DEVELOPMENT AND APPLICATIONS OF SEMI-STOCHASTIC COUPLED-CLUSTER METHODS FOR OPEN-SHELL SYSTEMS AND ELECTRONIC EXCITATIONS IN MOLECULES

By

Arnab Chakraborty

A DISSERTATION

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

Chemistry—Doctor of Philosophy

2024

ABSTRACT

It is well established that the exponential wave function ansatz of coupled-cluster (CC) theory and its equation-of-motion (EOM) extension to excited states are among the most appealing ways to describe the electronic structure of molecules. One of the key challenges in the development of the CC and EOMCC methodologies is the incorporation of many-electron correlation effects due to higher–than–two-body components of the cluster and EOM excitation operators, needed to achieve a quantitative description, without running into the usually prohibitive computational costs of the CC approach with singles, doubles, and triples (CCSDT) and its excited-state EOMCCSDT extension, the CC method with singles, doubles, triples, and quadruples (CCSDTQ) and its excited-state EOMCCSDTQ extension, *etc.*, and without resorting to perturbative CCSD(T)-type ideas that fail in multireference situations, such as bond breaking and excited states dominated by two-electron transitions. One of the most promising approaches in this area is the semi-stochastic CC(P;Q) methodology that identifies the most important higher–than–doubly excited determinants needed in the high-level CC/EOMCC calculations with the help of the stochastic configuration interaction (CI) and CC Quantum Monte Carlo (QMC) wave function propagations and uses the suitably designed deterministic iterative and noniterative steps of the CC(P;Q) formalism to converge the desired CCSDT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, *etc.*, energetics. In this dissertation, we first discuss our recent work on extending the semi-stochastic, CIQMC-driven, particle-conserving CC(P;Q) framework to excited electronic states and open-shell systems. We tested performance of the resulting methods by examining their ability to recover the high-level CCSDT/EOMCCSDT energetics in calculations of the electronic excitation spectra of the CH^+, CH, and CNC molecules and singlet–triplet gaps in a few biradical systems, including methylene, $(HFH)^-$, cyclobutadiene, cyclopentadienyl cation, and trimethylenemethane. The second part of this dissertation focuses on an alternative way of determining ground and excited states of open-shell systems within the EOMCC framework by turning to the single and double electron attachment (EA) and single and double ionization potential

(IP) EOMCC schemes. By generating ground and excited states of open-shell species, such as radicals and biradicals, with the help of suitably designed operators that can formally add electrons to or remove electrons from the parent closed-shell cores (an operation producing the appropriate multi-configurational reference space within a single-reference framework, while relaxing the remaining electrons), the resulting EA/IP- and DEA/DIP-EOMCC methods offer several advantages over the particle-conserving CC/EOMCC treatments, including rigorous spin and symmetry adaptation of the calculated electronic states and the ability of handling high- and low-spin states in an accurate and well-balanced manner. We demonstrate how to utilize the stochastic CIQMC wave function propagations, which are of particle-conserving character, in identifying the dominant higher-order 3-particle-2-hole ($3p$-$2h$)/3-hole-2-particle ($3h$-$2p$) and 4-particle-2-hole ($4p$-$2h$)/4-hole-2-particle ($4h$-$2p$) components of the respective particle-nonconserving electron attaching/ionizing operators, needed to obtain a quantitative description, without having to resort to the previously exploited user- and system-dependent active-space concepts. The effectiveness of the semi-stochastic, CIQMC-driven, EA/IP/DEA/DIP-EOMCC approaches will be illustrated by examining the adiabatic excitations in the C_2N, CNC, N_3, and NCO radicals and by revisiting the singlet–triplet gaps in methylene and trimethylenemethane.

To my parents, teachers, and friends.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my doctoral advisor, Professor Piotr Piecuch, for his invaluable advice, guidance, and support throughout my Ph.D. journey. He introduced me to the world of coupled-cluster and other many-body theories, which I sincerely appreciate. Professor Piecuch's readiness to engage in scientific discussions and his inspiring presence have been fundamental to my academic growth. I am grateful to him for sharing his deep knowledge, his patient mentoring, and his constant motivation to approach research with extreme care.

I would also like to acknowledge the current and previous members of my guidance committee: Professor Katharine Hunt, Professor James Jackson, Professor Angela Wilson, and Professor Benjamin Levine, for their support, time, and advice. I would like to thank Professor Marcos Dantus for the opportunity to work on the extensive study of H_3^+ generation from the double ionization of organic halogens and pseudo halogens. This collaboration has taught me how to communicate ideas to a diverse group of scientists. I would also like to thank Professors Madhav Ranganathan, Srihari Keshavamurthy, and Nishanth Nair from the Indian Institute of Technology Kanpur; Professor Satrajit Adhikari from the Indian Association for the Cultivation of Science, and Professor Pinaki Chaudhury from the University of Calcutta for sparking my interest in theoretical chemistry.

Furthermore, I would like to thank Dr. Kathryn Severin and Dr. Elizabeth McGaw, under whom I have had so many wonderful semesters of teaching opportunities for CEM 395 and CEM 495, and Professor Katharine Hunt, under whom I had the great opportunity to teach CEM 484. These teaching assignments gave me the scope to develop my communication skills and helped me understand the fundamental concepts of physical chemistry more deeply.

I am also grateful to the current and former Piecuch group members: Dr. Jun Shen, Dr. Suhita Basumallick, Dr. J. Emiliano Deustua, Dr. Ilias Magoulas, Dr. Stephen Yuwono, Mr. Karthik Gururangan, Mr. Tiange Deng, Ms. Swati Priyadarsini, and Mr. Agnibha Hanra. Their support, help, and encouragement made my Ph.D. journey unforgettable.

I would also like to thank all my friends from the East Lansing area. My Ph.D. experience would not have been so enjoyable without you all.

I want to thank my parents, Mr. Hari Sadhan Chakraborty and Mrs. Mira Chakraborty, for their continuous support throughout my life.

Thank you all!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Electronic structure theory is a major branch of quantum mechanics that aims to elucidate the behavior of electrons in atoms, molecules, and condensed matter systems. It provides a fundamental framework for understanding and predicting the properties of matter at the microscopic level. At its core, lies the time-independent electronic Schrödinger equation, which, within the Born–Oppenheimer approximation, has the form

$$H\Psi_\mu(\mathbf{X};\mathbf{R}) = E_\mu(\mathbf{R})\Psi_\mu(\mathbf{X};\mathbf{R}), \tag{1.1}$$

where H is the electronic Hamiltonian, Ψ_μ, which depends on the coordinates of all electrons $\mathbf{X}$ and parametrically on the coordinates of the nuclei $\mathbf{R}$ that provide the external potential for the motion of electrons, is the many-electron wave function characterizing the μ^{th} electronic state, with $\mu = 0$ denoting the ground state and $\mu > 0$ designating excited states, and E_μ that parametrically depends on the nuclear coordinates $\mathbf{R}$ is the corresponding total electronic energy. In the non-relativistic description, the electronic Hamiltonian operator for a system containing N electrons and M nuclei can be expressed in atomic units as

$$H = \sum_{i=1}^{N}(-\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}}) + \sum_{i<j}^{N}\frac{1}{r_{ij}} = Z + V, \tag{1.2}$$

where

$$Z = \sum_{i=1}^{N} z_i, \tag{1.3}$$

with $z_i = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}}$, is the one-body term that contains the kinetic energy of electrons and the coulombic attraction between the electrons and the frozen nuclei (with Z_A denoting the nuclear charge associated with the A^{th} nucleus) and

$$V = \sum_{i<j}^{N} v_{ij}, \tag{1.4}$$

with $v_{ij} = \frac{1}{r_{ij}}$, is the two-body operator describing the electron-electron coulombic repulsions. Although it has been nearly a century since the Schrödinger equation was proposed, we can

solve this equation analytically only for the one-electron systems, such as the hydrogen atom or the H_2^+ molecule. Nevertheless, due to the great theoretical and computational advances in the last few decades, it is possible to obtain numerical solutions for the description of increasingly large and complex molecular systems. However, an accurate determination of the electronic spectra and potential energy surfaces (PESs) of closed-shell species and open-shell systems, such as radicals and biradicals, along bond breaking coordinates, multiple bond breaking, and excited states dominated by two-electron transitions continue to pose a challenge. Much of this challenge originates from the fact that the above problems are governed by a complicated combination of two different types of many-electron correlation effects, including dynamical correlations, which result from the motion of electrons trying to avoid each other, especially at shorter distances, due to repulsive electron-electron interactions, and nondynamical or static correlations that result from electron entanglement at longer electron-electron separations. One of the biggest challenges of electronic structure theory is to design computationally affordable ways of describing dynamical and nondynamical correlations accurately and in a balanced manner.

In principle, one can try to address this challenge by turning to the full configuration interaction (FCI) approach, where the many-electron wave function is expressed as a linear combination of all possible Slater determinants within a basis set of one-electron functions, *i.e.*, spin-orbitals. Inserting this expansion into the electronic Schrödinger equation results in an eigenvalue problem, involving the matrix elements of H, which is equivalent to diagonalizing the Hamiltonian matrix in the many-electron Hilbert space spanned by the Slater determinants. This procedure provides the exact numerical solution of the Schrödinger equation in a basis, and the only remaining step then is to gradually increase the basis set dimension to reach the complete basis set (CBS) limit, which provides the true, numerically exact solution of the Schrödinger equation for the electronically bound states. Unfortunately, this seemingly straightforward approach bears a dimensionality that scales factorially with respect to the system size (*i.e.*, number of electrons and the basis set dimension), render-

ing it impractical for most of the systems encountered in chemical sciences. With only a few electrons in realistic basis sets, the computational cost of FCI diagonalization becomes unmanageable even with the most powerful supercomputers available today. Indeed, the combinatorial scaling of the computational costs associated with diagonalizing the Hamiltonian matrix in the complete N-electron Hilbert space with the number of electrons and one-electron basis functions is so severe that relying solely on the growth of the processing speed of classical computers is far from being sufficient. Quantum computers might address this issue, but devices in this category capable of producing noiseless FCI solutions for realistic electronic structure problems do not exist. Therefore, one of the main efforts in the area of quantum chemistry method development has been to formulate computationally tractable approaches capable of capturing FCI-quality many-electron correlation effects for a broad spectrum of chemically relevant systems without running into the prohibitive computational costs associated with the FCI approach.

Since constructing the many-electron wave functions using all possible Slater determinants and numerically exact coefficients at these determinants is not feasible, the simplest alternative, at least for the ground state, would be to use single determinants to represent them. This leads to the Hartree–Fock (HF) approximation, which is one of the most well-known independent particle models in which we apply the variational principle to optimize a single Slater determinant via a self consistent-field (SCF) procedure. Although the HF method is capable of recovering total ground-state electronic energies of molecular systems to within about 1%, it is very far from being a good solution, since it neglects many-electron correlation effects needed to describe covalent chemical bond breaking, electronic excitation, electron attachment, and ionization phenomena, and the vast majority of molecular properties of interest in chemistry. It also does not capture important physical effects, such as dispersion in van der Waals molecules, and is particularly bad in describing dissociations of low-spin species into high-spin open-shell fragments. This can be illustrated with the example of the single bond dissociation of F_2 into two F atoms [1]. In this case, the restricted

HF (RHF) solution, which requires each spatial molecular orbital (MO) to be either empty or doubly occupied by electrons of opposite spins, overbinds the dissociation energy by an order of magnitude, forcing F_2 to dissociate into F^+ and F^-, when it should dissociate into neutral F atoms. On the other hand, the unrestricted HF (UHF) method, which uses different spatial orbitals for different spins, produces spin-symmetry-broken solutions without binding the molecule at all. All of the above problems in the HF approximation originate from the mean-field treatment of electron-electron interactions that results in essentially uncorrelated motion of electrons. However, despite all the deficiencies of the HF theory, the HF determinants often serve as a good starting point for the so-called *post*-HF methods that generate the remaining Slater determinants in the wave function as particle-hole (p-h) excitation from them. In this way, we can capture the relevant many-electron correlation effects, whose strength in the ground state can be measured by the correlation energy, defined as the difference between the exact or FCI and HF energies, and describe all of the above phenomena that are driven by electron correlation.

Among all the correlated, post-HF, approaches that have been developed over the years to provide accurate approximate solutions to the Schrödinger equation, methods based on the coupled-cluster (CC) theory [2–7] and their suitable extensions to open-shell and excited states are widely recognized as offering the best balance between accuracy and computational cost, thus emerging as one of the most prominent alternatives to FCI [8, 9]. Historically, the CC theory emerged as an infinite-order generalization of the finite-order many-body perturbation theory (MBPT) by summing the linked wave function and connected energy diagrams to infinite order, using the linked [10–13] and connected [12, 13] cluster theorems. These two theorems provide the CC theory with some very important characteristics of the exact FCI description. For example, CC approaches are size extensive, *i.e.*, there is no loss of accuracy with an increase of the system size, because, contrary to truncated CI methods, only connected diagrams are present in the energy expressions and there are no unlinked contributions in the CC wave functions. Also, the CC theory relies on an exponential

ansatz for the wave function, and this allows for the separability or size consistency of the wave function in the non-interacting limit, provided the underlying reference state is also separable, enabling the CC methods to deal with fragmentation phenomena. These important properties, combined with the fast convergence of the truncated CC methods toward the exact FCI limit in the majority of molecular applications, have established the CC theory as the *de facto* standard for high-accuracy electronic structure calculations.

There are two distinct conceptual frameworks within the CC theory depending on the dimensionality of the reference or model space used as the zeroth-order description of the many-electron problem of interest. The oldest formalism, which uses a single Slater determinant as the reference state and Fermi vacuum, defines the single-reference (SR) CC methods. One can also consider multi-reference (MR) CC approaches where multi-dimensional model spaces, spanned by multiple Slater determinants, provide the zeroth-order states. Since many of the problems of interest, such as low-lying excited states dominated by two-electron transitions, electronic spectra of radical and biradical species, and excited-state potentials along bond stretching coordinates have an intrinsically MR character, one naturally tends to think of resorting to MRCC theories [8, 9, 14–18] and other MR theories [19–23] that are specifically designed to deal with these kinds of situations. Unfortunately, even the best MRCC theories suffer from various mathematical and numerical problems, such as unphysical [14, 15, 24] and singular [14, 15, 25–28] solutions, intruder states [14, 15, 24], and intruder solutions [15, 24]. Also, the multi-dimensional reference spaces grow factorially with respect to the number of active electrons and orbitals, which makes it very difficult to perform MRCC (in fact, any MR) calculations when the numbers of active orbitals and electrons become larger, which is often the case when polyradical species and transition metal atoms are involved or truncations used in MRCC and other MR methods are too severe. On the other hand, as further elaborated below, the SRCC methods can be applied to most of the MR situations relevant to chemistry, while being much easier to understand and apply compared to their MR counterparts. Therefore, this dissertation work focuses on SRCC-type

ideas, where the relevant dynamical as well as nondynamical correlation effects are formally recovered through conventional p-h excitations from a single Slater determinant defining the Fermi vacuum.

The SRCC theory uses an exponential ansatz to construct the N-electron wave function, $|\Psi_0\rangle = e^T|\Phi\rangle$, where $|\Phi\rangle$ is the reference determinant (typically, but not necessarily the HF determinant) and $T = \sum_{n=1}^{m_A} T_n$ is the cluster operator, with T_n representing its n-particle-n-hole (np-nh) component and $m_\mathrm{A} \leq N$ designating the truncation level. When $m_\mathrm{A} = N$, the full CC method is obtained which is equivalent to the FCI approach, but for practical reasons m_A is almost always set at a much lower value resulting in the hierarchy of conventional SRCC approximations, such as CC with singles (S) and doubles (D) or CCSD [29–32], where m_A is set to 2, CC with singles (S), doubles (D), and triples (T) or CCSDT [33–35], where $m_\mathrm{A} = 3$, CC with singles (S), doubles (D), triples (T), and quadruples (Q) or CCSDTQ [36–39], where $m_\mathrm{A} = 4$, and so on. One of the interesting facts regarding the CC theory is that the use of an exponential wave operator enables the CC theory to include additional excitations, not explicitly included in the calculations, through product or "disconnected" excitations. In other words, even in low orders like CCSD, where only single and double excitations out of the HF reference are included in the cluster operator T, the various product terms that originate due to the expansion of the exponential operator in a Taylor series lead to some triple, some quadruple, some pentuple, *etc.* excitations out of the reference determinant. The CC formalism can also be extended to excited states with the help of, for example, the equation-of-motion (EOM) CC methodology [40–44] or its linear response (LR) CC [45–49] and symmetry-adapted-cluster (SAC) CI [50] counterparts. In this dissertation research, the focus is on the EOMCC formalism that adopts the ansatz $|\Psi_\mu\rangle = R_\mu e^T|\Phi\rangle$, where $R_\mu = r_{\mu,0}\mathbb{1} + \sum_{n=1}^{m_A} R_{\mu,n}$ is a linear, CI-like, excitation operator, $r_{\mu,0}$ and $R_{\mu,n}$ are the zero- and n-body components of R_μ, and m_A is the level truncation in R_μ, which usually is, but does not have to be [44], the same as that used in T. Similar to the ground-state SRCC theory, different values of m_A result in different truncated EOMCC schemes, such

as EOMCCSD [41–43] when $m_{\rm A} = 2$, EOMCCSDT [51–53] when $m_{\rm A} = 3$, EOMCCSDTQ [54, 55] when $m_{\rm A} = 4$, *etc.* The SRCC framework and its EOMCC extension are very efficient in accounting for dynamical correlations, providing rapid convergence to the exact FCI limit, but the basic CCSD and EOMCCSD approaches have problems in situations involving stronger nondynamical correlations that originate from electronic quasi-degeneracies and MR character of the states of interest. One can deal with the latter situations, by incorporating higher–than–two-body components of the cluster and EOM excitation operators, as is done in the high-level CCSDT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, *etc.*, methods. The "only" problem is that the inclusion of higher–than–two-body components of the T and R_μ operators in the calculations poses another challenge, namely, large, often prohibitive, computational costs that limit the resulting approaches to smaller many-electron systems. For example, the computational time scaling of CCSD/EOMCCSD is $n_o^2 n_u^4$, or $\mathscr{N}^6$, where n_o and n_u are the numbers of correlated occupied and unoccupied orbitals, respectively, and $\mathscr{N}$ is a measure of the system size, which is manageable for systems with about 100 correlated electrons and, with the help of local correlation techniques, for systems with dozens or hundreds of atoms, but CCSDT/EOMCCSDT and CCSDTQ/EOMCCSDTQ scale as $n_o^3 n_u^5$ ($\mathscr{N}^8$) and $n_o^4 n_u^6$ ($\mathscr{N}^{10}$), respectively, limiting their use to systems with a few non-hydrogen atoms. It is, therefore, important to come up with approximations within the CC and EOMCC formalisms that would allow one to incorporate higher–than–two-body components of the cluster and EOM excitation operators at small fractions of the computational costs compared to their full CCSDT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, *etc.*, parents and without much loss of accuracy, while avoiding failures of the perturbative CC/EOMCC approximations of the CCSD(T) [56] and similar types in MR cases and situations characterized by large and nonperturbative correlation effects.

Among the various methodologies developed to tackle the above problems, the completely renormalized (CR) CC/EOMCC [57–70] and the active-space CC/EOMCC [38, 51, 52, 71–80] approaches have become especially valuable. In the former case, one uses the formalism

of the method of moments of CC (MMCC) equations [57–68, 70, 80–82] to add *a posteriori*, noniterative, and state-specific corrections to the energies obtained in the lower-order CC/EOMCC calculations. For example, in the ground-state CR-CC(2,3) [63–66, 68] and CR-CC(2,4) [64, 65, 67, 83, 84] approaches one corrects the CCSD energies for the correlation effects due to the connected triply, in the former case, and triply and quadruply excited clusters, in the latter case, which are disregarded in the initial CCSD calculation. In the case of the excited-states, the CR-EOMCC methods, such as CR-EOMCC(2,3) [65, 68] and δ-CR-EOMCC(2,3) [69], correct the total or vertical excitation energies obtained with EOMCCSD. The CR-CC/EOMCC approaches can be very accurate, eliminating failures of perturbative methods, such as CCSD(T) or EOMCCSD(T) [85, 86], in the examination of bond breaking, excited states dominated by two-electron transitions, and other MR situations, but they do not allow for the relaxation of the lower-order cluster and EOM excitation operator components, such as T_1 and T_2 or $R_{\mu,1}$ and $R_{\mu,2}$, in the presence of their higher-order counterparts, such as T_3 or $R_{\mu,3}$, which may lead to substantial inaccuracies in certain reactions involving biradical transition states, singlet–triplet gaps in some biradicals, and certain cases of doubly excited states [87–89]. In the active-space CC/EOMCC methodologies, higher–than–two-body components of the cluster and EOM excitation operators are obtained iteratively, so that the low-order components of T and R_μ can relax in the presence of their higher-order counterparts, which are down-selected using small subsets of orbitals, called active orbitals, to reduce computational costs compared to CCSDT/EOMCCSDT and similar high-level approaches. For example, in CCSDt/EOMCCSDt [38, 51, 52, 72–77, 79] and CCSDtq/EOMCCSDtq [38, 71, 73–76, 79], one considers all singles and doubles, treating T_1, T_2, $R_{\mu,1}$ and $R_{\mu,2}$ exactly, but only relatively small subsets of triples (t) or triples and quadruples (tq), selected using active orbitals are included in the definitions of T_3, T_4, $R_{\mu,3}$, and $R_{\mu,4}$. Although it is true that the inclusion of the leading higher–than–two-body components of T and R_μ defined by active orbitals dramatically reduces the computational costs compared to the parent approaches, such as CCSDT/EOMCCSDT or CCSDTQ/EOMCCS-

DTQ, the active-space CC/EOMCC methods are not free of problems. For example, although the active-space CC/EOMCC approaches allow the lower-rank components of T and R_μ to be relaxed in presence of the higher-rank ones, they miss some important higher-order dynamical correlations because they neglect the higher–than–two-body T and R_μ amplitudes outside the active-space definitions. Fortunately, this can be successfully addressed by merging the CR and active-space CC/EOMCC methodologies with the help of the CC(P;Q) [70, 84, 87, 88] formalism, which is of main significance for the development of new ideas pursued in this dissertation.

