys. 116, 7411 (2002).
- 88. K. KOWALSKI and P. PIECUCH, J. Chem. Phys. 120, 1715 (2004).
- 89. P. PIECUCH and K. KOWALSKI, Int. J. Mol. Sci. 3, 676 (2002).
- 90. K. KOWALSKI and P. PIECUCH, Mol. Phys. 102, 2425 (2004).
- 91. I. S. O. PIMIENTA, K. KOWALSKI, and P. PIECUCH, J. Chem. Phys. 119, 2951 (2003).
- 92. P. PIECUCH, K. KOWALSKI, and I. S. O. PIMIENTA, Int. J. Mol. Sci. 3, 475 (2002).
- 93. P. PIECUCH, M. WŁOCH, M. LODRIGUITO, and J. R. GOUR, in Recent Advances in the Theory of Chemical and Physical Systems, edited by S. WILSON, J.-P. JULIEN, J. MARUANI, E. BRÄNDAS, and G. DELGADO-BARRIO, volume XXX of Progress in Theoretical Chemistry and Physics, pp. XXX-XXXX, Springer, Berlin, 2005, in press.
- 94. P. PIECUCH, S. A. KUCHARSKI, K. KOWALSKI, and M. MUSIAŁ, Comp. Phys. Commun. 149, 71 (2002).
- 95. M. J. MCGUIRE, P. PIECUCH, K. KOWALSKI, S. A. KUCHARSKI, and M. MUSIAŁ, J. Phys. Chem. A 108, 8878 (2004).
- 96. M. J. MCGUIRE and P. PIECUCH, J. Am. Chem. Soc. 127, 2608 (2005).
- 97. H. WEYL, The Classical Groups, Their Invariants and Representations, Princeton University Press, Princeton, N.J., 1946.

- 98. C. D. SHERRILL and H. F. SCHAEFER III, Adv. Quantum Chem. 34, 143 (1999).
- 99. M. NOOIJEN, Phys. Rev. Lett. 84, 2108 (2000).
- 100. H. NAKATSUJI, J. Chem. Phys. 113, 2949 (2000).
- 101. H. NAKATSUJI and E. R. DAVIDSON, J. Chem. Phys. 115, 2000 (2001).
- 102. H. NAKATSUJI, J. Chem. Phys. 115, 2465 (2001).
- 103. H. NAKATSUJI, J. Chem. Phys. 116, 1811 (2002).
- 104. T. VAN VOORHIS and M. HEAD-GORDON, J. Chem. Phys. 115, 5033 (2001).
- 105. P. PIECUCH, K. KOWALSKI, P.-D. FAN, and K. JEDZINIAK, *Phys. Rev. Lett.* 90, 113001 (2003).
- 106. M. NOOIJEN and V. LOTRICH, J. Chem. Phys. 113, 4549 (2000).
- 107. H. NAKATSUJI, Phys. Rev. A 14, 41 (1976).
- 108. S. RONEN, Phys. Rev. Lett. 91, 123002 (2003).
- 109. E. R. DAVIDSON, Phys. Rev. Lett. 91, 123001 (2003).
- 110. D. A. MAZZIOTTI, Phys. Rev. A 69, 012507 (2004).
- 111. D. MUKHERJEE and W. KUTZELNIGG, Chem. Phys. Lett. 397, 174 (2004).
- 112. W. KUTZELNIGG and D. MUKHERJEE, Phys. Rev. A 71, 022502 (2005).
- 113. D. HORN and M. WEINSTEIN, Phys. Rev. D 30, 1256 (1984).
- 114. K. JANKOWSKI, L. MEISSNER, and J. WASILEWSKI, Int. J. Quantum Chem. 28, 931 (1985).