In a CC(P;Q) computation, the CC/EOMCC energies obtained in the subspace of the many-electron Hilbert space called the P-space, designated as $\mathcal{H}^{(P)}$, *i.e.*, the energies obtained with conventional (CCSD/EOMCCSD, CCSDT/EOMCCSDT, *etc.*) or unconventional (*e.g.*, CCSDt/EOMCCSDt, CCSDtq/EOMCCSDtq) truncations of the T and R_μ operators, are corrected for the missing correlation effects captured with the help of the complementary subspace of the many-electron Hilbert space called the Q space, denoted as $\mathcal{H}^{(Q)}$, using noniterative moment corrections similar to those used in the CR-CC/EOMCC approaches. The conventional truncations in the iterative P-space CC/EOMCC calculations, combined with similarly conventional definitions of the Q spaces based on the many-body excitation ranks of the determinants included in them, give rise to the aforementioned CR-CC/EOMCC methods, such as CR-CC(2,3), CR-CC(2,4), or CR-EOMCC(2,3), whereas the unconventional truncations based on exploiting the active-space ideas yield the CC(t;3), CC(t,q;3), CC(t,q;3,4), *etc.*, hierarchy, where one first solves the CCSDt [CC(t;3)] or CCSDtq [CC(t,q;3) and CC(t,q;3,4)] equations, which recover much of the nondynamical and some dynamical correlations, while correcting the resulting energies for the missing triples [CC(t;3) and CC(t,q;3)] or missing triples and quadruples [CC(t,q;3,4)]. These methods reproduce the results obtained with their respective parent CC approaches, such as CCSDT, if CC(t;3) is used, or CCSDTQ, if one uses CC(t,q;3) or CC(t,q;3,4), to within small fractions of a millihartree and at small fractions of the computational effort [70, 84, 87–90], but due to

the selection of user- and system-dependent active orbitals, they are no longer black-box approaches. Naturally, to address this concern, the next development efforts should be directed towards finding computationally efficient ways by which the selection of the $\mathcal{H}^{(P)}$ and $\mathcal{H}^{(Q)}$ spaces in the CC(P;Q) calculations is made fully automatic. In this way, we might achieve the desired relaxation of the lower-rank components of T and R_μ in the presence of their higher-rank counterparts without resorting to the user- and system-dependent concepts of the active orbitals exploited in the CC(t;3), CC(t,q;3), CC(t,q;3,4), *etc.*, hierarchy. Clearly, this should be done in such a manner that the resulting CC(P;Q) computations rapidly converge the target CCSDT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, *etc.*, energetics, at small fractions of computational costs of the parent approaches, even when higher–than–two-body components of the T and R_μ operators become large and nonperturbative.

In search of the algorithms enabling an automated determination of the P and Q spaces for CC(P;Q) computations, the Piecuch research group proposed a new class of hybrid CC(P;Q) methods that can loosely be categorized into three different approaches. The first one is based on merging the deterministic CC(P;Q) theory with the stochastic quantum Monte Carlo (QMC) wave function propagations in the many-electron Hilbert space defining the configuration interaction QMC (CIQMC) [91–94] and the CC Monte Carlo (CCMC) [95–98] methods, resulting in the called the semi-stochastic or QMC-driven CC(P;Q) schemes [99–102]. In the second approach, the deterministic CC(P;Q) framework is combined with one of the selected-CI [103–119] methods, such as CIPSI [105, 116–118], resulting in the selected-CI-driven CC(P;Q) methodology [120]. These hybrid CC(P;Q) schemes rely on external sources, such as CIQMC, CCMC, and selected CI, for efficiently identifying the leading determinants in many-electron wave functions in a black-box manner, which are then used to construct the P and Q spaces for the CC(P;Q) calculations. Very recently, a third algorithm has been proposed where the automated design of the P and Q spaces for CC(P;Q) calculations is accomplished using the generalized moments of the CC/EOMCC equations. This algorithm is referred to as the adaptive CC(P;Q) approach [121]. This

dissertation focuses on the semi-stochastic or CIQMC-driven CC(P;Q) framework and one of the main topics of interest is the extension and application of the semi-stochastic CC(P;Q) methodology to ground and excited states of open-shell systems, such as radicals, and other interesting problems involving non-singlet electronic states, such as singlet–triplet gaps in biradicals. However, before discussing the QMC-driven CC(P;Q) methodology, we briefly review the key concepts of QMC, as applied to electronic systems.

The history of QMC methods dates back to the pioneering work of Metropolis and Ulam [122], which paved they way for solving deterministic problems, such as differential equations, via stochastic sampling procedures [122–124]. This resulted in the development of methods such as the variational Monte Carlo (VMC) [125, 126], where the parameters of a trial wave function are optimized by using a stochastic sampling procedure to calculate the expectation values of the Hamiltonian. In another approach, which is commonly referred to as diffusion Monte Carlo (DMC) [127–129], one evolves a trial wave function according to the imaginary-time form of the Schrödinger equation, which results in projecting the exact ground state at the limit of infinite time, (see, *e.g.*, Refs. [130–132] for further details). The DMC methods are easily parallelizable across multiple nodes and operate in the real space of $3N$ particle coordinates, removing any limitations arising from the use of finite basis sets encountered in quantum chemistry, making them very attractive. Unfortunately, in the context of electronic structure theory, they suffer from a major issue: the DMC methods are unaware of the nodal structure of the target wave function and if the wave function propagation is run without any constraints, it projects out the true mathematical ground-state of the spin-free Hamiltonian, which is bosonic in nature, violating the Pauli exclusion principle and resulting in the so-called "boson catastrophe" or "fermion sign problem". A common way to mitigate this is to use the "fixed-node approximation" [133–136], where nodal structures of approximate wave functions obtained in some inexpensive quantum chemistry calculations are imposed as constraints, but, as a result, the exactness of the DMC propagations in the limit of infinite imaginary time is compromised.

In recent years, Full Configuration Interaction Quantum Monte Carlo (FCIQMC) and its truncated variants [91–94] have emerged as methods to circumvent the fermion sign problem inherent in DMC propagations. In these approaches, the imaginary-time propagation of wave functions in real space is substituted by the propagation of CI expansions within the many-electron Hilbert space defined by Slater determinants. FCIQMC and its truncated variants aim to achieve convergence to their respective targets (*e.g.*, FCI for FCIQMC, CISD for CISD-MC, CISDT for CISDT-MC, CISDTQ for CISDTQ-MC, *etc.*) in the limit of infinite imaginary time employing a walker population dynamics algorithm. This algorithm samples the many-electron Hilbert space, allocating more walkers to determinants of higher importance and fewer walkers to those of lesser importance. While the original CIQMC algorithm of Ref. [91] exhibits slow convergence towards its target CI solutions, recent advancements, such as the initiator approximation [92] and its adaptive shift modifications [93, 94], have been introduced to accelerate the convergence of CIQMC. Along with CIQMC, Coupled Cluster Monte Carlo (CCMC) methods have been developed [95–98], substituting linear CI expansions with the exponential CC ansatz when propagating wave functions in the many-electron Hilbert space.

As all stochastic approaches, the CIQMC/CCMC algorithms suffer from large numerical noise in the early stages of wave function propagations, but they are excellent in identifying the leading determinants in the wave functions of interest. This observation forms the basis of the semi-stochastic CC(P;Q) methods, initially proposed in Ref. [99] and further developed in Refs. [100–102] (see, also Ref. [137]), where lists of Slater determinants extracted from CIQMC/CCMC propagations are used to construct the P spaces for the underlying CC(P) and EOMCC(P) computations, and then the remaining determinants relevant to the CC/EOMCC target approach, such as CCSDT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, *etc.*, not captured by CIQMC/CCMC are used to build the Q spaces defining the noniterative CC(P;Q) corrections. The purely stochastic approach to solving the electronic Schrödinger equation using the CIQMC [91–94] and CCMC [95–98] algorithms is plagued by the noise

in energies and properties associated with the varying walker populations and this can only be reduced by using very long propagation times, large walker populations, and additional techniques such as blocking analysis. This limits the applicability of the purely stochastic CIQMC and CCMC calculations to about 40-50 correlated electrons. On the other hand, the semi-stochastic CC(P;Q) formalism of Refs. [99–102], which limits the use of CIQMC or CCMC to short wave function propagations sufficient to the identify the dominant wave function components that are subsequently handled by the deterministic CC(P)/EOMCC(P) computations followed by the CC(P;Q) corrections to capture the remaining information, eliminates the stochastic noise and, as shown in Refs. [99–102], enables us to reproduce the target high-level CC/EOMCC energetics out of the early stages of QMC propagations. The semi-stochastic CC(P;Q) formalism [99–102] and its underlying CC(P)/EOMCC(P) counterpart [99–102, 137] have shown a lot of promise in the calculations for singlet ground and excited states, including PESs along bond breaking coordinates [99–101], but prior to my involvement, nothing has been done for open-shell systems and non-singlet electronic states. In this dissertation, I will discuss our efforts on extending and applying the semi-stochastic CC(P;Q) algorithm to ground and excited states of open-shell systems, such as radicals, and other problems involving non-singlet electronic states, such as singlet–triplet gaps in biradicals [100, 102].

Another way to tackle the challenging problem of open-shell systems, such as radicals and biradicals, within the SRCC framework is to use the formalism of the electron-attachment (EA) [78, 79, 138–142] and ionization-potential (IP) [78, 79, 140–146] EOMCC theories, and their double electron-attachment (DEA) and double ionization-potential (DIP) extensions [80, 147–157]. Here, the wave functions of the ground and excited states of $(N \pm 1)$-electron (EA/IP) and $(N \pm 2)$-electron (DEA/DIP) open-shell systems are constructed using linear EOM-type operators, which attach electrons to or remove electrons from the ground-state CC wave functions of an N-electron closed-shell core. The use of a closed-shell reference state ensures that the EA/IP and DEA/DIP EOMCC wave functions are automatically

orthogonally spin adapted and, thus, preserve the spin symmetry of the Hamiltonian, providing an elegant and powerful way of studying radicals, biradicals, and other systems that differ by one or two electrons from the underlying closed-shell cores. In analogy to the previously discussed particle-conserving EOMCC schemes, where the reference and target systems contain the same number of electrons, the EA/IP and DEA/DIP EOMCC methodologies also suffer from the fact that their lower level approximations, such as EA-EOMCCSD and IP-EOMCCSD, truncated at the 2-particle-1-hole ($2p$-$1h$) and 2-hole-1-particle ($2h$-$1p$) components, respectively, and their DEA/DIP counterparts truncated at $3p$-$1h$ and $3h$-$1p$ terms [147–156] cannot accurately describe the electronic spectra of radicals and biradicals [78–80, 140–142, 153–156]. To obtain accurate spectra of radicals and biradicals, one needs to incorporate $3p$-$2h$/$3h$-$2p$ and $4p$-$2h$/$4h$-$2p$ terms in the respective EA/IP and DEA/DIP EOMCC approaches. Unfortunately, as soon as we incorporate these terms, the calculations become very (usually prohibitively) expensive, limiting their applicability to smaller many-electron systems. In analogy to the particle conserving CC/EOMCC methods, this difficulty can be addressed with the help of user- and system-dependent active orbitals to reduce computational costs [78–80, 140–142, 154–156], but, once again, the question emerges if one could automate the procedure of selecting the higher-rank many-body components of the relevant R_μ-type electron-attaching or ionizing operators, such as the $3p$-$2h$, $3h$-$2p$, $4p$-$2h$, and $4h$-$2p$ terms. Answering this question using the aforementioned stochastic CIQMC ideas within the EA/IP and DEA/DIP EOMCC frameworks has been one of my other method development research projects and is the second main objective of this dissertation. The effectiveness and utility of the semi-stochastic, CIQMC-driven, CC(P;Q) and EA/IP/DEA/DIP EOMCC approaches and computer codes developed in this dissertation research in applications involving ground and excited states of open-shell species and lowest-energy singlet and triplet states of biradical systems will be illustrated through several molecular examples.

CHAPTER 2

BACKGROUND INFORMATION

2.1 Single-Reference Coupled-Cluster Theory and Its Equation-of-Motion Extension to Excited Electronic States

As already mentioned in the Introduction, in the SRCC formalism, the exact ground-state wave function of an N-electron system is expressed as

$$|\Psi_0\rangle = e^T|\Phi\rangle, \tag{2.1}$$

with

$$T = \sum_{n=1}^{N} T_n, \tag{2.2}$$

where $|\Phi\rangle$ is an independent-particle-model reference state that serves as the Fermi vacuum (usually, a HF Slater determinant) and T is the cluster operator. The n-body or np-nh component of the cluster operator is defined as

$$T_n = \sum_{\substack{i_1<\cdots<i_n \\ a_1<\cdots<a_n}} t_{a_1\ldots a_n}^{i_1\ldots i_n} E_{i_1\ldots i_n}^{a_1\ldots a_n}, \tag{2.3}$$

where $t_{a_1\ldots a_n}^{i_1\ldots i_n}$ are the cluster amplitudes and $E_{i_1\ldots i_n}^{a_1\ldots a_n}$ are the elementary np-nh excitation operators defined as $E_{i_1\ldots i_n}^{a_1\ldots a_n} = a^{a_1}\ldots a^{a_n}a_{i_n}\ldots a_{i_1}$, with a^p and a_p representing the usual creation and annihilation operators associated with spin-orbital $|p\rangle$, respectively, which generate excited Slater determinants $|\Phi_{i_1\ldots i_n}^{a_1\ldots a_n}\rangle$ when acting on $|\Phi\rangle$. Here, we use indices $i_1, i_2\ldots$ or $i, j, \ldots$ to denote the occupied spin-orbitals and indices $a_1, a_2, \ldots$ or $a, b, \ldots$ to denote the unoccupied spin-orbitals in the reference determinant $|\Phi\rangle$. In practice, we truncate the many-body expansion defining the cluster operator T at an excitation rank $m_{\rm A} < N$ and different choices of $m_{\rm A}$ result in the conventional CC hierarchy. For example, $m_{\rm A} = 2$ results in the basic CCSD approximation in which T is defined as $T^{(\rm CCSD)} = T_1 + T_2$, $m_{\rm A} = 3$ produces CCSDT [$T^{(\rm CCSDT)} = T_1 + T_2 + T_3$], $m_{\rm A} = 4$ generates the CCSDTQ approach [$T^{(\rm CCSDTQ)} = T_1 + T_2 + T_3 + T_4$], *etc.* To obtain the many-body components T_n and the

energy corresponding to a particular truncation of the cluster operator, denoted here as method A, one first inserts the exponential form of the wave function, Eq. (2.1), into the Schrödinger equation to get its connected-cluster form,

$$\bar{H}|\Phi\rangle = E_0|\Phi\rangle, \tag{2.4}$$

where $\bar{H} = e^{-T}He^{T} = (He^{T})_C$ is the similarity-transformed Hamiltonian and the subscript C designates the connected part of a given operator product. Next, Eq. (2.4) is projected onto the set of excited determinants $|\Phi^{a_1\ldots a_n}_{i_1\ldots i_n}\rangle$ defined by the level of truncation $m_{\rm A} \leq N$ and T is replaced by $T^{(\rm A)} = \sum_{n=1}^{m_{\rm A}} T_n$ to obtain the system of non-linear equations,

$$\langle\Phi^{a_1\ldots a_n}_{i_1\ldots i_n}|\bar{H}^{(\rm A)}|\Phi\rangle = 0, \quad i_1 < \cdots < i_n, \quad a_1 < \cdots < a_n, \quad n = 1, \ldots, m_{\rm A}, \tag{2.5}$$

which are solved iteratively to yield the relevant cluster amplitudes $t^{i_1\ldots i_n}_{a_1\ldots a_n}$. The ground-state SRCC energy, $E_0^{(\rm A)}$, is obtained by calculating the expectation value of the similarity-transformed Hamiltonian $\bar{H}^{(\rm A)} = e^{-T^{(\rm A)}}He^{T^{(\rm A)}} = (He^{-T^{(\rm A)}})_C$ with respect to the Fermi vacuum $|\Phi\rangle$,

$$E_0^{(\rm A)} = \langle\Phi|\bar{H}^{(\rm A)}|\Phi\rangle. \tag{2.6}$$

As an illustration, in the basic CCSD approach, where $m_{\rm A} = 2$, one must first solve the non-linear equations $\langle\Phi^a_i|\bar{H}^{(\rm CCSD)}|\Phi\rangle = 0$ and $\langle\Phi^{ab}_{ij}|\bar{H}^{(\rm CCSD)}|\Phi\rangle = 0$, where $\bar{H}^{(\rm CCSD)} = e^{-T_1-T_2}He^{T_1+T_2} = (He^{T_1+T_2})_C$, to obtain the cluster amplitudes t^i_a and t^{ij}_{ab} and, subsequently, the CCSD energy is computed as $E_0^{(\rm CCSD)} = \langle\Phi|\bar{H}^{(\rm CCSD)}|\Phi\rangle$.

The SRCC theory can be extended to excited states using the EOM formalism, where one applies a linear, CI-like, excitation operator R_μ to the CC ground state defined by Eq. (2.1). Hence, the EOMCC ansatz for the wave function of the μ^{th} excited state, $|\Psi_\mu\rangle$, is

$$|\Psi_\mu\rangle = R_\mu|\Psi_0\rangle = R_\mu e^{T}|\Phi\rangle, \tag{2.7}$$

where

$$R_\mu = r_{\mu,0}\mathbb{1} + \sum_{n=1}^{N} R_{\mu,n}, \tag{2.8}$$

with

$$R_{\mu,n} = \sum_{\substack{i_1<\cdots<i_n \\ a_1<\cdots<a_n}} r^{i_1\ldots i_n}_{\mu,a_1\ldots a_n} E^{a_1\ldots a_n}_{i_1\ldots i_n} \tag{2.9}$$

representing the n-body (np-nh) component of R_μ and $\mathbb{1}$ designating the unit operator (we use the convention in which $\mu = 0$ represents the ground state and $\mu > 0$ designates excited states). Typically, we truncate the excitation operator R_μ at the same level of truncation as the cluster operator T and, as a result, we obtain the hierarchy of EOMCC approximations similar to the ground state problem. Thus, $m_\mathrm{A} = 2$ gives rise to the basic EOMCCSD approximation [$R^{(\mathrm{EOMCCSD})}_\mu = r_{\mu,0}\mathbb{1} + R_{\mu,1} + R_{\mu,2}$], $m_\mathrm{A} = 3$ produces EOMCCSDT [$R^{(\mathrm{EOMCCSDT})}_\mu = r_{\mu,0}\mathbb{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3}$], $m_\mathrm{A} = 4$ results in the EOMCCSDTQ approach [$R^{(\mathrm{EOMCCSDTQ})}_\mu = r_{\mu,0}\mathbb{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3} + R_{\mu,4}$], *etc.* To obtain the equations for the excitation amplitudes $r^{i_1\ldots i_n}_{\mu,a_1\ldots a_n}$ and energies $E^{(A)}_\mu$, one inserts Eq. (2.7) into the Schrödinger equation and projects the resulting equation onto the set of excited Slater determinants $|\Phi^{a_1\ldots a_n}_{i_1\ldots i_n}\rangle = E^{a_1\ldots a_n}_{i_1\ldots i_n}|\Phi\rangle$ with $m_\mathrm{A} \leq N$, to obtain an eigenvalue problem

$$\langle\Phi^{a_1\cdots a_n}_{i_1\cdots i_n}|(\bar{H}^{(\mathrm{A})}_{\mathrm{open}} R^{(\mathrm{A})}_{\mu,\mathrm{open}})_C|\Phi\rangle = \omega^{(\mathrm{A})}_\mu r^{i_1\cdots i_n}_{\mu,a_1\cdots a_n}, \tag{2.10}$$

$$i_1 < \cdots < i_n, \quad a_1 < \cdots < a_n, \quad n = 1, \ldots, m_\mathrm{A},$$

where $\bar{H}^{(\mathrm{A})}_{\mathrm{open}} = \bar{H}^{(\mathrm{A})} - \bar{H}^{(\mathrm{A})}_{\mathrm{closed}} = \bar{H}^{(\mathrm{A})} - E^{(\mathrm{A})}_0\mathbb{1}$ and $R^{(\mathrm{A})}_{\mu,\mathrm{open}} = R^{(\mathrm{A})}_\mu - r_{\mu,0}\mathbb{1}$ are the open parts of $\bar{H}^{(\mathrm{A})}$ and $R^{(\mathrm{A})}_\mu$ having external Fermion lines and $\omega^{(\mathrm{A})}_\mu = E^{(\mathrm{A})}_\mu - E^{(\mathrm{A})}_0$ is the vertical excitation energy corresponding to the excited state. After solving Eq. (2.10) (typically, using the iterative Davidson-type diagonalization procedure generalized to non-Hermitian eigenvalue problems) for amplitudes $r^{i_1\ldots i_n}_{\mu,a_1\ldots a_n}$ and excitation energies $\omega^{(\mathrm{A})}_\mu$, we determine the coefficient $r_{\mu,0}$ entering R_μ, which is nonzero for excited states of the same symmetry as the ground state, by using the expression

$$r_{\mu,0} = \langle\Phi|(\bar{H}^{(\mathrm{A})}_{\mathrm{open}} R^{(\mathrm{A})}_{\mu,\mathrm{open}})_C|\Phi\rangle/\omega^{(\mathrm{A})}_\mu. \tag{2.11}$$

It is important to point out that the similarity-transformed Hamiltonian $\bar{H}^{(\mathrm{A})}$ is not Hermitian. This implies that for the CC and EOMCC theories, the "bra" and "ket" eigen-

vectors of $\bar{H}^{(\mathrm{A})}$ corresponding to a given eigenvalue are completely different. Therefore, we have to distinguish between the right or ket CC and EOMCC states defined by Eqs. 2.1 and 2.7, respectively, and their left or bra counterparts. The appropriate ansatz for the left CC ($\mu = 0$) and EOMCC ($\mu > 0$) states matching their $|\Psi_0\rangle$ and $|\Psi_\mu\rangle$ is as follows:

$$\langle\tilde{\Psi}_\mu| = \langle\Phi|L_\mu e^{-T}, \tag{2.12}$$

where L_μ is a hole-particle de-excitation operator that satisfies the biorthonormality condition

$$\langle\tilde{\Psi}_\mu|\Psi_\nu\rangle = \langle\Phi|L_\mu R_\nu|\Phi\rangle = \delta_{\mu\nu}. \tag{2.13}$$

Here, $\delta_{\mu\nu}$ is the Kronecker delta and the de-excitation operator L_μ is defined as

$$L_\mu = \delta_{\mu 0}\mathbb{1} + \sum_{n=1}^{N} L_{\mu,n}, \tag{2.14}$$

where

$$L_{\mu,n} = \sum_{\substack{i_1<\cdots<i_n \\ a_1<\cdots<a_n}} l^{a_1\ldots a_n}_{\mu,i_1\ldots i_n}(E^{a_1\ldots a_n}_{i_1\ldots i_n})^\dagger \tag{2.15}$$

is the n-body (np-nh) component of L_μ. In analogy to T and R_μ, we approximate L_μ by $L^{(\mathrm{A})}_\mu$, in which the summation over n is truncated at the many-body rank m_A, and determine the amplitudes $l^{a_1\ldots a_n}_{\mu,i_1\ldots i_n}$ by solving the left eigenvalue problem

$$\begin{gathered}\delta_{\mu 0}\langle\Phi|\bar{H}^{(\mathrm{A})}_{\mathrm{open}}|\Phi^{a_1\ldots a_n}_{\mu,i_1\ldots i_n}\rangle + \langle\Phi|L^{(\mathrm{A})}_{\mu,\mathrm{open}}\bar{H}^{(\mathrm{A})}_{\mathrm{open}}|\Phi^{a_1\ldots a_n}_{\mu,i_1\ldots i_n}\rangle = \omega^{(\mathrm{A})}_\mu l^{a_1\ldots a_n}_{\mu,i_1\ldots i_n}, \\ i_1 < \cdots < i_n, \quad a_1 < \cdots < a_n, \quad n = 1, \ldots, m_\mathrm{A}.\end{gathered} \tag{2.16}$$

This left eigenvalue problem is crucial because it allows one to compute properties other than energy within the CC and EOMCC formalisms. As it turns out, it is also essential in defining robust noniterative energy corrections to lower-level CC/EOMCC theories, used by the CR-CC(2,3), CR-EOMCC(2,3), and other biorthogonal CR-CC/EOMCC approaches and their CC(P;Q) generalization, which we discuss next.

2.2 The CC($P;Q$) Formalism

The CC($P;Q$) framework [70, 84, 87, 88], a generalization of the biorthogonal MMCC [63–65, 68, 70] formalism to unconventional truncations in T and R_μ, comprises two steps. In the first step, abbreviated as CC(P) for the ground ($\mu = 0$) state and EOMCC(P) for excited ($\mu > 0$) states, one solves the CC/EOMCC equations in the subspace of the N-electron Hilbert space referred to as the P space, denoted as $\mathcal{H}^{(P)}$. This space is spanned by the excited determinants $|\Phi_K\rangle = E_K|\Phi\rangle$ which, together with the reference determinant $|\Phi\rangle$, dominate the ground- and excited-state wave functions $|\Psi_\mu\rangle$ of interest (E_K is the elementary particle–hole excitation operator generating $|\Phi_K\rangle$ from $|\Phi\rangle$; for the sake of brevity of this description, we assume that ground and excited states have the same symmetry; excited states having different symmetries than the ground state are addressed later). This is done in a usual way adopted in all single-reference CC and EOMCC calculations, *i.e.*, one starts with obtaining the cluster operator

$$T^{(P)} = \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} t_K E_K, \tag{2.17}$$

with t_K representing the corresponding cluster amplitudes by solving the system of equations

$$\langle \Phi_K|\bar{H}^{(P)}|\Phi\rangle = 0, \quad |\Phi_K\rangle \in \mathcal{H}^{(P)}, \tag{2.18}$$

where $\bar{H}^{(P)} = e^{-T^{(P)}} H e^{T^{(P)}} = (He^{T^{(P)}})_C$ is the relevant similarity-transformed Hamiltonian, and determining the corresponding ground-state energy

$$E_0^{(P)} = \langle \Phi|\bar{H}^{(P)}|\Phi\rangle. \tag{2.19}$$

Then, the similarity-transformed Hamiltonian $\bar{H}^{(P)}$ is diagonalized in the P space $\mathcal{H}^{(P)}$ to determine the excited-state EOMCC(P) energies $E_\mu^{(P)}$ and the corresponding EOM excitation and de-excitation operators,

$$R_\mu^{(P)} = r_{\mu,0}\mathbb{1} + \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} r_{\mu,K} E_K \tag{2.20}$$

and

$$L_{\mu}^{(P)} = \delta_{\mu 0}\mathbb{1} + \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} l_{\mu,K}(E_K)^{\dagger}, \tag{2.21}$$

respectively, where $r_{\mu,K}$ and $l_{\mu,K}$ designate the relevant amplitudes, which define the EOMCC(P) ket states

$$|\Psi_{\mu}^{(P)}\rangle = R_{\mu}^{(P)} e^{T^{(P)}} |\Phi\rangle \tag{2.22}$$

and the CC(P)/EOMCC(P) bra states

$$\langle\tilde{\Psi}_{\mu}^{(P)}| = \langle\Phi| L_{\mu}^{(P)} e^{-T^{(P)}} \tag{2.23}$$

satisfying $\langle\tilde{\Psi}_{\mu}^{(P)}|\Psi_{\nu}^{(P)}\rangle = \delta_{\mu\nu}$. Once all of this is done, one proceeds to the second step, which is the calculation of the noniterative corrections $\delta_{\mu}(P;Q)$ to the CC(P) and EOMCC(P) energies $E_{\mu}^{(P)}$ that account for the remaining many-electron correlation effects of interest captured with the help of the another subspace of the N-electron Hilbert space, referred to as the Q space and designated as $\mathcal{H}^{(Q)}$ [$\mathcal{H}^{(Q)} \subseteq (\mathcal{H}^{(0)} \oplus \mathcal{H}^{(P)})^{\perp}$, where $\mathcal{H}^{(0)}$ is a one-dimensional subspace spanned by $|\Phi\rangle$]. The expression for these corrections is

$$\delta_{\mu}(P;Q) = \sum_{|\Phi_K\rangle \in \mathcal{H}^{(Q)}} \ell_{\mu,K}(P) \mathfrak{M}_{\mu,K}(P), \tag{2.24}$$

where

$$\mathfrak{M}_{0,K}(P) = \langle\Phi_K|\bar{H}^{(P)}|\Phi\rangle, \tag{2.25}$$

when $\mu = 0$, and

$$\mathfrak{M}_{\mu,K}(P) = \langle\Phi_K|\bar{H}^{(P)} R_{\mu}^{(P)}|\Phi\rangle, \tag{2.26}$$

when $\mu > 0$, are the generalized moments of the CC(P) and EOMCC(P) equations that correspond to projections of the P space CC/EOMCC equations on to the complementary Q-space determinants $|\Phi_K\rangle \in \mathcal{H}^{(Q)}$ and the coefficients $\ell_{\mu,K}(P)$ multiplying moments $\mathfrak{M}_{\mu,K}(P)$ in Eq. (2.24) are calculated with the expression

$$\ell_{\mu,K}(P) = \langle\Phi| L_{\mu}^{(P)} \bar{H}^{(P)}|\Phi_K\rangle / D_{\mu,K}(P), \tag{2.27}$$

in which the $D_{\mu,K}(P)$ denominators are given by

$$D_{\mu,K}(P) = E_{\mu}^{(P)} - \langle \Phi_K | \bar{H}^{(P)} | \Phi_K \rangle \tag{2.28}$$

(one could replace the Epstein–Nesbet $D_{\mu,K}(P)$ denominators entering Eq. (2.27) $\ell_{\mu,K}(P)$ by their Møller–Plesset analogs, but, as shown in the past, for example in Refs. [63, 64, 66, 68, 84, 88, 101], the Epstein–Nesbet form is generally more effective). The final CC(P;Q) electronic energies for the ground ($\mu = 0$) and excited ($\mu > 0$) states are determined as

$$E_{\mu}^{(P+Q)} = E_{\mu}^{(P)} + \delta_{\mu}(P;Q). \tag{2.29}$$

Now, a question arises as to how to define the P and Q spaces entering the CC(P;Q) framework to obtain accurate ground- and excited-state energetics that match the quality of the target high-level CC/EOMCC calculations without incurring the substantial computational costs of methods, such as CCSDT/EOMCCSDT, CCSDTQ/EOMCCSDTQ, *etc.* The simplest possibility is to rely on the conventional choices, where the P and Q spaces are defined based on the many-body ranks of the excited determinants $|\Phi_K\rangle$ included in them. For example, if we want to correct the CCSD and EOMCCSD energies for triples using the CC(P;Q) formulas, the P space $\mathscr{H}^{(P)}$ is spanned by the singly and doubly excited determinants $|\Phi_i^a\rangle$ and $|\Phi_{ij}^{ab}\rangle$, respectively, and the Q space $\mathscr{H}^{(Q)}$ by the triply excited determinants $|\Phi_{ijk}^{abc}\rangle$. As already alluded to above, the resulting CR-CC(2,3) and CR-EOMCC(2,3) corrections to the CCSD and EOMCCSD energies have been very successful, but, by decoupling the low-order T_n and $R_{\mu,n}$ components with $n \leq 2$ from their higher-order, such as T_3 and $R_{\mu,3}$ counterparts, they may not be as accurate as desired, for example in situations where T_3 and $R_{\mu,3}$ components become large, resulting in substantial inaccuracies and difficulties in balancing ground- and excited-state energies. One can address this problem by using active orbitals to enrich the relevant P spaces with the dominant higher–than–doubly excited determinants, as in the CC(t;3), CC(t,q;3), CC(t,q;3,4), *etc.*, methodologies mentioned in the Introduction, but the resulting approaches are no longer computational black boxes. The semi-stochastic CC(P;Q) approach to ground- and excited-state calculations [99–102],

which is described in detail in Chapter 3, exploits the CIQMC or CCMC propagations to identify the leading higher–than–doubly excited determinants pertinent to the CC/EOMCC calculations of interest in a black-box manner, while using corrections $\delta_\mu(P;Q)$ to capture the remaining correlations that the CC(P)/EOMCC(P) energies at a given QMC propagation time do not describe, eliminates the above concerns. Since one of the objectives of this dissertation project is to extend the semi-stochastic CC($P;Q$) ideas to the EA/IP/DEA/DIP EOMCC methods, these particle nonconserving approaches are discussed next.

2.3 Particle Nonconserving Equation-of-Motion Coupled-Cluster Theories: The Single and Double Electron Attachment and Ionization Potential Methods

Radicals, biradicals, and other open-shell species and their low-lying excited states constitute a major challenge to CC and other *ab initio* methods due to, in most cases, their inherent multiconfigurational character that is difficult to capture using the low-rank approximations, such as CCSD, EOMCCSD, and their linear response analogs. One can always think of using genuine MRCC and other MR methodologies, but as pointed out in the Introduction, their routine use is not straightforward and they have their own, often severe, challenges. The failures of CCSD/EOMCCSD in the presence of electronic quasi-degeneracies in open-shell systems imply the need to incorporate higher–than–double excitations in the CC/EOMCC wave functions, but this results in several computational and formal difficulties. As already discussed, the full inclusion of higher–than–two-body components in the cluster and EOM excitation operators improves the results, but the associated computer costs become very (often prohibitively) large as we increase the system size. One can think of resorting to the active-space ideas, giving rise to methods such as CCSDt/EOMCCSDt, CCSDtq/EOMCCSDtq *etc.*, but with the usual spin-integrated spin-orbital formulation, neither these methods nor any other particle conserving CC/EOMCC methodology properly account for the spin symmetry of the calculated states of open-shell systems. There is, however, an interesting, underappreciated, alternative discussed in this section and pursued in my doctoral work. One can deal with the above problems by utilizing the flexibility of the EOMCC formal-

ism and resorting to the particle-nonconserving EA/IP and DEA/DIP EOMCC frameworks. These approaches provide an elegant way of obtaining orthogonally spin-adapted results for the ground and excited states of radicals and biradicals by diagonalizing the similarity-transformed Hamiltonian of a closed-shell CC theory in the subspaces of the Fock space obtained by adding one or two electrons to (EA, DEA) or removing one or two electrons from (IP, DIP) a related closed-shell core.

In these kinds of methods, one starts with a CC ground-state wave function $|\Psi_0\rangle = e^T|\Phi\rangle$ of an N-electron closed-shell system for which the corresponding correlation energy is defined as

$$\Delta E_0^{(N)} = E_0^{(N)} - \langle\Phi|H|\Phi\rangle = (H_N e^T)_{C,\text{closed}}, \tag{2.30}$$

where $\langle\Phi|H|\Phi\rangle$ is the reference energy, $H_N = H - \langle\Phi|H|\Phi\rangle$ is the Hamiltonian operator in the normal ordered form relative to the N-electron Fermi vacuum $|\Phi\rangle$, the superscript (N) indicates the number of electrons in the closed-shell system, and T is the corresponding cluster operator. In the EA and IP-EOMCC theories, the similarity-transformed Hamiltonian,

$$\bar{H}_{N,\text{open}} = (H_N e^T)_{C,\text{open}} = e^{-T} H_N e^T - \Delta E_0^{(N)}, \tag{2.31}$$

corresponding to the N-electron closed-shell CC theory, is diagonalized in the $(N+1)$-electron (EA) and $(N-1)$-electron (IP) subspaces, $\mathscr{H}^{(N+1)}$ and $\mathscr{H}^{(N-1)}$, respectively, of the Fock space. In doing so, the ground and excited-state wave functions of the target $(N+1)$- and $(N-1)$-electron systems of interest are defined as

$$|\Psi_\mu^{(N\pm1)}\rangle = R_\mu^{(N\pm1)}|\Psi_0^{(N)}\rangle = R_\mu^{(N\pm1)} e^T|\Phi\rangle, \tag{2.32}$$

where the cluster operator T is obtained by solving the usual ground-state CC equations for the N-electron closed-shell system. $R_\mu^{(N+1)}$ is an electron-attaching operator defined as

$$R_\mu^{(N+1)} = \sum_{n=0}^{M_R} R_{\mu,(n+1)p\text{-}nh}, \tag{2.33}$$

where

$$R_{\mu,(n+1)p\text{-}nh} = \sum_{\substack{i_1>\cdots>i_n \\ a<a_1<\cdots<a_n}} r^{i_1\ldots i_n}_{\mu,aa_1\ldots a_n} E^{aa_1\ldots a_n}_{i_1\ldots i_n} (n \geq 1). \tag{2.34}$$

$R^{(N-1)}_\mu$ is an electron ionizing operator defined as

$$R^{(N-1)}_\mu = \sum_{n=0}^{M_R} R_{\mu,(n+1)h\text{-}np}, \tag{2.35}$$

with

$$R_{\mu,(n+1)h\text{-}np} = \sum_{\substack{i>i_1>\cdots i_n \\ a_1<\cdots<a_n}} r^{ii_1\ldots i_n}_{\mu,a_1\ldots a_n} E^{a_1\ldots a_n}_{ii_1\ldots i_n} (n \geq 1). \tag{2.36}$$

Here, $M_R = N$ in the exact theory and $M_R < N$ in approximate EA/IP EOMCC schemes. Similarly to the particle-conserving EOMCC calculations, we insert Eq. (2.32) into the Schrödinger equation to obtain the EOMCC-type eigenvalue problem for the $|\Psi^{(N\pm1)}_\mu$ states or the $R^{(N\pm1)}_\mu$ operators that define them,

$$(\bar{H}_{N,\text{open}} R^{(N\pm1)}_\mu)_C |\Phi\rangle = \omega^{(N\pm1)}_\mu R^{(N\pm1)}_\mu |\Phi\rangle, \tag{2.37}$$

where

$$\omega^{(N\pm1)}_\mu = E^{(N\pm1)}_\mu - E^{(N)}_0 \tag{2.38}$$

are the corresponding electron attaching ($\omega^{(N+1)}_\mu$) or electron ionizing ($\omega^{(N-1)}_\mu$) energies. After solving this eigenvalue problem, one obtains the electron-attachment or ionization energies directly. As the similarity-transformed Hamiltonian, $\bar{H}_{N,\text{open}}$, of the underlying closed-shell N-electron system commutes with S^2 and S_z spin operators, the EA-EOMCC and IP-EOMCC wave functions obtained by diagonalizing $\bar{H}_{N,\text{open}}$ in the appropriate subspaces of $\mathscr{H}^{(N+1)}$ and $\mathscr{H}^{(N-1)}$ are automatically the eigenfunctions of both the S^2 and S_z operators and thus free from the spin-contamination issues present in the usual open-shell EOMCC schemes employing unrestricted HF or restricted open-shell HF (ROHF) references.

The most basic approximations in the EA- and IP-EOMCC categories are the EA-EOMCCSD or EA-EOMCC(2p-1h) and IP-EOMCCSD or IP-EOMCC(2h-1p) approaches, where the cluster operator

$$T = T_1 + T_2 \tag{2.39}$$

is obtained by solving the CCSD equations for the N-electron closed-shell system, the electron attaching operator $R_\mu^{(N+1)}$ defining the (N+1)-electron target states $|\Psi_\mu^{(N+1)}\rangle$ is given by

$$R_\mu^{(N+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} = \sum_a r_{\mu,a} a^a + \sum_{j,a<b} r^j_{\mu,ab} a^a a^b a_j, \tag{2.40}$$

and the ionizing operator $R_\mu^{(N-1)}$ needed to generate the (N-1)-electron states $|\Psi_\mu^{(N-1)}\rangle$ is defined as

$$R_\mu^{(N-1)} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} = \sum_i r^i_\mu a_i + \sum_{i>j,b} r^{ij}_{\mu,b} a^b a_j a_i. \tag{2.41}$$

In these two cases, we end up with diagonalizing the similarity-transformed Hamiltonian of CCSD, $(H_N e^{T_1+T_2})_{C,\text{open}}$, in the space spanned by $|\Phi^a\rangle = a^a|\Phi\rangle$ and $|\Phi^{ab}_j\rangle = a^a a^b a_j|\Phi\rangle$ determinants for EA-EOMCC(2p-1h) and the space spanned by the $|\Phi_i\rangle = a_i|\Phi\rangle$ and $|\Phi^b_{ij}\rangle = a^b a_j a_i|\Phi\rangle$ determinants when the IP-EOMCC(2h-1p) calculations are performed. Although the EA-EOMCC(2p-1h) and IP-EOMCC(2h-1p) methods can be useful in obtaining the lowest electron affinities and ionization potentials, they struggle to describe electron attachment and ionization energies corresponding to excited states of anions and cations, what is particularly relevant for this dissertation, the electronic excitations in radicals, especially when the target states have inherently multiconfigurational character. The good news is that one can usually address this issue by incorporating the higher-rank 3p-2h/3h-2p components in the $R_\mu^{(N\pm1)}$ operator expressions *i.e.*, by using the

$$R_\mu^{(N+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h}, \tag{2.42}$$

and

$$R_\mu^{(N-1)} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p} \tag{2.43}$$

operators in the EA/IP-EOMCC computations. Now, there are two options: either one uses Eqs. (2.42) and (2.43) for the electron attaching and ionizing operators and the similarity-transformed Hamiltonian of CCSD, $\bar{H}^{(\mathrm{CCSD})}_{N,\mathrm{open}} = (H_N e^{T_1+T_2})_{C,\mathrm{open}}$, resulting in the EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p) approaches, respectively, or, one keeps the same definitions of the $R^{(N\pm1)}_\mu$ operators as in Eqs. (2.42) and (2.43), but replaces $\bar{H}^{(\mathrm{CCSD})}_{N,\mathrm{open}}$ by the similarity-transformed Hamiltonian of the CCSDT theory, $\bar{H}^{(\mathrm{CCSDT})}_{N,\mathrm{open}} = (H_N e^{T_1+T_2+T_3})_{C,\mathrm{open}}$, resulting in the EA-EOMCCSDT and IP-EOMCCSDT methods. Both classes of methods are similarly effective at describing the electronic spectra of the $(N \pm 1)$ electron radical systems, so we will focus on EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p), which do not require solving the CCSDT equations, but they may still be too expensive in practice due to the relatively high computational costs of diagonalizing $\bar{H}^{(\mathrm{CCSD})}_{N,\mathrm{open}}$ in spaces containing $|\Phi^{abc}_{jk}\rangle = a^a a^b a^c a_k a_j |\Phi\rangle$ and $|\Phi^{bc}_{ijk}\rangle = a^b a^c a_k a_j a_i |\Phi\rangle$ determinants that scale as $\mathscr{N}^7$ with the system size.

One way to deal with the problem of high costs of EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p) computations is to resort to the active-space ideas, similar to the ground-state CC and excited-state EOMCC cases, by including only the leading components of the $R_{\mu,3p\text{-}2h}/R_{\mu,3h\text{-}2p}$ operators defined by active orbitals. These active-space components, defined in Refs. [78, 79, 140–142], will be denoted as $r_{\mu,3p\text{-}2h}/r_{\mu,3h\text{-}2p}$. Therefore, one incorporates all singly and doubly excited amplitudes of the cluster operator T_1 and T_2, all attaching $R_{\mu,1p}$ and $R_{\mu,2p\text{-}1h}$ amplitudes, and all electron ionizing $R_{\mu,1h}$ and $R_{\mu,2h\text{-}1p}$ amplitudes, but only small subsets of 3p-2h/3h-2p amplitudes of $R^{(N\pm1)}_\mu$ defined by active orbitals, giving rise to the active-space EA/IP-EOMCC methods known as EA-EOMCCSDt and IP-EOMCCSDt or EA-EOMCC(3p-2h)$\{N_u\}$ and IP-EOMCC(3h-2p)$\{N_o\}$, where N_o and N_u are the numbers of active occupied and active unoccupied orbitals in $|\Phi\rangle$, respectively [78, 79, 140–142]. In the EA-EOMCCSDt = EA-EOMCC(3p-2h)$\{N_u\}$ approach, the cluster operator T of the N-electron closed-shell core obtained with CCSD is defined as

$$T = T_1 + T_2 \tag{2.44}$$

and the electron attaching operator given by Eq. (2.42) is replaced by

$$R_\mu^{(N+1)}\{N_u\} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} + r_{\mu,3p\text{-}2h}, \tag{2.45}$$

where

$$r_{\mu,3p\text{-}2h} = \sum_{\substack{j>k \\ \mathbf{A}<b<c}} r^{jk}_{\mu,\mathbf{A}bc} a^{\mathbf{A}} a^b a^c a_k a_j. \tag{2.46}$$

In IP-EOMCCSDt = IP-EOMCC(3h-2p)$\{N_o\}$, the same definition of the cluster operator is used and the ionizing operator given by Eq. (2.43) is replaced by

$$R_\mu^{(N-1)}\{N_o\} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} + r_{\mu,3h\text{-}2p}, \tag{2.47}$$

with

$$r_{\mu,3h\text{-}2p} = \sum_{\substack{\mathbf{I}>j>k \\ b<c}} r^{\mathbf{I}jk}_{\mu,bc} a^b a^c a_k a_j a_{\mathbf{I}}. \tag{2.48}$$

The bold capital indices in the above expressions for $r_{\mu,3p\text{-}2h}$ and $r_{\mu,3h\text{-}2p}$ denote active spin-orbitals in the respective categories (**I** for occupied and **A** for unoccupied). The EA-EOMCCSDt = EA-EOMCC(3p-2h)$\{N_u\}$ and IP-EOMCCSDt = IP-EOMCC(3h-2p)$\{N_o\}$ approaches are computationally much more affordable than their EA-EOMCC(3p-2h)/IP-EOMCC(3h-2p) and EA-EOMCCSDT/IP-EOMCCSDT counterparts. They scale as $N_u n_o^2 n_u^4$ and $N_o n_o^2 n_u^4$, respectively, in steps needed to diagonalize the similarity-transformed Hamiltonian, as opposed to $n_o^2 n_u^5$ and $n_o^3 n_u^4$, respectively ($N_o < n_o, N_u << n_u$) when 3p-2h/3h-2p components are treated fully.

The EA/IP-EOMCC approaches can be extended to attach or remove more than one electrons from a suitable closed-shell reference state, since the R_μ operator can be designed to handle multiple electron attachments or removals, leading to the multiple electron attachment or ionization potential EOMCC approaches. For instance, attaching or removing two electrons from a closed-shell reference state resulting from CC calculations, while relaxing the remaining electrons, gives rise to the DEA or DIP EOMCC approaches. In the DEA/DIP-EOMCC schemes, the ground- and excited-state wave functions of the $(N+2)$-electron (the

DEA case) or $(N-2)$-electron (the DIP case) systems, where N is the number of electrons in the underlying closed-shell species, are represented as

$$|\Psi_\mu^{(N\pm2)}\rangle = R_\mu^{(N\pm2)} e^T |\Phi\rangle. \tag{2.49}$$

Here, once again, T is the cluster operator obtained from SRCC calculations for the N-electron reference system (using the N-electron reference determinant $|\Phi\rangle$, which defines the Fermi vacuum),

$$R_\mu^{(N+2)} = \sum_{n=0}^{M_R} R_{\mu,(n+2)p\text{-}nh} \tag{2.50}$$

is the double electron attachment operator, and

$$R_\mu^{(N-2)} = \sum_{n=0}^{M_R} R_{\mu,(n+2)h\text{-}np} \tag{2.51}$$

is the double ionizing operator, where $M_R \leq N$ defines the truncation level. The individual many-body components of the $R_\mu^{(N\pm2)}$ operators are given by the expressions

$$R_{\mu,(n+2)p\text{-}nh} = \sum_{\substack{i_1>\cdots>i_n \\ a<b<a_1<\cdots<a_n}} r^{i_1\ldots i_n}_{\mu,aba_1\ldots a_n} a^a a^b a^{a_1} \ldots a^{a_n} a_{i_n} \ldots a_{i_1}, \tag{2.52}$$

in the DEA case, and

$$R_{\mu,(n+2)h\text{-}np} = \sum_{\substack{i>j>i_1>\cdots>i_n \\ a_1<\cdots<a_n}} r^{iji_1\ldots i_n}_{\mu,a_1\ldots a_n} a^{a_1} \ldots a^{a_n} a_{i_n} \ldots a_{i_1} a_j a_i, \tag{2.53}$$

in the case of the DIP formalism. Similarly to the EA/IP-EOMCC approaches, inserting Eq. (2.49) into the Schrödinger equation yields the corresponding DEA/DIP-EOMCC eigenvalue problem

$$(\bar{H}_{N,\mathrm{open}} R_\mu^{(N\pm2)})_C |\Phi\rangle = \omega_\mu^{(N\pm2)} R_\mu^{(N\pm2)} |\Phi\rangle. \tag{2.54}$$

The DEA-EOMCC and DIP-EOMCC methods are particularly suitable for biradical systems, since they are capable of generating an appropriate multideterminantal reference space by adding or removing two electrons from a closed-shell core, but in order to get quantitative accuracies in most of the situations encountered in chemistry, especially for the singlet

manifolds and singlet–triplet gaps, one needs to incorporate the $R_{\mu,4p\text{-}2h}$ components, in the DEA-EOMCC case, and the $R_{\mu,4h\text{-}2p}$ components, when the DIP-EOMCC methodology is employed, in the corresponding $R_{\mu}^{(N\pm2)}$ operators, giving rise to the DEA-EOMCC($4p$-$2h$) and DIP-EOMCC($4h$-$2p$) approaches, respectively [80, 153–156]. Due to the enormous numbers of the $4p$-$2h$/$4h$-$2p$ amplitudes in the majority of realistic applications, the DEA/DIP-EOMCC approaches with a full treatment of $4p$-$2h$/$4h$-$2p$ correlations are very expensive, but this can again be dealt with by turning to the much more practical active-space DEA-EOMCC($4p$-$2h$) $\{N_u\}$ and DIP-EOMCC($4h$-$2p$) $\{N_o\}$ approaches where the $4p$-$2h$ and $4h$-$2p$ amplitudes are down-selected using active orbitals [80, 153–156]. This can be done by using the following definitions of the $R_{\mu}^{(N\pm2)}$ operators [80, 153–156]:

$$R_{\mu}^{(N+2)}\{N_u\} = R_{\mu,2p} + R_{\mu,3p\text{-}1h} + r_{\mu,4p\text{-}2h}, \tag{2.55}$$

in the DEA-EOMCC($4p$-$2h$)$\{N_u\}$ case, and

$$R_{\mu}^{(N-2)}\{N_o\} = R_{\mu,2h} + R_{\mu,3h\text{-}1p} + r_{\mu,4h\text{-}2p}, \tag{2.56}$$

in the case of the DIP-EOMCC($4h$-$2p$)$\{N_o\}$ approach, where

$$r_{\mu,4p\text{-}2h} = \sum_{\substack{k>l \\ \mathbf{A}<\mathbf{B}<c<d}} r^{kl}_{\mu,\mathbf{AB}cd} a^{\mathbf{A}} a^{\mathbf{B}} a^{c} a^{d} a_{l} a_{k} \tag{2.57}$$

and

$$r_{\mu,4h\text{-}2p} = \sum_{\substack{\mathbf{I}>\mathbf{J}>k>l \\ c<d}} r^{\mathbf{IJ}kl}_{\mu,cd} a^{c} a^{d} a_{l} a_{k} a_{\mathbf{J}} a_{\mathbf{I}}, \tag{2.58}$$

with the capital bold indices representing active spin-orbitals in the respective occupied ($\mathbf{I}$, $\mathbf{J}$) and unoccupied ($\mathbf{A}$, $\mathbf{B}$) categories.

As with all active-space CC/EOMCC methods, the introduction of active orbitals, which are user and system dependent, into Eqs. (2.45)–(2.48) and Eqs. (2.55)–(2.58) means that the resulting EA-EOMCC($3p$-$2h$)$\{N_u\}$, IP-EOMCC($3h$-$2p$)$\{N_o\}$, DEA-EOMCC($4p$-$2h$)$\{N_u\}$, and DIP-EOMCC($4h$-$2p$)$\{N_o\}$ methods are not black-box schemes. The novel semi-stochastic

EA/IP/DEA/DIP-EOMCC(P) approaches to ground and excited states of radical and biradical species, which we describe in Chapter 4, exploit the CIQMC propagations to automatically identify the leading $3p$-$2h$/$3h$-$2p$/$4p$-$2h$/$4h$-$2p$ determinants pertinent to the calculations of interest, and thus eliminate the above concerns. Before introducing these and other semi-stochastic methods in Chapter 3, we provide a brief overview of the CIQMC methodology exploited in such approaches.

2.4 Overview of Configuration Interaction Quantum Monte Carlo

In this section, we summarize the key ingredients of the FCIQMC methodology and its truncated CIQMC counterparts. The FCIQMC approach introduced in Refs. [91, 92] shares similarities with both the deterministic FCI and stochastic diffusion Monte Carlo (DMC) methods. As mentioned in the Introduction, unlike the conventional DMC methodology, FCIQMC does not suffer from the Fermion sign problem [158]. The basic idea of FCIQMC is to propagate the imaginary-time Schrödinger equation written here(in atomic units),

$$\frac{\partial|\Psi\rangle}{\partial\tau} = -H|\Psi\rangle, \tag{2.59}$$

in the many-electron Hilbert space spanned by Slater determinants. It can be shown that by starting from a trial wave function $|\Psi(\tau=0)\rangle$, such as the HF Slater determinant, that has a nonzero overlap with the exact FCI state $|\Psi_0\rangle$, and by applying the energy shift S such that it approaches the FCI energy E_0 as $\tau\to\infty$, in the infinite imaginary-time limit, we project out $|\Psi_0\rangle$,

$$|\Psi_0\rangle = \lim_{\substack{\tau\to\infty \\ S\to E_0}} e^{-\tau(H-S)}|\Psi(\tau=0)\rangle. \tag{2.60}$$

In the many-electron Hilbert space, the propagated wave function $|\Psi(\tau)\rangle = e^{-\tau(H-S)}|\Psi(\tau=0)\rangle$ can be represented as a CI expansion

$$|\Psi(\tau)\rangle = c_0(\tau)|\Phi_0\rangle + \sum_K c_K(\tau)|\Phi_K\rangle, \tag{2.61}$$

where the time-dependent CI coefficients at the Slater determinants $|\Phi_K\rangle$ satisfy the system of coupled differential equations [91]

$$\frac{\partial c_K(\tau)}{\partial \tau} = -(H_{KK} - S)c_K(\tau) - \sum_{L(\neq K)} H_{KL}c_L(\tau), \tag{2.62}$$

where $H_{KL} = \langle\Phi_L|H|\Phi_L\rangle$ are matrix elements of the Hamiltonian involving the various Slater determinants $|\Phi_K\rangle$. It can be readily demonstrated that in the $\tau \to \infty$ limit, with S converging to E_0, the system of equations given by Eq. (2.62) becomes equivalent to the conventional CI eigenvalue problem

$$\sum_L H_{KL}c_L(\infty) = E_0 c_K(\infty). \tag{2.63}$$

Thus, when all possible Slater determinants are allowed in the wave function propagation defined by Eq. (2.62), we obtain results equivalent to FCI. Furthermore, if we confine ourselves to a subset of Slater determinants corresponding to one of the truncated CI methods, such as CISD, CISDT, CISDTQ, *etc.*, the $\tau = \infty$ limit of Eq. (2.62) results in the corresponding truncated CI solution. The main challenge with the system of equations represented by Eq. (2.62) lies in the impracticality of achieving its direct deterministic integration. Thus, in the spirit of DMC, the FCIQMC and truncated CIQMC schemes introduced in Refs. [91] and [92] employ a walker population dynamics to determine the time-dependent CI coefficients $c_K(\tau)$ and, consequently, the energy $E(\tau)$ corresponding to $|\Psi(\tau)\rangle$. The walkers, designated here by α, are fictitious particles that populate Slater determinants and have signs $s_\alpha = \pm 1$, and the CI coefficients $c_K(\tau)$ can be calculated as the signed sums of the walkers associated with particular determinants $|\Phi_K\rangle$, *i.e.*,

$$c_K(\tau) \sim N_K = \sum_\alpha s_\alpha \delta_{KK_\alpha}, \tag{2.64}$$

where $|\Phi_{K_\alpha}\rangle$ designates the determinant on which the walker α is located and s_α is the associated sign. The resulting CIQMC algorithm defining the walker population dynamics equivalent to Eq. (2.62) comprises three main steps that are executed at each time step τ of the imaginary-time propagation:

1. *The spawning step*: For each walker α located on determinant $|\Phi_{K_\alpha}\rangle$, we attempt to spawn a new "child" walker on another determinant $|\Phi_L\rangle$, selected through a normalized probability

$$p_s(L|K_\alpha) = \frac{\Delta\tau|H_{K_\alpha L}|}{p_{gen}(L|K_\alpha)}, \tag{2.65}$$

which is compared to a number between 0 and 1 obtained with the random number generator. In this step, if spawning occurs, the total number of walkers increases. Here, $\Delta\tau$ represents the time step used in the imaginary-time propagation.

2. *The diagonal death/cloning step*: For each walker, we compute the probability

$$p_d(K_\alpha) = \Delta\tau(H_{K_\alpha K_\alpha} - S). \tag{2.66}$$

If $p_d > 0$, the walker is removed from the simulation, and if $p_d < 0$, the walker is cloned with the probability $|p_d|$. The cloning (or birth) step is usually rare.

3. *The annihilation step*: Subsequent to the previous two steps, the annihilation step is performed, wherein all pairs of walkers with opposite sign are removed from each determinant.

A schematic representation of how the CIQMC algorithm works is depicted in Fig. 2.1. After reaching a sufficiently large number of walkers, the correlation energy and the walker populations stabilize using a suitable energy shift S. Since the propagation is performed in the space of Slater determinants, the resulting wave function has the proper Fermionic symmetry. As already alluded to above, if the spawning of walkers is restricted to a certain level of truncation, we can obtain truncated CIQMC approaches, such as CISDT-MC and CISDTQ-MC to mention two representative examples, where spawning of walkers on determinants higher than triples, in the CISDT-MC case, or higher than quadruples in the CISDTQ-MC case, is forbidden. By replacing the time-dependent CI expansion, Eq. (2.61), by the analogous CC ansatz, one obtains the CCMC approach introduced in Ref. [95] (see,

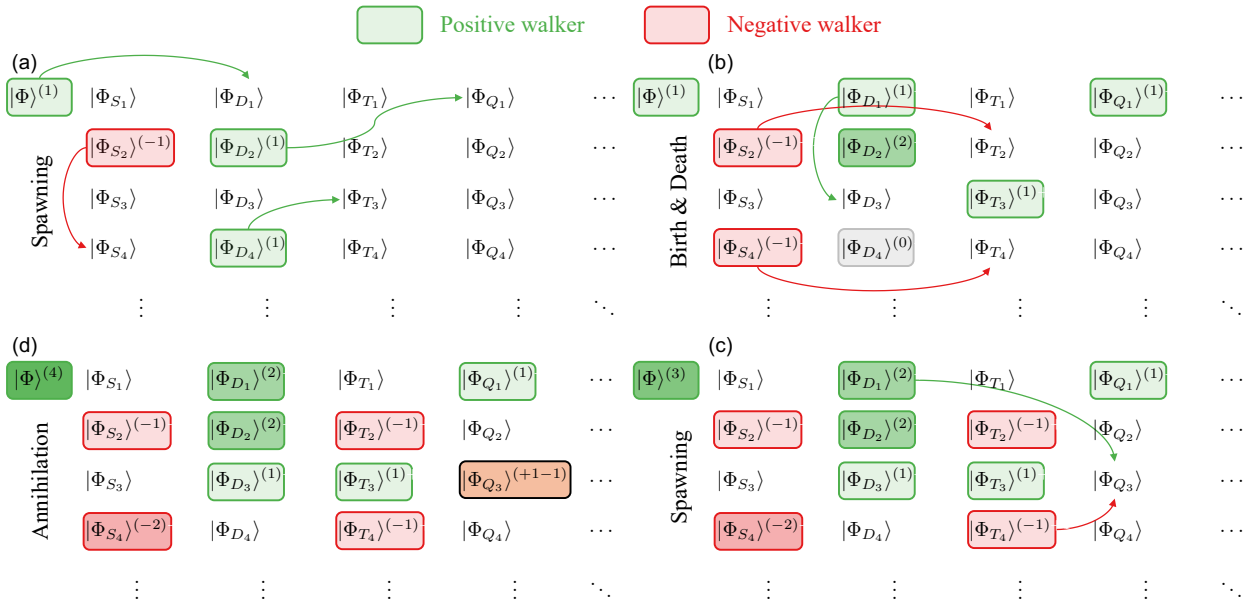

Figure 2.1 A schematic representation of the CIQMC algorithm. $|\Phi\rangle$ represents the reference determinant, while $\{|\Phi_{S_1}\rangle, |\Phi_{S_2}\rangle, |\Phi_{S_3}\rangle, |\Phi_{S_4}\rangle, \cdots\}$ denote singly excited determinants, $\{|\Phi_{D_1}\rangle, |\Phi_{D_2}\rangle, |\Phi_{D_3}\rangle, |\Phi_{D_4}\rangle, \cdots\}$ doubly excited determinants, $\{|\Phi_{T_1}\rangle, |\Phi_{T_2}\rangle, |\Phi_{T_3}\rangle, |\Phi_{T_4}\rangle, \cdots\}$ triples, $\{|\Phi_{Q_1}\rangle, |\Phi_{Q_2}\rangle, |\Phi_{Q_3}\rangle, |\Phi_{Q_4}\rangle, \cdots\}$ quadruples, *etc.* The number of walkers on a particular determinant is indicated as a superscript within parenthesis. Green and red rectangles distinguish positive and negative walkers, respectively, with darker shades indicating a higher number of walkers on the determinants. Here, the simulation starts with one walker placed on the reference determinant $|\Phi\rangle$. Panel (a) illustrates spawning steps, panel (b) depicts the death of a walker on the determinant $|\Phi_{D_4}\rangle$ with a gray box, panel (c) shows more spawning events and panel (d) displays the annihilation step on $|\Phi_{Q_3}\rangle$ with an orange rectangular box.

also, Refs. [96–98]). Several ideas have been explored to improve the convergence of the QMC algorithm. One of them is the initiator CIQMC (i-CIQMC) approach of Ref. [92], which we use in our work, and its i-CCMC counterpart, where spawning is allowed as long as there is at least a minimum walker population, n_a, on the parent determinant. In the next chapters, we discuss how we can take advantage of the CIQMC algorithm to accelerate our CC(P;Q) and EA/IP/DEA/DIP-EOMCC computations, while making them fully automated and free from the user- and system-dependent active orbitals when identifying the appropriate P spaces [CC(P;Q)] or the leading 3p-2h/3h-2p/4p-2h/4h-2p amplitudes.

CHAPTER 3

THE SEMI-STOCHASTIC CC(P;Q) METHODOLOGY FOR GROUND AND EXCITED STATES

In the previous chapters, we have discussed how the deterministic CC(P;Q) theory that combined the CR-CC/EOMCC and active space ideas emerged as a very robust tool for studying challenging chemical systems in a computationally affordable way. The only drawback, however, is the need to choose a suitable set of active orbitals that can generate an ensemble of determinants, which define the proper active space, and this selection is system and user dependent. On the other hand, the QMC methods based on propagating the CI ansatz according to the imaginary time Schrödinger equation in the many electron Hilbert space spanned by Slater determinants are very efficient in identifying the most important determinants, based on probabilistic arguments, in an automated manner, but they spend a lot of time balancing the corresponding coefficients. In this chapter we discuss a novel hybrid approach, namely the semi-stochastic CC(P;Q) formalism, that merges the deterministic CC(P;Q) framework with the stochastic wave function sampling of CIQMC methodologies [99–102]. In particular, we focus on my work on extending the semi-stochastic CC(P;Q) methodology to ground and excited states of open-shell systems and non-singlet excited states. In doing so, we follow the results presented in Refs. [100] and [102], with a focus on my contributions to these papers.

3.1 Theory

The CIQMC algorithm needs a very long propagation time to converge a stable wave function, but the leading determinants are identified much sooner, and on the other hand the deterministic CC(P;Q) framework allows for very accurate energetics even in challenging MR situations as long as the lower clusters, such as T_1 and T_2, are allowed to relax in the presence of the higher order T_3, T_4, *etc.*, clusters. This observation was utilized in Ref. [99], where it was demonstrated that one could use the information about the leading determinants captured during the early stages of *i*-CIQMC to create lists of determinants defining

P spaces for CC(P) calculations and then use the deterministic, noniterative, CC(P;Q) corrections to capture the correlations effects missing in the P-space CC(P) calculations. Later in Ref. [137], the semi-stochastic CC(P) approach was extended to the excited state EOMCC(P) method that provided a fast convergence to target energetics in a very straightforward manner, without resorting to more-complex excited state i-CIQMC frameworks of Refs. [159, 160]. In this work we discuss the extension of the semi-stochastic ground-state CC(P;Q) [99] and EOMCC(P) [137] approaches to open-shell systems and non-singlet excited states and the CC(P;Q) corrections to EOMCCC(P), as described in Refs. [100] and [102].

The key steps of the semi-stochastic CC(P;Q) [99–102, 137] algorithm for ground and excited states that resulted from the merger of deterministic CC(P;Q) algorithm and the stochastic CIQMC wave function sampling, are described below.

1. Initiate a CIQMC run for the ground state and, if the system of interest has spin, spatial, or other symmetries, the analogous QMC propagation for the lowest state of each irreducible representation (irrep) to be considered in the CC(P;Q) calculations by placing a certain number of walkers on the appropriate reference function(s) $|\Phi\rangle$ (*e.g.*, the restricted Hartree–Fock (RHF) or restricted open- shell Hartree–Fock (ROHF) determinants).

2. At some propagation time $\tau > 0$, *i.e.*, after a certain number of CIQMC time steps, called MC iterations, extract a list or, if states belonging to multiple irreps are targeted, lists of determinants relevant to the desired CC(P;Q) computations from the QMC propagation(s) initiated in step 1 to determine the P space or spaces needed to set up the ground-state CC(P) and excited state EOMCC(P) calculations (for example, *cf.* Fig. 3.1). If the goal is to converge the CCSDT/EOMCCSDT-level energetics, the P space for the CC(P) calculations and the EOMCC(P) calculations for excited states belonging to the same irrep as the ground state is defined as all singly and

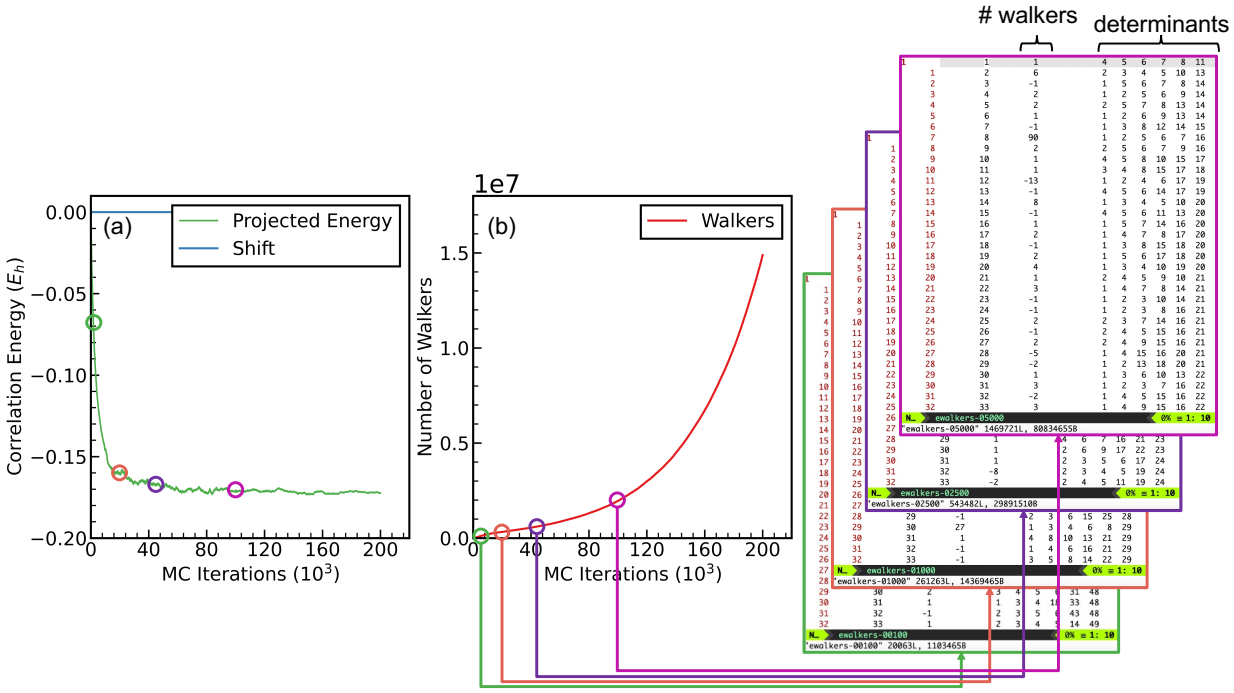

Figure 3.1 A schematic illustration depicting the construction of P-spaces in CC(P) and EOMCC(P) computations. Panel (a) showcases the stabilization of correlation energy (green line) and the corresponding increase in the total number of walkers is shown in panel (b) (red line). On the right four snapshots from a QMC calculation are presented, featuring the lists of determinants picked up by the QMC algorithm at various time steps (green for 2000, orange for 20000, violet for 50000, and magenta for 100000 QMC iterations). It is evident that QMC deems some determinants more important than others by placing more walkers on them.

doubly excited determinants and a subset of triply excited determinants, where each triply excited determinant in the subset is populated by a minimum of n_P positive or negative walkers (in this work, $n_P = 1$). For the excited states belonging to other irreps, the P space defining the CC(P) problem is the same as that used in the case of the ground state, but the lists of triply excited determinants defining the EOMCC(P) diagonalizations are provided by the CIQMC propagations for the lowest-energy states of these irreps. One proceeds in a similar way when the goal is to converge other types of high-level CC/EOMCC energetics. For example, if we want to obtain the results of the CCSDTQ/EOMCCSDTQ quality, we also have to extract the lists of quadruples, in addition to the triples, from the CIQMC runs to define the corresponding P spaces.

3. Solve the CC(P) and EOMCC(P) equations in the P space or spaces obtained in the previous step. If we are targeting the CCSDT/EOMCCSDT-level energetics and the excited states of interest belong to the same irrep as the ground state, we define $T^{(P)} = T_1 + T_2 + T_3^{(\mathrm{MC})}$, $R_\mu^{(P)} = r_{\mu,0}\mathbb{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3}^{(\mathrm{MC})}$, and $L_\mu^{(P)} = \delta_{\mu,0}\mathbb{1} + L_{\mu,1} + L_{\mu,2} + L_{\mu,3}^{(\mathrm{MC})}$, where the list of triples $T_3^{(\mathrm{MC})}$, $R_{\mu,3}^{(\mathrm{MC})}$, and $L_{\mu,3}^{(\mathrm{MC})}$ is extracted from the ground-state CIQMC propagation at time τ. For the excited states belonging to other irreps, we construct the similarity-transformed Hamiltonian $\bar{H}^{(P)}$, to be diagonalized in the EOMCC steps, in the same way as in the ground-state computations, but then use the CIQMC propagations for the lowest states of these irreps to define the lists of triples in $R_{\mu,3}^{(\mathrm{MC})}$ and $L_{\mu,3}^{(\mathrm{MC})}$. We follow a similar procedure when targeting the CCSDTQ/EOMCCSDTQ-level energetics in which case $T^{(P)} = T_1+T_2+T_3^{(\mathrm{MC})}+T_4^{(\mathrm{MC})}$, $R_\mu^{(P)} = r_{\mu,0}\mathbb{1}+R_{\mu,1}+R_{\mu,2}+R_{\mu,3}^{(\mathrm{MC})}+R_{\mu,4}^{(\mathrm{MC})}$, and $L_\mu^{(P)} = \delta_{\mu,0}\mathbb{1}+L_{\mu,1}+L_{\mu,2}+L_{\mu,3}^{(\mathrm{MC})}+L_{\mu,4}^{(\mathrm{MC})}$.

4. Correct the CC(P) and EOMCC(P) energies for the missing correlations of interest that were not captured by the CIQMC propagations at the time τ the lists of the P-space excitations were created (the remaining triples if the goal is to recover the CCSDT/EOMCCSDT energetics, the remaining triples and quadruples if one targets CCSDTQ/EOMCCSDTQ, *etc.*) using the CC(P;Q) corrections $\delta_\mu(P;Q)$ defined by Eq. (2.24).

5. Check the convergence of the resulting $E_\mu^{(P+Q)}$ energies calculated using Eq. (2.29) by repeating steps 2–4 at some later CIQMC propagation time $\tau' < \tau$. If the $E_\mu^{(P+Q)}$ energies do not change within a given convergence threshold, we can stop the calculations. One can also stop them if τ in steps 2–4 is chosen such that the stochastically determined P space(s) contain sufficiently large fraction(s) of higher–than–doubly excited determinants relevant to the target CC/EOMCC level. Our unpublished tests using the CC(P;Q)-based CC(t;3) corrections to the EOMCCSDt energies, the ground-state semi-stochastic CC(P;Q) calculations reported in Ref. [99], and the excited-state

CC(P;Q) calculations using i-FCIQMC to generate the underlying P spaces performed in this work indicate that one should be able to reach millihartree or sub-millihartree accuracies relative to the parent CC/EOMCC computations, when the stochastically determined P spaces contain as little as ∼5–10% and no more than ∼30–40% of higher–than–double excitations of interest, although this may need further study.

Similarly to the semi-stochastic form of the ground-state CC(P;Q) methodology introduced in Ref. [99], the above algorithm offers significant savings in the computational effort compared to the fully deterministic, high-level, EOMCC approaches it targets. These savings originate from three factors. First, the computational times associated with the early stages of the i-CIQMC walker propagations are very short compared to the corresponding converged runs. Second, the CC(P) calculations and the subsequent EOMCC(P) diagonalizations offer significant speedups compared to their CC/EOMCC parents, when the corresponding excitation manifolds contain small fractions of higher–than–doubly excited determinants. For example, as pointed out in Refs. [99, 137], when the most expensive $\langle\Phi^{abc}_{ijk}|[H,T_3]|\Phi\rangle$ (or $\langle\Phi^{abc}_{ijk}|[\bar{H}^{(2)},T_3]|\Phi\rangle$, where $\bar{H}^{(2)}=\exp(-T_1-T_2)H\exp(T_1+T_2)$) and $\langle\Phi^{abc}_{ijk}|[\bar{H}^{(P)},R_{\mu,3}]|\Phi\rangle$ terms in the CCSDT and EOMCCSDT equations are isolated and reprogrammed using techniques similar to implementing selected CI approaches, combined with sparse matrix multiplication and index rearrangement routines (rather than conventional many-body diagrams that assume continuous excitation manifolds labelled by occupied and unoccupied orbitals from the respective ranges of indices; generally, the stochastically determined lists of excitations do not form continuous manifolds that could be *a priori* identified), one can speed up their determination by a factor of up to $(D/d)^2$, where D is the number of all triples and d is the number of triples included in the stochastically determined P space. Other terms, such as $\langle\Phi^{abc}_{ijk}|[H,T_2]|\Phi\rangle$ and $\langle\Phi^{abc}_{ijk}|[\bar{H}^{(P)},R_{\mu,2}]|\Phi\rangle$ or $\langle\Phi^{ab}_{ij}|[H,T_3]|\Phi\rangle$ and $\langle\Phi^{ab}_{ij}|[\bar{H}^{(P)},R_{\mu,3}]|\Phi\rangle$, when treated in a similar manner, may offer additional speedups, on the order of (D/d), too. Our current CC(P) and EOMCC(P) routines are not as efficient yet, but the speedups that scale linearly with (D/d) in the most expensive $\langle\Phi^{abc}_{ijk}|[H,T_3]|\Phi\rangle$ and $\langle\Phi^{abc}_{ijk}|[\bar{H}^{(P)},R_{\mu,3}]|\Phi\rangle$ contri-

butions are attainable. The third factor contributing to major savings in the computational effort offered by the semi-stochastic CC(P;Q) approach is the observation that the determination of the noniterative correction $\delta_\mu(P;Q)$ for a given electronic state μ is much less expensive than the time required to complete a single iteration of the target CC/EOMCC calculation (in the case of the calculations aimed at the CCSDT/EOMCCSDT energetics, the computational time associated with each $\delta_\mu(P;Q)$ scales no worse than $\sim 2n_o^3n_u^4$, as opposed to the $n_o^3n_u^5$ scaling of every CCSDT and EOMCCSDT iteration).

Before going to the next section, we must discuss an interesting aspect of the semi-stochastic CC(P)/EOMCC(P) and CC(P;Q) methodologies. The CC(P) and EOMCC(P) energies at $\tau = 0$ are identical to the energies obtained in the CCSD and EOMCCSD calculations and that the corresponding $\tau = 0$ CC(P;Q) corrections are equivalent to those of CR-CC(2,3) (the ground state) and CR-EOMCC(2,3) (excited states). It should also be noted that the CC(P) and EOMCC(P) energies at $\tau = \infty$ are identical to the energies obtained in the full CCSDT and EOMCCSDT calculations. The semi-stochastic CC(P;Q) calculations recover the CCSDT and EOMCCSDT energetics in this limit too, although the $\tau = \infty$ values of the $\delta_\mu(P;Q)$ corrections are zero in this case, since the $\tau = \infty$ P spaces contain all the triples, *i.e.*, the corresponding Q-space triples lists are empty. These relationships between the semi-stochastic CC(P), EOMCC(P), and CC(P;Q) approaches and the fully deterministic CCSD/EOMCCSD, CR-CC(2,3)/CREOMCC(2,3), and CCSDT/EOMCCSDT methodologies were helpful in examining the correctness of our codes. They also point to the ability of the CC(P), EOMCC(P), and CC(P;Q) calculations driven by the information extracted from CIQMC to offer a systematically improvable description as τ approaches ∞.

3.2 Electronic Excitations in CH^+, CH, and CNC

In order to explore the performance of the semi-stochastic CC(P;Q) approach to excited states proposed in this work and examine, in particular, the ability of the noniterative $\delta_\mu(P;Q)$ corrections to accelerate the convergence of the CIQMC-driven EOMCC(P) calcu-

lations toward the desired EOMCC energetics, represented in this study by EOMCCSDT, we carried out benchmark calculations for the frequently studied vertical excitations in CH^+ ion at the equilibrium [Table 3.1 and Fig. 3.2(a) and (b)] and stretched [Table 3.2 and Fig. 3.2(c) and (d)] geometries, which were previously used to test the EOMCC(P) framework [137] and which was a useful system to check the correctness of our codes, and the adiabatic excitations in the challenging open-shell CH (Table 3.3) and CNC (Table 3.4) systems, which have low lying excited states dominated by two-electron transitions that require at least the EOMCCSDT theory level to obtain a reliable description [62, 68, 78, 79, 141, 161–163]. The CH^+ ion was described by the [5s3p1d/3s1p] basis set of Ref. [164] and we used the aug-cc-pVDZ [165, 166] and DZP[4s2p1d] [167, 168] bases for the CH and CNC species, respectively. Following Refs. [99, 137] (*cf.*, also, Ref. [169]), we used the HANDE software package [170, 171] to execute the stochastic i-FCIQMC runs, needed to generate the lists of triply excited determinants included in the CC(P) and EOMCC(P) calculations. Our standalone CC/EOMCC codes, interfaced with the RHF, ROHF, and integral routines in the GAMESS program suite [172–174], were used to carry out the required CC(P), EOMCC(P), CC(P;Q), and fully deterministic (CCSD/EOMCCSD and CCSDT/EOMCCSDT) computations (the Q spaces used to construct the CC(P;Q) corrections to the CC(P) and EOMCC(P) energies consisted of the triples not captured by the i-FCIQMC runs at the corresponding propagation times τ). Each i-FCIQMC run was initiated by placing 1500 walkers on the relevant reference function (see Tables 3.1–3.4 for the details) and we set the initiator parameter n_a at 3. All of the i-FCIQMC propagations used the time step τ of 0.0001 a.u. In the post-ROHF computations for the CH and CNC species, the core electrons corresponding to the 1s shells of the carbon and nitrogen atoms were kept frozen. In the case of CH^+, we correlated all electrons.

3.2.1 CH^+

We begin our discussion of the numerical results with the CH^+ ion, where we investigated the three lowest excited states of the $^1\Sigma^+$ symmetry (labelled as 2 $^1\Sigma^+$, 3 $^1\Sigma^+$, and 4 $^1\Sigma^+$;

the ground state is designated as 1 $^1\Sigma^+$), two lowest states of the $^1\Pi$ symmetry (1 $^1\Pi$ and 2 $^1\Pi$), and two lowest $^1\Delta$ states (1 $^1\Delta$ and 2 $^1\Delta$). While these results were obtained mainly by Dr. Stephen Yuwono, with minor contributions from me, I will begin with them, since they are good for setting the stage for the subsequent discussion of the electronic excitation spectra of open-shell CH and CNC species and singlet–triplet gaps in biradicals, where I was a lead contributor. Two C–H internuclear separations were considered, the equilibrium distance $R = R_\mathrm{e} = 2.13713$ bohr [Table 3.1 and Fig. 3.2(a) and (b)] and the stretched $R = 2R_\mathrm{e}$ geometry [Table 3.2 and Fig. 3.2(c) and (d)]. Following the semi-stochastic CC(P;Q) algorithm, as described above, and our interest in converging the CCSDT/EOMCCSDT energetics, the cluster and right and left EOM operators used in the calculations for the $^1\Sigma^+$ states were approximated by $T^{(P)} = T_1 + T_2 + T_3^{(\mathrm{MC})}$, $R_\mu^{(P)} = r_{\mu,0}\mathbb{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3}^{(\mathrm{MC})}$, and $L_\mu^{(P)} = \delta_{\mu,0}\mathbb{1} + L_{\mu,1} + L_{\mu,2} + L_{\mu,3}^{(\mathrm{MC})}$, respectively, where the list of triples defining the three-body components $T_3^{(\mathrm{MC})}$, $R_{\mu,3}^{(\mathrm{MC})}$, $L_{\mu,3}^{(\mathrm{MC})}$ at a given time τ was obtained from the ground-state i-FCIQMC propagation at the same value of τ. The $T_3^{(\mathrm{MC})}$ component of $T^{(P)}$ used in the CC(P;Q) computations of the $^1\Pi$ and $^1\Delta$ states, needed to determine the similarity-transformed Hamiltonian $\bar{H}^{(P)}$ to be diagonalized in the subsequent EOMCC steps, was defined in the same way as in the case of the $^1\Sigma^+$ states, but the lists of triples entering the $R_{\mu,3}^{(\mathrm{MC})}$ component of $R_\mu^{(P)}$ and the $L_{\mu,3}^{(\mathrm{MC})}$ component of $L_\mu^{(P)}$ were obtained differently. They were extracted from the i-FCIQMC runs for the lowest states within the irreps of C_{2v} relevant to the symmetries of interest, meaning the $^1\mathrm{B}_1$ (C_{2v}) component of 1 $^1\Pi$ for the $^1\Pi$ states and the $^1\mathrm{A}_2$ (C_{2v}) component of 1 $^1\Delta$ for the $^1\Delta$ states (C_{2v} is the largest Abelian subgroup of the true point group of CH^+, $\mathrm{C}_{\infty v}$; our codes cannot handle non-Abelian symmetries). As implied by Eq. 2.24, the $\delta_\mu(P;Q)$ corrections to the CC(P) and EOMCC(P) energies at a given time τ were computed using the $\mathfrak{M}_{\mu,K}(P)$ and $\ell_{\mu,K}(P)$ amplitudes corresponding to the triply excited determinants $|\Phi_K\rangle$ not captured by i-FCIQMC at the same τ.

As pointed out in Refs. [51, 52, 137], the 2 $^1\Sigma^+$, 2 $^1\Pi$, 1 $^1\Delta$, and 2 $^2\Delta$ states of CH^+ at $R = R_\mathrm{e}$ and all of the excited states of the stretched $\mathrm{CH}^+/R = 2R_\mathrm{e}$ system, which we

calculated in this work, are characterized by substantial MR correlations that originate from large two-electron excitation contributions (the 2 $^2\Delta$ state at $R = 2R_e$ also has significant triple excitations [51, 52, 137]). It is therefore not surprising that the basic EOMCCSD level, equivalent to the EOMCC(P) calculations at $\tau = 0$, performs poorly for all of these states, producing very large errors relative to EOMCCSDT that are about 12, 20, and 34–35 millihartree for the 2 $^1\Sigma^+$, 2 $^1\Pi$, and both $^1\Delta$ states, respectively, at $R = R_e$ and ∼14–144 millihartree when the excited state at $R = 2R_e$ are considered (see Tables 3.1 and 3.2). The EOMCCSD energies for the 3 $^1\Sigma^+$, 4 $^1\Sigma^+$, and $^1\Pi$, states at the equilibrium geometry, which are dominated by one-electron transitions, are more accurate, but errors on the order of 3–6 millihartree still remain. As shown in Tables 3.1 and 3.2, the CR-EOMCC(2,3) triples correction to EOMCCSD, equivalent to the CC($P;Q$) calculations at $\tau = 0$, offers substantial improvements, as exemplified by the small errors, on the order of 1–3 millihartree, for the majority of excited states of CH^+ considered in this subsection, but there are cases, especially the 4 $^1\Sigma^+$ and 2 $^1\Delta$ states at $R = R_e$, where the differences between the CR-EOMCC(2,3) and parent EOMCCSDT energies, which are about 12 millihartree in the former case and more than 63 millihartree in the case of the latter state, remain very large. This is related to the substantial coupling of the one- and two-body components of the cluster and EOM excitation and deexcitation operators with their three-body counterparts, which the CR-EOMCC(2,3) corrections to EOMCCSD neglect. Our group's older active-space EOMCCSDt calculations for CH^+ reported in Refs. [51, 52] and the more recent semi-stochastic EOMCC(P) calculations for the same system described in Ref. [137] are indicating that the incorporations of the leading triples in the relevant P spaces, which allows the one- and two-body components of T, R_μ, and L_μ to relax in the presence of their three-body counterparts, is the key to improve the results of the CR-EOMCC(2,3) calculations.

This is exactly what we observe in Tables 3.1 and 3.2 and Fig. 3.2. In agreement with our previous work [137], by enriching the P spaces used in the CC(P) and EOMCC(P) computations with the subsets of triples captured during i-FCIQMC propagations, the results

greatly improve, allowing us to reach the millihartree or sub-millihartree accuracy levels for all the calculated excited states of CH^+ at both nuclear geometries considered in this work when the stochastically determined P spaces contain about 20–30% of all triples. The CC(P;Q) corrections to the EOMCC(P) energies based on Eq. (2.24) accelerate the convergence toward EOMCCSDT even further. As shown in Tables 3.1 and 3.2 and Fig. 3.2, these corrections are so effective that we reach the millihartree or sub-millihartree accuracy levels relative to the parent EOMCCSDT energetics almost instantaneously, *i.e.*, out of early stages of the *i*-FCIQMC propagations, when no more that 5–10% of all triples are included in the relevant P spaces. This is true even when the highly complex 4 $^1\Sigma^+$ and 2 $^1\Delta$ states at $R = R_e$, for which the EOMCCSD calculations produce the massive, ∼33 and ∼144 millihartree, errors, which remain large (about 13 and 63 millihartree, respectively) at the CR-EOMCC(2,3) level. As shown in Table 3.2, the CC(P;Q) corrections to the EOMCC(P) energies, which account for the missing triples that the *i*-FCIQMC propagations at a given time τ did not capture, allow us to reach the sub-millihartree accuracy levels with less than 5% (the 2 $^1\Delta$ state) or ∼10% (the 4 $^1\Sigma^+$ state) of triples in the relevant P spaces. The uncorrected EOMCC(P) calculations display the relatively fast convergence toward EOMCCSDT as well, but they reach similar accuracies at later propagation time τ, when about 15% (the 2 $^1\Delta$ state) or 25% (the 4 $^1\Sigma^+$ state) of triples are captured by *i*-FCIQMC. Obviously, the details of the rate of convergence of the semi-stochastic CC(P;Q) calculations toward EOMCCSDT, especially when one wants to tighten it, depend on the specific excited state being calculated, but, as shown in Tables 3.1 and 3.2, once about 20% of triples are captured by the *i*-FCIQMC propagations, we recover the EOMCCSDT energetics for all the calculated excited states of CH^+ at both geometries examined in this study to within 0.1 millihartree or better.

Interestingly, there is a great deal of consistency between the behavior of the uncorrected semi-stochastic EOMCC(P) approach, in which the lists of triples defining the relevant P spaces are extracted from *i*-FCIQMC propagations, and the fully deterministic EOMCCSDt

calculations for CH^+ reported in Refs. [51, 52], in which the leading triples were identified using active orbitals. Indeed, once the stochastically determined P spaces extracted from i-FCIQMC capture about 20–30% of all triples, which in the case of CH^+ system examined here is achieved after 50000 or fewer $\Delta\tau = 0.0001$ a.u. MC iterations, the energies resulting from the EOMCC(P) computations become very similar to those obtained with the EOMCCSDt method using the active space that consists of the highest-energy occupied (3σ) and three lowest-energy unoccupied ($1\pi_x$, $1\pi_y$, and 4σ) orbitals. Following, the definitions of the 'little t' T_3 and $R_{\mu,3}$ operators adopted in EOMCCSDt, for the state symmetries considered in this work, the active spaces consisting of 3σ, $1\pi_x$, $1\pi_y$, and 4σ valence orbitals amount to about 26–29% of all triples included in the respective EOMCC diagonalization spaces [51, 52]. This suggests that the types and values of the triply excited amplitudes defining the $R_{\mu,3}$ components of the EOM operators R_μ, which characterize the EOMCCSDt computations reported in Refs. [51, 52], and those that define the $R^{(\mathrm{MC})}_{\mu,3}$ components obtained in the i-FCIQMC-driven EOMCC(P) calculations performed after 50000 MC iterations using $\Delta\tau = 0.0001$ a.u. should be similar too. This is illustrated in Fig. 3.3, where we compare the distributions of the differences between the $R^{(\mathrm{MC})}_{\mu,3}$ amplitudes and their full EOMCCSDT counterparts resulting from the EOMCC(P) computations at 4000 [Fig. 3.3(a)], 10000 [Fig. 3.3(b)], and 50000 [Fig. 3.3(c)] MC iterations for the 2 $^1\Sigma^+$ state of CH^+ at $R = 2R_\mathrm{e}$ with the analogous distribution characterizing the $R_{\mu,3}$ amplitudes obtained with the EOMCCSDt approach using the 3σ, $1\pi_x$, $1\pi_y$, and 4σ active orbitals to define the corresponding triples space [Fig. 3.3(d)]; all EOM vectors R_μ needed to construct Fig. 3.3, corresponding to the EOMCC(P) EOMCCSDt, and EOMCCSDT calculations, were normalized to unity). As shown in Fig. 3.3 [*cf.* Fig. 3.3(c) and 3.3(d)], the small differences between the $R^{(\mathrm{MC})}_{\mu,3}$ amplitudes resulting from the EOMCC(P) calculations performed after 50000 MC iterations and the $R_{\mu,3}$ amplitudes obtained with EOMCCSDT, including their numerical values and distribution, closely resemble those characterizing the active-space EOMCCSDt computations reported in Refs. [51, 52]. This is in perfect agreement with the small errors relative to EOM-

CCSDT characterizing the two calculations, which are 0.302 millihartree in the former case (*cf.* Table 3.2) and 0.576 millihartree in the case of EOMCCSDt [51, 52]. When we start using considerably smaller fractions of triples and, as a consequence, significantly smaller P spaces in the EOMCC(P) calculations, which is what happens when the underlying i-FCIQMC propagation is terminated too soon, the differences between the $R_{\mu,3}^{(\mathrm{MC})}$ amplitudes resulting from the EOMCC(P) calculations and their EOMCCSDT counterparts, including their values and distribution, and the errors in the EOMCC(P) energies relative to EOMCCSDT increase. This can be seen in Fig. 3.3, especially when one compares panel (a), which corresponds to the EOMCC(P) calculations performed after 4000 MC iterations that use only 7% of triples, with panel (d) representing EOMCCSDt, which uses a much larger fraction of triple excitations ($\sim$30%), and in Table 3.2, where the error in EOMCC(P) energy of the 2 $^1\Sigma^+$ state of CH^+ at $R = 2R_e$ relative to EOMCCSDT obtained after 4000 MC iterations, of 4.263 millihartree, is $\sim$14 times larger than the analogous error obtained after 50000 MC steps.

The above analysis, which could be repeated for the remaining states of CH^+, reaching similar conclusions, has several interesting implications for the semi-stochastic CC($P;Q$) methodology pursued in this study, which will be examined by us in the future. It suggests, for example, that the CC(P)/EOMCC(P) and CC($P;Q$) approaches using CIQMC propagations to determine the lists of higher–than–double excitations in the corresponding P spaces can be regarded as natural alternatives to the fully deterministic active-space EOMCC methods, such as EOMCCSDt, and their CC($P;Q$)-corrected counterparts, such as CC(t;3) [70, 84, 87], whose performance in excited-state calculations will be an interesting thing for a future study. It also suggests that the fractions of higher–than–double excitations used to define the stochastically determined P spaces, needed to achieve high accuracies observed in the semi-stochastic CC($P;Q$) calculations discussed in this work, should decrease with the basis set. It was already observed in the previous ground-state semi-stochastic work [99], and we anticipate that the same will remain true in the CIQMC-driven excited-state

CC(P;Q) calculations. While this remark requires a separate thorough study, beyond the scope of this initial work on the excited-state CC(P;Q) we can rationalize it by referring to the analogies between the semi-stochastic CC(P)/EOMCC(P) and CC(P;Q) approaches and their deterministic CCSDt/EOMCCSDt and CC(t;3) counterparts. Indeed, the aforementioned (D/d) ratio that controls the speedups offered by the CC(P)/EOMCC(P) and CC(P;Q) calculations becomes $(n_o/N_o)(n_u/N_u)$ when the active-space CCSDt/EOMCCSDt and CC(t;3) calculations, based on the ideas laid down in Refs. [51, 52, 70, 77, 79, 87], are considered, where N_o and N_u are the numbers of active occupied and active unoccupied orbitals, respectively, which either do not grow with the basis set or grow with it very slowly compared to n_o and n_u.

Finally, before moving to the next molecular example, we would like to point out that, in analogy to the CC(P;Q)-based CC(t;3), CC(t,q;3), and CC(t,q;3,4) calculations using active orbitals to define the underlying P spaces (see, *e.g.*, Ref. [84]), one is better off by using smaller P spaces in the semi-stochastic CC(P)/EOMCC(P) considerations, which can be extracted out of the early stages of CIQMC propagations, and capturing the remaining correlations using noniterative CC(P;Q) corrections, than by running long-time CIQMC simulations to generate larger P spaces for the uncorrected CC(P)/EOMCC(P) calculations. This can be seen in Table 3.1 and 3.2 for CH^+ and in the remaining Tables 3.3 and 3.4 discussed in the next two subsections. We illustrate this remark by inspecting the EOMCC(P) and CC(P;Q) calculations for the 4 $^1\Sigma^+$ state of CH^+. As shown in Table 3.1, one needs to capture about 50% of triples in the P space to reach 0.1 millihartree accuracy relative to EOMCCSDT at $R = R_e$ using the uncorrected EOMCC(P) approach. When the CC(P;Q) correction is employed, only 15% of triples are needed to reach the same accuracy level. At the more challenging $R = 2R_e$ geometry (Table 3.2), one reaches a $\sim$0.1 millihartree accuracy level with about 40% of triples in the P space when using the uncorrected EOMCC(P) approach. This fraction reduces to about 20%, without any accuracy loss, when the CC(P;Q) correction is added to the EOMCC(P) energy. Based on the information provided in Sec-

tion 3.1, running the EOMCC(P) calculations with a smaller fraction of triples in the P space offers much larger savings in the computational effort than the additional time spent on determining the CC($P;Q$) correction, which is, as pointed out above, considerably less expensive than a single EOMCCSDT iteration. For example, in the pilot implementation of the excited-state EOMCC(P) and CC($P;Q$) approaches aimed at recovering EOMCCSDT energetics, employed in this work, the uncorrected EOMCC(P) run using 50% of triples in the P space, needed to reach a $\sim$0.1 millihartree accuracy relative to EOMCCSDT for the 4 $^1\Sigma^+$ state of CH^+ at $R = R_e$, is about twice as fast as the corresponding EOMCCSDT calculation. The EOMCC(P) diagonalization that forms part of the analogous CC($P;Q$) run, which needs only 15% of triples in the P space to reach the same accuracy level, is about 6 times faster than EOMCCSDT. The noniterative CC($P;Q$) correction is so inexpensive here that one can largely ignore the computational costs associated with its determination in this context [*cf.* Ref. [90] for the analogous comments made in the context of comparing costs of the CCSDt computations with those of CC(t;3)].

Table 3.1 Convergence of the CC(P)/EOMCC(P) and CC($P;Q$) energies toward CCSDT/EOMCCSDT for CH^+, calculated using the [5s3p1d/3s1p] basis set of Ref. [164], at the C–H internuclear distance $R = R_e = 2.13713$ bohr. The P spaces used in the CC(P) and EOMCC(P) calculations were defined as all singles, all doubles, and subsets of triples extracted from i-FCIQMC propagations for the lowest states of the relevant symmetries. Each i-FCIQMC run was initiated by placing 1500 walkers on the appropriate reference function [the RHF determinant for the $^1\Sigma_g^+$ states, the $3\sigma \rightarrow 1\pi$ state of the $^1B_1(C_{2v})$ symmetry for the $^1\Pi$ states, and the $3\sigma^2 \rightarrow 1\pi^2$ state of the $^1A_2(C_{2v})$ symmetry for the $^1\Delta$ states], setting the initiator parameter n_a at 3, and the time step $\Delta\tau$ at 0.0001 a.u. The Q spaces used in constructing the CC($P;Q$) corrections consisted of the triples not captured by i-FCIQMC. Adapted from Ref. [100].

MC iter. (10^3)	$1\,^1\Sigma_g^+$			$2\,^1\Sigma_g^+$		$3\,^1\Sigma_g^+$		$4\,^1\Sigma_g^+$		$1\,^1\Pi$			$2\,^1\Pi$		$1\,^1\Delta$			$2\,^1\Delta$	
	P[a]	$(P;Q)$[b]	%T[c]	P[a]	$(P;Q)$[b]	P[a]	$(P;Q)$[b]	P[a]	$(P;Q)$[b]	P[a]	$(P;Q)$[b]	%T[c]	P[a]	$(P;Q)$[b]	P[a]	$(P;Q)$[b]	%T[c]	P[a]	$(P;Q)$[b]
0[d]	1.845	0.063	0	19.694	1.373	3.856	0.787	5.537	0.954	3.080	0.792	0	11.656	2.805	34.304	−0.499	0	34.685	0.350
2	1.071	0.024	7	11.004	0.909	3.248	0.587	4.826	−4.469	0.772	0.179	13	3.746	0.530	1.492	0.151	10	5.951	0.432
4	0.423	0.009	15	5.474	0.090	1.893	0.047	1.980	0.100	0.513	0.102	20	1.852	0.128	0.525	0.051	16	2.542	0.128
6	0.249	0.003	20	4.712	0.111	1.268	0.046	1.077	0.068	0.213	0.054	25	0.957	0.073	0.471	0.028	18	1.892	0.094
8	0.181	0.003	23	1.371	0.112	0.643	0.067	0.702	0.075	0.170	0.058	27	0.743	0.060	0.240	0.021	22	0.940	0.057
10	0.172	0.004	24	1.572	0.061	0.295	0.044	0.385	0.026	0.118	0.046	29	0.411	0.047	0.198	0.017	24	0.877	0.041
50	0.077	0.001	37	0.755	0.026	0.139	0.037	0.208	0.032	0.053	0.027	43	0.157	0.027	0.039	0.008	42	0.133	0.011
100	0.044	0.000	48	0.277	0.009	0.007	0.013	0.155	0.017	0.021	0.013	57	0.063	0.012	0.014	0.005	56	0.043	0.005
150	0.015	0.000	59	0.085	0.005	0.058	0.006	0.041	0.007	0.008	0.005	71	0.020	0.004	0.004	0.002	71	0.008	0.003
200	0.006	0.000	69	0.024	0.002	0.014	0.002	0.002	0.003	0.004	0.003	82	0.008	−0.001	0.003	0.002	82	0.003	0.002
∞[e]	−38.019516			−37.702621		−37.522457		−37.386872		−37.900921			−37.498143		−37.762113			−37.402308	

[a]Errors in the CC(P) (the $1\,^1\Sigma_g^+$ ground state) and EOMCC(P) (excited states) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree.

[b]Errors in the CC($P;Q$) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree.

[c]The %T values are the percentages of triples captured during the i-FCIQMC propagations for the lowest state of a given symmetry [the $1\,^1\Sigma_g^+ = 1\,^1A_1(C_{2v})$ ground state for the $^1\Sigma_g^+$ states, the $^1B_1(C_{2v})$ component of the $1\,^1\Pi$ state for the $^1\Pi$ states, and the $^1A_2(C_{2v})$ component of the $1\,^1\Delta$ state for the $^1\Delta$ states].

[d]The CC(P) and EOMCC(P) energies at $\tau = 0.0$ a.u. are identical to the energies obtained in the CCSD and EOMCCSD calculations. The $\tau = 0.0$ a.u. CC($P;Q$) energies are equivalent to the CR-CC(2,3) (the ground state) and the CR-EOMCC(2,3) (excited states) energies.

[e]The CC(P) and EOMCC(P) energies at $\tau = \infty$ a.u. are identical to the energies obtained in the CCSDT and EOMCCSDT calculations.

Table 3.2 Same as Table 3.1 for the stretched C–H internuclear distance $R = 2R_e = 4.27426$ bohr. Adapted from Ref. [100].

MC iter. (10^3)	$1\,^1\Sigma_g^+$			$2\,^1\Sigma_g^+$		$3\,^1\Sigma_g^+$		$4\,^1\Sigma_g^+$		$1\,^1\Pi$			$2\,^1\Pi$		$1\,^1\Delta$			$2\,^1\Delta$	
	P	$(P;Q)$	%T	P	$(P;Q)$	P	$(P;Q)$	P	$(P;Q)$	P	$(P;Q)$	%T	P	$(P;Q)$	P	$(P;Q)$	%T	P	$(P;Q)$
0	5.002	0.012	0	17.140	1.646	19.929	−2.871	32.639	12.657	13.552	2.303	0	21.200	−1.429	44.495	−4.526	0	144.414	−63.405
2	1.588	0.031	3	5.209	0.478	12.524	−2.079	33.400	14.297	1.398	0.306	7	1.644	−0.060	1.372	0.046	6	13.363	0.368
4	0.504	0.015	7	4.263	−1.741	6.383	−0.760	12.671	2.178	0.712	0.058	12	0.724	0.050	0.451	0.014	9	3.338	0.130
6	0.275	0.002	11	1.405	0.047	1.352	0.051	5.870	0.593	0.409	0.033	14	0.612	0.031	0.422	0.022	12	2.340	0.063
8	0.263	0.004	12	1.543	0.065	1.173	0.020	4.406	0.699	0.436	0.050	16	0.457	−0.002	0.253	0.007	13	2.088	0.021
10	0.148	0.003	14	0.792	0.094	0.613	0.047	2.331	0.342	0.227	0.039	17	0.220	0.014	0.122	−0.001	14	0.862	0.038
50	0.030	0.000	26	0.302	0.002	0.339	0.007	0.457	0.013	0.061	0.007	30	0.079	0.060	0.047	0.005	26	0.288	0.005
100	0.009	0.000	39	0.103	0.003	0.119	0.006	0.110	0.011	0.013	0.002	41	0.016	0.004	0.013	0.004	36	0.038	0.000
150	0.004	0.000	52	0.031	0.000	0.035	0.003	0.076	0.006	0.005	0.002	52	0.007	0.002	0.005	0.001	47	0.014	0.000
200	0.001	0.000	63	0.024	0.000	0.019	0.000	−0.006	0.001	0.002	0.001	65	0.001	0.000	0.001	0.000	57	0.003	0.000
∞	−37.900394			−37.704834		−37.650242		−37.495275		−37.879532			−37.702345		−37.714180			−37.494031	

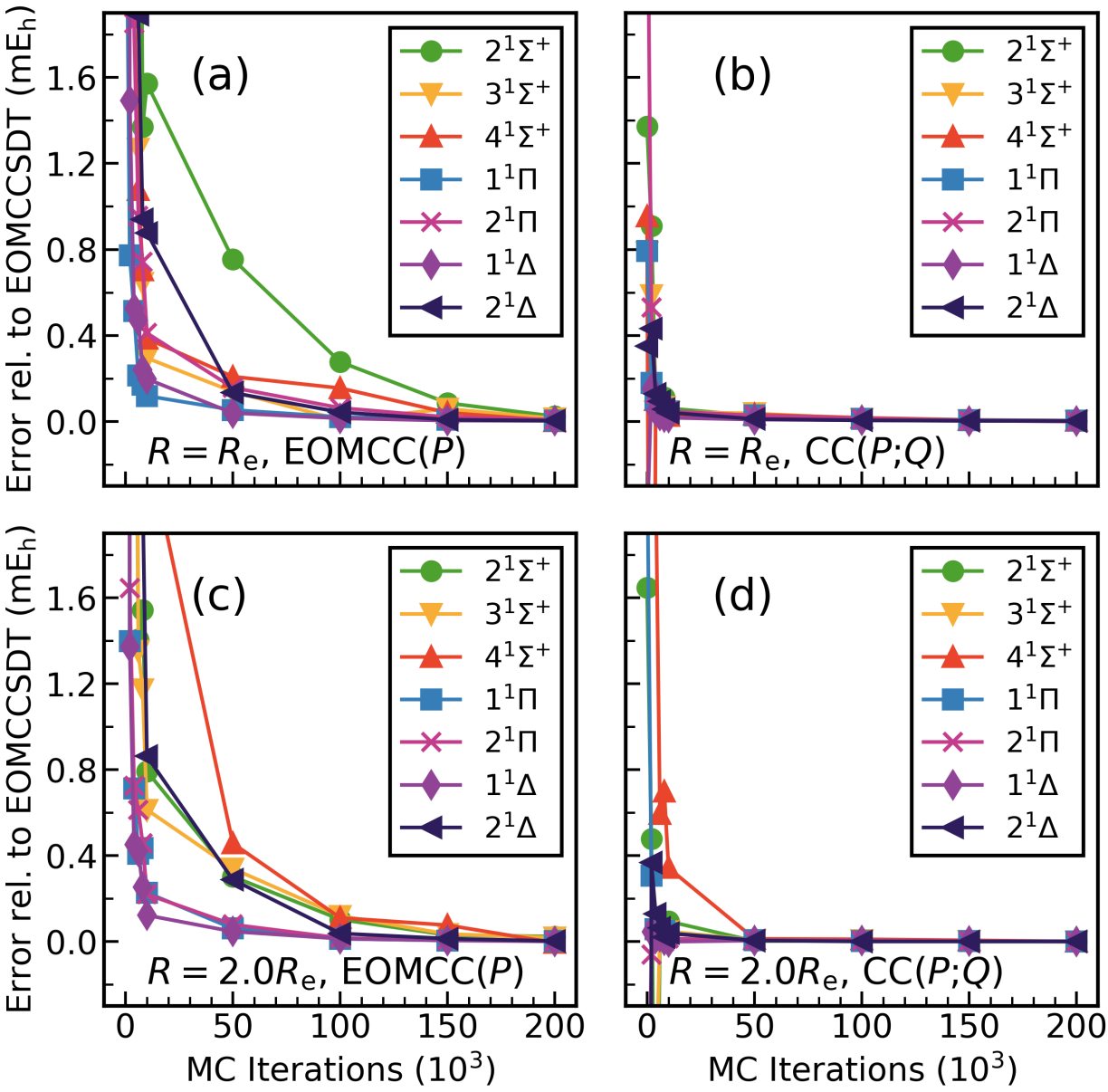

Figure 3.2 Convergence of the EOMCC(P) [panels (a) and (c)] and CC(P;Q) [panels (b) and (d)] energies toward EOMCCSDT for the three lowest-energy excited states of the $^1\Sigma^+$ symmetry, two lowest states of the $^1\Pi$ symmetry, and two lowest $^1\Delta$ states of the CH^+ ion, as described by the [5s3p1d/3s1p] basis set of Ref. [164], at the C–H internuclear distance R set at $R_e = 2.13713$ bohr [panels (a) and (b)] and $2R_e = 4.27426$ bohr [panels (c) and (d)]. Adapted from Ref. [100].

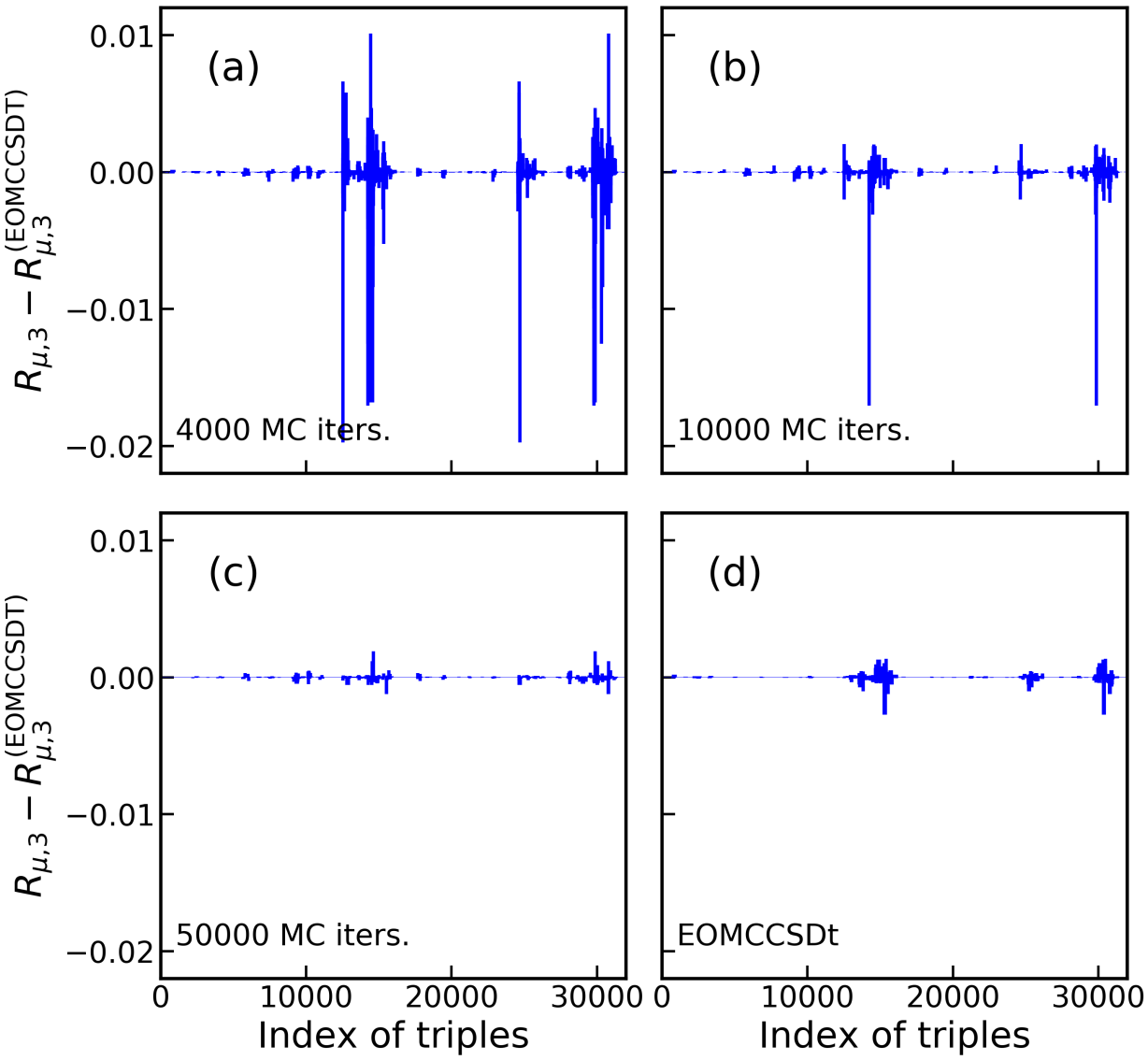

Figure 3.3 The distributions of the differences between the $R_{\mu,3}^{(\mathrm{MC})}$ amplitudes and their EOMCCSDT counterparts resulting from the EOMCC(P) computations at (a) 4000, (b) 10,000, and (c) 50,000 MC iterations using $\tau = 0.0001$ a.u. for the $2\,{}^1\Sigma^+$ state of CH^+ at $R = 2R_e$ with the analogous distribution characterizing the $R_{\mu,3}$ amplitudes obtained with the EOMCCSDt approach employing the 3σ, $1\pi_x$, $1\pi_y$, and 4σ active orbitals to define the corresponding triples space [panel (d)]. All vectors R_μ needed to construct panels (a)–(d) were normalized to unity. Adapted from Ref. [100].

3.2.2 CH

This subsection and Section 3.3 focus on the results of semi-stochastic CC($P;Q$) calculations for open-shell species and biradicals reported in Refs. [100] and [102] and obtained by me. In this subsection, we discuss the results for the CH radical that I contributed to Ref. [100].

Similar convergence patterns in the semi-stochastic EOMCC(P) and CC($P;Q$) calculations are observed for the CH radical (see Table 3.3). In this case, following our earlier deterministic EOMCC work, including the CR-EOMCC [62, 68] and electron-attachment (EA) EOMCC [68, 78, 141] approaches, and a wide range of EOMCC computations, including the high EOMCCSDT and EOMCCSDTQ levels, published by Hirata [161], along with the X $^2\Pi$ ground state, we examined the three low-lying doublet excited states, designated as A $^2\Delta$, B $^2\Sigma^-$, and C $^2\Sigma^+$, which belong to different irreducible representations than that of the ground state. In analogy to the aforementioned EOMCC studies of CH [62, 68, 78, 141, 161], the relevant CC(P) (the X $^2\Pi$ state) and EOMCC(P) (excited states) electronic energies and their CC($P;Q$) counterparts were determined at the corresponding experimentally derived equilibrium C–H distances, which are 1.1197868 Å for the X $^2\Pi$ state [175], 1.031 Å for the A $^2\Delta$ state [175], 1.1640 Å for the B $^2\Sigma^-$ state [176], and 1,1143 Å for the C $^2\Sigma^+$ state [177] (*cf.* Table 3.3). SInce all of our CC(P)/EOMCC(P) and CC($P;Q$) calculations, starting from the $\tau = 0$ CCSD/EOMCCSD and CR-EOMCC(2,3) levels and ending up with the larger values of τ needed to examine the convergence toward the parent CCSDT/EOMCCSDT energetics, were performed using the ROHF reference determinant, we also computed the ROHF-based CCSDT/EOMCCSDT energies, which formally correspond to the $\tau = \infty$ CC(P)/EOMCC(P) and CC($P;Q$) results. We had to do it, since the previously published CCSDT/EOMCCSDT results [161] relied on the unrestricted Hartree–Fock rather than the ROHF reference.

In analogy to CH^+, the lists of triples defining the $T_3^{(MC)}$ component of the cluster operator $T^{(P)}$ and the $R_{\mu,3}^{(MC)}$ and the $L_{\mu,3}^{(MC)}$ components of the EOM excitation and deexcita-

tion operators, $R_\mu^{(P)}$ and $L_\mu^{(P)}$, respectively, used in the CC(P), EOMCC(P), and CC($P;Q$) calculations for the CH radical, were extracted from the i-FCIQMC propagations for the lowest-energy states of the relevant irreps of C_{2v}, namely, the $^2B_2(C_{2v})$ component of the X $^2\Pi$ state, the lowest state of the $^2A_1(C_{2v})$ symmetry in the case of the A $^2\Delta$ and C $^2\Sigma^+$ states, and the lowest $^2A_2(C_{2v})$ state when considering the B $^2\Sigma^-$ state (again, we used C_{2v}, which is the largest Abelian subgroup of the true point group of CH, $C_{\infty v}$).

As explained in the Piecuch group's earlier papers [62, 68, 78, 141] and as shown in Ref. [161], all three excited states of the CH radical considered here, especially B $^2\Sigma^-$ and C $^2\Sigma^+$, which are dominated by two-electron excitations (*cf.* the reduced excitation level (REL) diagnostic values in Tables II and III of Ref. [68] or Table II of Ref. [62]), constitute a significant challenge, requiring the full EOMCCSDT treatment to obtain a reliable adiabatic excitation spectrum. This can be seen by inspecting the $\tau = 0$ EOMCC(P) *i.e.*, EOMCCSD, energies for the A $^2\Delta$, B $^2\Sigma^-$, and C $^2\Sigma^+$ states of CH shown in Table 3.3, which are characterized by the ∼13, ∼39, and ∼44 millihartree errors relative to EOMCCSDT, respectively. The CR-EOMCC(2,3) triples corrections to EOMCCSD, represented in Table 3.3 by the $\tau = 0$ CC($P;Q$) values, help, especially in the case of the C $^2\Sigma^+$ state, but the situation is far from ideal, since errors on the order of 8 and 5 millihartree for the A $^2\Delta$ and B $^2\Sigma^-$ states, respectively, remain. The situation considerably improves when we turn to the semi-stochasticCC($P;Q$) calculations, which incorporate the leading triples in the relevant P spaces by extracting them from the corresponding i-FCIQMC propagations and correct the resulting energies for the remaining triple excitations that were not captured by i-FCIQMC at a given time τ. As shown in Table 3.3, in the case of the A $^2\Delta$ and B $^2\Sigma^-$ states, which are not only challenging to EOMCCSD, but also to CR-EOMCC(2,3), we can reach comfortable 1–2 millihartree errors relative to EOMCCSDT using the semi-stochasticCC($P;Q$) corrections developed in this work once the relevant P spaces contain about 20–40% of all triples. With ∼50% triples in the same P space, the CC($P;Q$) energies of the A $^2\Delta$ and B $^2\Sigma^-$ states are within fractions of a millihartree from EOMCSDT. These are consider-

able improvements relative to the purely deterministic EOMCCSD and CR-EOMCC(2,3) computations, which give ∼13–39 and ∼5–8 millihartree errors, respectively, for the same two states, and the semi-stochasticEOMCC(P) calculations that reach 1–2 millihartree accuracy levels with about 70–80% triples in the respective P spaces. In the case of the C $^2\Sigma^+$ state, which is a major challenge to EOMCCSD, but not to CR-EOMCC(2,3), the behavior of the EOMCC(P) and CC($P;Q$) approaches is different, since the CC($P;Q$) corrections obtained with the help of some triples in the P space captured by i-FCIQMC are no longer needed to obtain the well-converged energetics, *i.e.*, the $\tau = 0$ CC($P;Q$) result, where the P space is spanned by singles and doubles only, is sufficiently accurate, but it is still interesting to observe that one can tighten the convergence further, reaching stable < 0.1 millihartree errors relative to EOMCCSDT with about 50% of all triples in the P space. In analogy to the A $^2\Delta$ and B $^2\Sigma^-$ states, it is also interesting to observe a reasonably smooth convergence of the uncorrected EOMCC(P) energies toward EOMCCSDT. It is clear from the results presented in Table 3.3 that the CC($P;Q$) corrections to the semi-stochastic CC(P) and EOMCC(P) energies offer considerable speedups compared to the uncorrected CC(P)/EOMCC(P)calculations, not only for the closed-shell molecules, such as CH^+, but also when examining open-shell species.

Table 3.3 Convergence of the CC(P)/EOMCC(P) and CC($P;Q$) energies toward CCSDT/EOMCCSDT for CH, calculated using the aug-cc-pVDZ basis set. The P spaces used in the CC(P) and EOMCC(P) calculations were defined as all singles, all doubles, and subsets of triples extracted from *i*-FCIQMC propagations for the lowest states of the relevant symmetries. Each *i*-FCIQMC run was initiated by placing 1500 walkers on the appropriate reference function [the ROHF $^2B_2(C_{2v})$ determinant for the X $^2\Pi$ state, the $1\pi \rightarrow 4\sigma$ state of the $^2A_1(C_{2v})$ symmetry for the A $^2\Delta$ and C $^2\Sigma^+$ states, and the $3\sigma \rightarrow 1\pi$ state of the $^2A_2(C_{2v})$ symmetry for the B $^2\Sigma^-$ state], setting the initiator parameter n_a at 3, and the time step $\Delta\tau$ at 0.0001 a.u. The Q spaces used in constructing the CC($P;Q$) corrections consisted of the triples not captured by *i*-FCIQMC. Adapted from Ref. [100].

MC iter. (10^3)	X $^2\Pi$			A $^2\Delta$			B $^2\Sigma^-$			C $^2\Sigma^+$		
	P[a]	$(P;Q)$[b]	%T[c]	P[a]	$(P;Q)$[b]	%T[c]	P[a]	$(P;Q)$[b]	%T[c]	P[a]	$(P;Q)$[b]	%T[c]
0	2.987	0.231	0.0	13.474	7.727	0.0	38.620	−4.954	0.0	43.992	0.087	0.0
2	2.405	0.170	13.8	13.009	7.395	9.8	10.602	−1.848	18.5	40.700	−0.689	9.8
4	1.413	0.086	41.7	10.907	5.288	19.3	7.066	−1.259	38.9	31.017	−0.319	19.7
6	0.883	0.035	58.9	10.119	4.577	27.2	3.452	−0.371	53.2	26.364	−0.508	28.8
8	0.603	0.022	66.8	7.764	2.436	34.6	2.309	−0.149	61.4	20.545	−0.412	34.3
10	0.495	0.019	72.6	6.987	2.170	38.1	1.965	−0.024	64.8	17.180	0.435	38.3
12	0.445	0.015	76.5	6.640	1.981	42.3	1.832	−0.081	69.5	16.929	0.029	42.5
14	0.389	0.013	77.5	7.040	1.887	45.7	1.180	0.030	72.2	13.114	0.253	45.1
16	0.309	0.008	79.2	6.047	1.667	48.3	1.303	0.012	75.6	7.646	−0.041	48.7
18	0.292	0.008	80.3	4.646	0.875	49.8	1.349	−0.062	77.5	5.312	0.011	50.1
20	0.243	0.006	82.2	3.809	0.754	52.6	0.796	0.038	79.5	4.691	0.108	52.2
50	0.150	0.002	89.1	1.367	0.112	74.1	0.298	0.038	91.6	1.436	0.070	74.0
100	0.055	0.002	95.3	0.177	0.017	91.7	0.144	0.014	98.3	0.204	0.013	91.3
150	0.025	0.000	98.1	0.042	−0.003	98.0	0.010	0.007	99.6	0.063	0.010	98.2
200	0.010	0.000	99.2	0.007	0.001	99.7	−0.001	−0.001	99.9	0.010	0.001	99.7
∞[e]	−38.387749			−38.276770			−38.267544			−38.238205		

[a]Errors in the CC(P) (the X $^2\Pi$ ground state) and EOMCC(P) (excited states) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree, calculated at the experimentally obtained equilibrium C–H distances used in Refs. [62, 68, 161], which are 1.1197868 Å for the X $^2\Pi$ state [175], 1.1031 Å for the A $^2\Delta$ state [175], 1.1640 Å for the B $^2\Sigma^-$ state [176], and 1.1143 Å for the C $^2\Sigma^+$ state [177]. The lowest-energy core orbital was frozen in all correlated calculations.

[b]Errors in the CC($P;Q$) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree, calculated at the experimentally determined equilibrium C–H distances as used in Refs. [62, 68, 161] (see footnote a for the C–H distances).

[c]The %T values are the percentages of triples captured during the *i*-FCIQMC propagations for the lowest state of a given symmetry [the $^2B_2(C_{2v})$ component of the X $^2\Pi$ ground state, the lowest $^2A_1(C_{2v})$ state for the A $^2\Delta$ and C $^2\Sigma^+$ states, and the lowest $^2A_2(C_{2v})$ state for the B $^2\Sigma^-$ state].

[d]The CC(P) and EOMCC(P) energies at $\tau = 0.0$ a.u. are identical to the energies obtained in the CCSD and EOMCCSD calculations. The $\tau = 0.0$ a.u. CC($P;Q$) energies are equivalent to the CR-CC(2,3) (the ground state) and the CR-EOMCC(2,3) (excited states) energies.

[e]The CC(P) and EOMCC(P) energies at $\tau = \infty$ a.u. are identical to the energies obtained in the ROHF-based CCSDT and EOMCCSDT calculations.

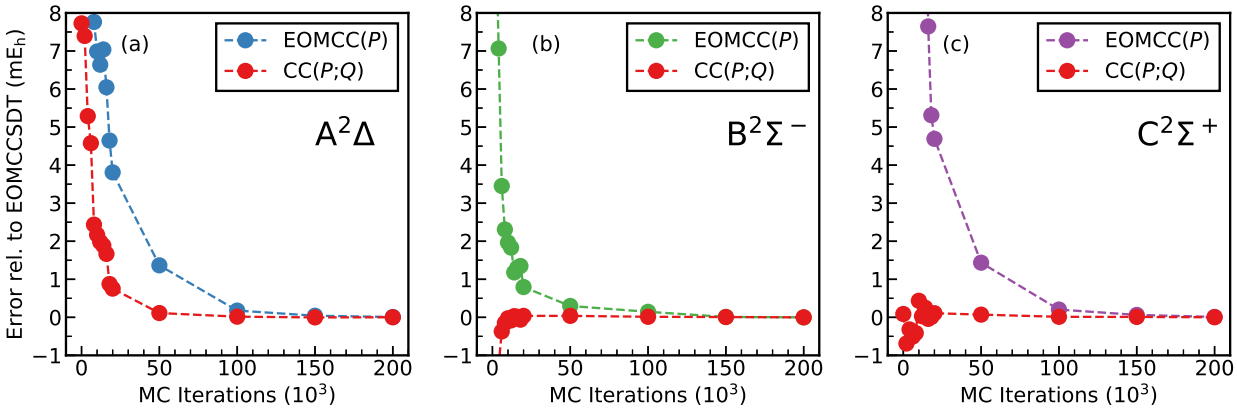

Figure 3.4 Convergence of the EOMCC(P) and CC($P;Q$) energies of the A $^2\Delta$ [panel (a)], B $^2\Sigma^-$ [panel (b)], and C $^2\Sigma^+$ [panel (c)] states of CH toward EOMCCSDT for the three lowest-energy excited states of CH calculated as described by the aug-cc-pVDZ basis set. The geometries used are the equilibrium C–H distances reported in Refs. [62, 68, 161], which are 1.1031 Å for the A $^2\Delta$ state [175], 1.1640 Å for the B $^2\Sigma^-$ state [176], and 1.1143 Å for the C $^2\Sigma^+$ state [177].

3.2.3 CNC

Our last example, which is also the largest many-electron system considered in the present study, is the linear, $D_{\infty h}$ symmetric, CNC molecule. Following our earlier CR-CC(2,3)/CR-EOMCC(2,3) and EA-EOMCC calculations for this challenging open-shell molecular species [68, 162, 163], we considered the X $^2\Pi_g$ ground state and the two low-lying doublet excited states, A $^2\Delta_u$ and B $^2\Sigma_u^+$. The *i*-FCIQMC-driven CC(P) ground-state and EOMCC(P) excited-state energies and the corresponding CC($P;Q$) corrections, along with their deterministic EOMCCSD, CR-EOMCC(2,3), and EOMCCSDT counterparts, were calculated using the equilibrium C–N distances optimized in Ref. [162] with EA-SAC-CI. They are 1.253 Å for the X $^2\Pi_g$ state, 1.256 Å for the A $^2\Delta_u$ state, and 1.259 Å for the B $^2\Sigma_u^+$ state. As in the case of the CH radical, we used the ROHF reference determinant. Following the computational protocol adopted in this study, and in analogy to the CH^+ and CH species, the lists of triples defining the $T_3^{\rm MC}$, $R_{\mu,3}^{\rm (MC)}$, and $L_{\mu,3}^{\rm (MC)}$ components used in the semi-stochastic CC(P),EOMCC(P), and CC($P;Q$) calculations for CNC were obtained using the *i*-FCIQMC propagations for the lowest-energy states of the relevant irreps of the largest Abelian subgroup of $D_{\infty h}$, *i.e.*, D_{2h}, meaning the $^2B_{2g}$ (D_{2h}) component of the X $^2\Pi_g$ state and the lowest state of the $^2B_{1u}$ (D_{2h}) symmetry in the case of the A $^2\Delta_u$ and B $^2\Sigma_u^+$ states.

As shown in Table 3.4 and in agreement with one of our previous studies [68], all three states of CNC considered in this work, especially A $^2\Delta_u$ and B $^2\Sigma_u^+$, are poorly described by CCSD and EOMCCSD, which produce more than 18, 31, and 111 millihartree errors, respectively, relative to the target EOMCCSDT energetics (see the $\tau = 0$ CC(P) and EOMCC(P) energies in Table 3.4). The excessively large, $>$ 111 millihartree, error in the EOMCCSD energy of the B $^1\Sigma_u^+$ state is related to its strongly multireference character dominated by two-electron excitations (*cf.* the REL values characterizing the excited states of CNC in Table IV of Ref. [68]). In the case of the ground state and the B $^2\Sigma_u^+$ excited state, the CR-CC(2,3) and CR-EOMCC(2,3) corrections to CCSD and EOMCCSD seem to be quite effective, reducing the large errors relative to CCSDT/EOMCCSDT observed in the CCSD and EOMCCSD

calculations to a sub-millihartree level, but the ∼16 millihartree error resulting from the CR-EOMCC(2,3) calculations for the A $^2\Delta_u$ state, while considerably lower than the $>$ 31 millihartree error obtained with EOMCCSD, is still rather large (see the $\tau = 0$ CC($P;Q$) energies in Table 3.4). By incorporating the dominant triply excited determinants captured by the *i*-FCIQMC propagations in the respective P spaces, the semi-stochasticCC(P) and EOMCC(P) approaches help, allowing us to reach stable ∼1–2 millihartree accuracy levels for the X $^2\Pi_g$ and A $^2\Delta_u$ states relative to the target CCSDT/EOMCCSDT energetics with about 50–60% triples, but the CC($P;Q$) corrections that account for the remaining triples, missing in the *i*-FCIQMC wave functions, are considerably more effective. In the case of the A $^2\Delta_u$ state, which poses problems to both EOMCCSD and CR-EOMCC(2,3), which give about 31 and 16 millihartree errors relative to EOMCCSDT, respectively, we reach a stable ∼1–2 millihartree accuracy level with about 30–40% triples in the corresponding P space, as opposed to the aforementioned 50–60% needed in the uncorrected EOMCC(P) run. The benefits of using the semi-stochastic CC($P;Q$) vs. deterministic CR-EOMCC(2,3) corrections for the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states are less obvious, but it is encouraging to observe the rapid convergence toward the target CCSDT and EOMCCSDT energetics in the former calculations. In particular, they allow us to lower the 0.4–0.5 millihartree Errors obtained with CR-EOMCC(2,3) to a 0.1 millihartree level with about 10% of all triples, identified by *i*-FCIQMC, in the case of the X $^2\Pi_g$ state and with ∼30–40% triples in the P space when the B $^2\Sigma_u^+$ state is considered. Once again, the CC($P;Q$) corrections to the energies resulting from the semi-stochastic CC(P) and EOMCC(P) calculations speed up the uncorrected CC(P)/EOMCC(P) computations, while allowing us to improve the CR-CC(2,3) and CR-EOMCC(2,3) energetics by bringing them very close to the CCSDT and EOMCCSDT levels at the fraction of the cost.

Table 3.4 Convergence of the CC(P)/EOMCC(P) and CC(P;Q) energies toward CCSDT/EOMCCSDT for CNC, calculated using DZP[4s2p1d] basis set. The P spaces used in the CC(P) and EOMCC(P) calculations were defined as all singles, all doubles, and subsets of triples extracted from i-FCIQMC propagations for the lowest states of the relevant symmetries. Each i-FCIQMC run was initiated by placing 1500 walkers on the appropriate reference function [the ROHF $^2\mathrm{B}_{2g}(\mathrm{D}_{2h})$ determinant for the X $^2\Pi_g$ state and the $3\sigma_u \rightarrow 1\pi_g$ state of the $^2\mathrm{B}_{1u}(\mathrm{D}_{2h})$ symmetry for the A $^2\Delta_u$ and B $^2\Sigma_u^+$ states], setting the initiator parameter n_a at 3, and the time step $\Delta\tau$ at 0.0001 a.u. The Q spaces used in constructing the CC(P;Q) corrections consisted of the triples not captured by i-FCIQMC. Adapted from Ref. [100].

MC iter. (10^3)	X $^2\Pi_g$			A $^2\Delta_u$			B $^2\Sigma_u^+$		
	P[a]	(P;Q)[b]	%T[c]	P[a]	(P;Q)[b]	%T[c]	P[a]	(P;Q)[b]	%T[c]
0	18.458	−0.495	0.0	31.157	16.017	0.0	111.307	−0.433	0.0
2	10.331	−0.043	13.2	18.835	9.114	6.5	81.493	−2.496	6.5
4	4.424	−0.029	33.2	10.637	5.717	16.1	53.677	−2.526	16.0
6	2.824	−0.011	44.1	7.555	4.199	22.7	35.539	−1.254	22.8
8	1.818	−0.013	49.9	6.181	3.090	27.5	26.767	−0.864	27.9
10	1.306	−0.006	53.3	5.187	2.441	30.8	21.337	−0.284	31.5
12	1.092	−0.003	56.5	4.162	1.778	34.0	17.056	0.196	34.3
14	0.911	−0.005	58.7	3.529	1.418	37.0	12.843	0.046	37.5
16	0.820	−0.003	60.6	3.106	1.149	39.5	9.197	0.134	39.9
18	0.651	−0.003	62.5	2.510	0.811	41.7	8.879	−0.034	42.4
20	0.610	−0.001	63.9	2.395	0.785	44.4	7.548	0.151	44.7
50	0.077	0.000	79.7	0.172	0.058	70.9	0.732	0.055	70.7
100	0.002	0.000	94.5	0.002	0.001	92.3	0.005	0.003	91.9
150	0.000	0.000	99.3	0.000	0.000	99.1	0.000	0.000	99.1
∞[e]	−130.421932			−130.276946			−130.252999		

[a]Errors in the CC(P) (X $^2\Pi_g$ state) and EOMCC(P) (the remaining states) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree, calculated at the experimentally obtained equilibrium C–N distances optimized in Ref. [162], which are 1.253 Å for the X $^2\Pi_g$ state, 1.256 Å for the A $^2\Delta_u$ state, and 1.259 Å for the B $^2\Sigma_u^+$ state. The three lowest-energy core orbital was frozen in all correlated calculations.

[b]Errors in the CC(P;Q) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree, calculated at the equilibrium C–N distances optimized in Ref. [162] (see footnote a for these C–N distances).

[c]The %T values are the percentages of triples captured during the i-FCIQMC propagations for the lowest state of a given symmetry [the $^2\mathrm{B}_{2g}(\mathrm{D}_{2h})$ component of the X $^2\Pi_g$ ground state and the lowest $^2\mathrm{B}_{1u}(\mathrm{D}_{2h})$ state for the A $^2\Delta$ and B $^2\Sigma_u^+$ states].

[d]The CC(P) and EOMCC(P) energies at $\tau = 0.0$ a.u. are identical to the energies obtained in the CCSD and EOMCCSD calculations. The $\tau = 0.0$ a.u. CC(P;Q) energies are equivalent to the CR-CC(2,3) (the ground state) and the CR-EOMCC(2,3) (excited states) energies.

[e]The CC(P) and EOMCC(P) energies at $\tau = \infty$ a.u. are identical to the energies obtained in the ROHF-based CCSDT and EOMCCSDT calculations.

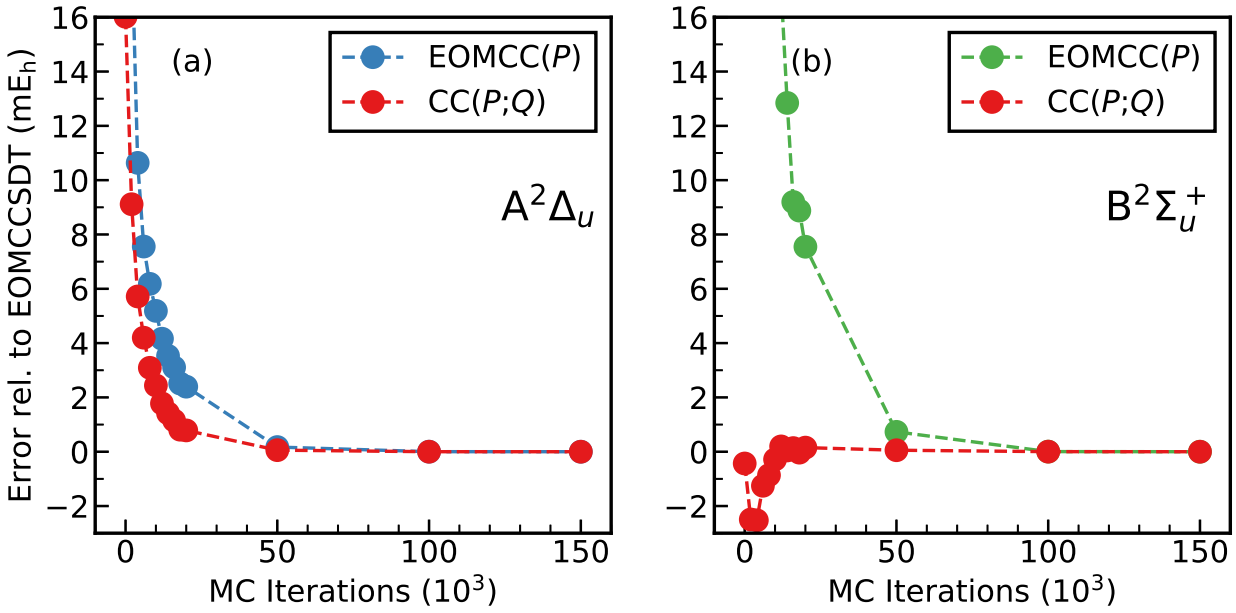

Figure 3.5 Convergence of the EOMCC(P) and CC($P;Q$) energies of the A $^2\Delta_u$ [panel (a)] and B $^2\Sigma_u^+$ [panel (b)] states of CNC toward EOMCCSDT for the two lowest-energy doublet excited states of CNC calculated as described by the DZP[4s2p1d] basis set. The geometries used are the equilibrium C–N distances reported in Refs. [68, 162], which are 1.256 Å for the A $^2\Delta_u$ state and 1.259 Å for the B $^2\Sigma_u^+$ state.

3.3 Singlet–Triplet Gaps in Methylene, $(HFH)^-$, Cyclobutadiene, Cyclopentadienyl Cation, and Trimethylenemethane

In order to assess the performance of our semi-stochastic, CIQMC-driven, CC(P;Q) methodology in converging the full CCSDT data for the singlet–triplet gaps and the corresponding singlet- and triplet-state energies of biradical systems, we applied it to methylene, $(HFH)^-$, cyclobutadiene, cyclopentadienyl cation, and trimethylenemethane. The results discussed in this section, which were all generated by me as part of this dissertation project, were reported in Ref. [102]. Following our earlier studies of the singlet–triplet gaps in the same systems using the deterministic CC(P;Q) [88] and DEA/DIP-EOMCC [80, 153–155] approaches, we used the aug-cc-pVTZ basis set [165, 166] for methylene, the 6-31G(d,p) basis [178, 179] for the $(HFH)^-$ ion, and the cc-pVDZ basis set [165] for cyclobutadiene, cyclopentadienyl cation, and trimethylenemethane. In the case of methylene and trimethylenemethane, we focused on the ability of the semi-stochastic CC(P;Q) algorithm to converge the adiabatic $\Delta E_{\text{S-T}}$ data obtained with CCSDT. When executing the semi-stochastic CC(P;Q) calculations for $(HFH)^-$, cyclobutadiene, and cyclopentadienyl cation, we focused on recovering the CCSDT values of the vertical singlet–triplet gaps. Throughout this work, we define $\Delta E_{\text{S-T}}$ as $E_{\text{S}} - E_{\text{T}}$, where E_{S} and E_{T} are the electronic energies of the corresponding singlet and triplet states, *i.e.*, the positive $\Delta E_{\text{S-T}}$ value implies that triplet is lower in energy.

All of the CC calculations reported in this section were performed using our group's standalone codes, interfaced with the RHF, ROHF, and integral transformation routines in the GAMESS package [172, 173] which were originally developed in Refs. [70, 84, 87, 88], and extended to the stochastically generated P spaces for the use in CC(P) and CC(P;Q) computations in Refs. [99–101, 137]. The *i*-FCIQMC [methylene, $(HFH)^-$, and cyclobutadiene] and *i*-CISDTQ-MC (cyclopentadienyl cation and trimethylenemethane) calculations, needed to generate the lists of triples for the semi-stochastic CC(P) and CC(P;Q) runs, were carried out with the HANDE software [170, 171]. Each of the *i*-FCIQMC and *i*-CISDTQ-MC propagations was initiated by placing 1500 walkers on the relevant reference determinant.

The CIQMC time step $\delta\tau$ and the initiator parameter n_a were set at 0.0001 a.u. and 3, respectively. In all post-Hartree–Fock calculations, the core MOs correlating with the 1s orbitals of the C and F atoms were kept frozen. If the true point group of the biradical system of interest was not Abelian, we used its largest Abelian subgroup, since our CC codes interfaced with GAMESS and the CIQMC routines in HANDE cannot handle non-Abelian symmetries.

3.3.1 Methylene

We begin the discussion of our results by analyzing the performance of the semi-stochastic CC(P;Q) approach in converging the CCSDT energies of the ground (X ^{3}B$_1$) and first-excited (A ^{1}A$_1$) states of the methylene/aug-cc-pVTZ system and the adiabatic gap between them. The C_{2v}-symmetric geometries of CH_2 in the two states, optimized using FCI and the [5s3p/3s] triple zeta basis set of Dunning [180] augmented with two sets of polarization functions (TZ2P), were taken from Ref. [181]. The electronically nondegenerate triplet ground state has a predominantly SR nature dominated by the $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^1(1b_1)^1$ configuration, whereas the first-excited singlet state exhibits a significant MR character requiring a linear combination of the $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2$ and doubly excited $(1a_1)^2(2a_1)^2(1b_2)^2(1b_1)^2$ closed-shell determinants for a proper zeroth-order description. Because of these fundamentally different characteristics of the X ^{3}B$_1$ and A ^{1}A$_1$ states, a well-balanced and accurate treatment of dynamical and nondynamical correlation effects is the key to a reliable description of the singlet–triplet gap in methylene. It is, therefore, unsurprising that one usually resorts to methods of the MRCI [181–186] or MRCC [187–190] type, or to the high-level SRCC theories that account for higher–than–doubly excited clusters in an iterative manner, such as full CCSDT used in Refs. [88, 191], to accomplish this goal (for other examples of high-level *ab initio* calculations for the X ^{3}B$_1$ and A ^{1}A$_1$ states of methylene, see Refs. [80, 153, 154, 192] and references therein). The CCSDT results for the adiabatic singlet–triplet gap in methylene, which are of interest in the present study, are indeed very accurate. As shown, for example, in Ref. [88], the difference between the adiabatic $\Delta E_{\text{S-T}}$ value ob-

tained in the CCSDT/TZ2P calculations and the corresponding FCI result of 11.14 kcal/mol [181] is only 0.11 kcal/mol or 38 cm^{-1}. As demonstrated in Ref. [88] as well, the purely electronic $A\,^1A_1 - X\,^3B_1$ separation resulting from the CCSDT computations using the aug-cc-pVTZ basis employed in this work is only about 0.15 kcal/mol ($\sim$50 cm^{-1}) higher than the experimentally derived value of 9.37 kcal/mol reported in Ref. [188], obtained by correcting the vibrationless adiabatic singlet–triplet gap determined in Ref. [193] for the relativistic and nonadiabatic (Born–Oppenheimer diagonal correction) effects estimated in Refs. [194] and [195], respectively. It is, therefore, interesting to examine if the semi-stochastic CC(P;Q) approach advocated in this work is capable of reproducing the high-quality CCSDT/aug-cc-pVTZ data for the $X\,^3B_1$ and $A\,^1A_1$ states of methylene and the adiabatic separation between them.

The results of our FCIQMC-driven CC(P;Q) calculations for the methylene/aug-cc-pVTZ system, reported as errors relative to the parent CCSDT data, and their CC(P) counterparts are shown in Table 3.5 and Fig. 3.6. The reference determinants $|\Phi\rangle$ used to initiate the i-FCIQMC propagations and to carry out the CC(P), CC(P;Q), CCSD, CR-CC(2,3), and CCSDT calculations were the ROHF determinant in the case of the $X\,^3B_1$ state and the RHF determinant for the $A\,^1A_1$ state. The subsets of triply excited determinants needed to construct the P spaces used in the CC(P) and CC(P;Q) computations at various propagation times τ were the $S_z = 1$ triples of the B_1 symmetry captured during the i-FCIQMC run for the $X\,^3B_1$ state and the $S_z = 0$ triples of the A_1 symmetry captured during the analogous run for the $A\,^1A_1$ state. Following the semi-stochastic CC(P;Q) algorithm described in Section 3.1, the Q spaces needed to determine corrections $\delta(P;Q)$ were defined as the remaining triples not captured by i-FCIQMC.

Let us start our analysis by examining the CC(P) and CC(P;Q) data at $\tau = 0$, where the P spaces do not contain any triply excited determinants. As shown in Table 3.5, the CC(P) energies of the $X\,^3B_1$ and $A\,^1A_1$ states at $\tau = 0$, which are equivalent to those obtained using conventional CCSD, are above their CCSDT [*i.e.*, $\tau = \infty$ CC(P)] counterparts by

4.187 and 5.918 millihartree, respectively. This translates into a 380 cm^{-1} or ∼11% error in the adiabatic $\Delta E_{S\text{-}T}$ value when compared to the 3328 cm^{-1} singlet–triplet gap obtained with CCSDT. The situation improves when the CC(P;Q) corrections $\delta(P;Q)$ due to T_3 correlation effects, calculated by placing all triply excited determinants in the respective Q spaces, are added to the CC(P) energies. The $\tau = 0$ CC(P;Q) or CR-CC(2,3) energy characterizing the X $^3B_1$ state is only 0.177 millihartree above the parent CCSDT value, which is an error reduction relative to CCSDT compared to the underlying CC(P) result by a factor of ∼24. The $\delta(P;Q)$ correction improves the $\tau = 0$ CC(P) energy of the more challenging A $^1A_1$ state as well, although the difference between the resulting CR-CC(2,3) energy and its CCSDT counterpart, of 0.656 millihartree, is almost 4 times larger than the analogous difference obtained for the X $^3B_1$ state. As a result, the 105 cm^{-1} error relative to CCSDT characterizing the adiabatic A $^1A_1$ − X $^3B_1$ separation obtained in the $\tau = 0$ CC(P;Q) or CR-CC(2,3) calculations, while considerably smaller than the 380 cm^{-1} obtained in the underlying CC(P) (*i.e.*, CCSD) runs, leaves room for further improvements. One can improve the CR-CC(2,3) energies of the X $^3B_1$ and A $^1A_1$ states and the gap between them by enriching the P spaces defining the CC(P) calculations with the leading triply excited determinants identified using active orbitals and correcting the resulting CCSDt energies for the remaining T_3 correlations that have not been captured by CCSDt [88], but our objective here is to examine if one can accomplish the same, or improve the CC(t;3) results reported in Ref. [88] even further, by turning to the more black-box semi-stochastic CC(P;Q) methodology, in which the dominant triply excited determinants are identified with CIQMC.

The results in Table 3.5 and Fig. 3.6 show that when the $\tau = 0$ P spaces are augmented with the subsets of triply excited determinants captured in the i-FCIQMC runs at $\tau > 0$ and, following the CC(P;Q) recipe, the resulting CC(P) energies are corrected for the remaining T_3 correlations, the convergence of the total electronic energies of the X $^3B_1$ and A $^1A_1$ states and the adiabatic separation between them toward their CCSDT parents is rapid. We can see this already in the early stages of the i-FCIQMC propagations. For example, at $\tau = 0.8$

a.u., *i.e.*, after only 8000 $\delta\tau = 0.0001$ a.u. MC iterations, the errors in the CC(P;Q) energies of the X ^{3}B$_1$ and A ^{1}A$_1$ states and the corresponding $\Delta E_{\text{S-T}}$ value relative to CCSDT are 0.049 millihartree, 0.106 millihartree, and 13 cm^{-1}, respectively, substantially improving the CR-CC(2,3) [*i.e.*, $\tau = 0$ CC(P;Q)] calculations, which give 0.177 millihartree for the X ^{3}B$_1$ state, 0.656 millihartree for the A ^{1}A$_1$ state, and 105 cm^{-1} for $\Delta E_{\text{S-T}}$. This confirms our expectation that the main source of errors in the CR-CC(2,3) computations, especially in the case of the more MR A ^{1}A$_1$ state, which is characterized by larger T_3 effects, is the use of the unrelaxed T_1 and T_2 amplitudes obtained with CCSD in constructing the correction due to triples. The FCIQMC-based CC(P;Q) calculations at $\tau = 0.8$ a.u., which use as little as 16% of all triply excited determinants to define the P space for the X ^{3}B$_1$ state and only 25% of all triples in the P space for the A ^{1}A$_1$ state, are also more accurate than the CC(t;3) computations reported in Ref. [88], which produced the 0.130 millihartree, 0.409 millihartree, and 61 cm^{-1} errors relative to CCSDT for the X ^{3}B$_1$ and A ^{1}A$_1$ energies and $\Delta E_{\text{S-T}}$, respectively. This is all very promising, especially if we realize that the *i*-FCIQMC propagations used to generate the lists of triples for our semi-stochastic CC(P;Q) runs, which work so well, are very far from convergence when $\tau = 0.8$ a.u. Indeed, as seen in Table 3.6, the total numbers of walkers at 8000 $\delta\tau = 0.0001$ a.u. MC iterations, which are 132689 in the case of the X ^{3}B$_1$ state and 165564 for the A ^{1}A$_1$ state, represent tiny fractions, 2.17% and 1.11%, respectively, of the total walker populations at $\tau = 20.0$ a.u., where we stopped our *i*-FCIQMC propagations (see Fig. 3.7 for a comparison of the rate of convergence of the CC(P), CC(P;Q), and the underlying *i*-FCIQMC calculations).

As demonstrated in Table 3.5 and Fig. 3.6, the convergence of the energies of the X ^{3}B$_1$ and A ^{1}A$_1$ states and the gap between them resulting from the FCIQMC-driven CC(P;Q) calculations remains fast at the larger propagation times τ as well. For example, if we allow *i*-FCIQMC to populate the respective P spaces with about 26–38% of all triples, which happens after 20000 $\delta\tau = 0.0001$ a.u. MC iterations, the CC(P;Q) energies of the X ^{3}B$_1$ and A ^{1}A$_1$ states and the resulting singlet–triplet gap become practically indistinguishable from

the parent CCSDT data, with errors in the total electronic energies and $\Delta E_{\text{S-T}}$ being only $\sim$20 microhartree and 2 cm^{-1}, respectively. As shown in Table 3.6 of the, walker populations characterizing the X ^{3}B$_1$ and A ^{1}A$_1$ states produced by *i*-FCIQMC at 20000 $\delta\tau = 0.0001$ a.u. MC time steps are still very small compared to the total numbers of walkers at $\tau = 20.0$ a.u., where we terminated our *i*-FCIQMC propagations (4.11% for the X ^{3}B$_1$ state and 2.17% in the case of the A ^{1}A$_1$ state). It is also interesting to note that the more MR A ^{1}A$_1$ state requires a higher fraction of triply excited determinants in the P space than its SR X ^{3}B$_1$ counterpart to achieve similar accuracy levels in the semi-stochastic CC(P;Q) computations for both states. For example, the *i*-FCIQMC propagation has to capture about 25% of all triples, for the inclusion in the P space, if we are to reduce errors relative to CCSDT in the CC(P;Q) calculations for the A ^{1}A$_1$ state to $\sim$0.1 millihartree. In the case of the X ^{3}B$_1$ state, the analogous fraction of triples is about 10% (cf. Table 3.5). This highlights the importance of balancing the SR triplet state with the more MR singlet state in obtaining accurate $\Delta E_{\text{S-T}}$ estimates, which is not a problem for the semi-stochastic CC(P;Q) methodology because the underlying *i*-FCIQMC wave function sampling is very effective in identifying the dominant higher–than–doubly excited determinants, to be included in the relevant P spaces, and the $\delta(P;Q)$ corrections to the CC(P) energies take care of the remaining correlation effects of interest.

Before concluding this subsection and discussing other molecular examples, we would like to comment on the effectiveness of the noniterative corrections $\delta(P;Q)$, adopted in the CC(P;Q) formalism, in accelerating convergence of the underlying CC(P) calculations toward CCSDT. The CC(P) and CC(P;Q) error curves shown in Fig. 3.6 illustrate this best. It is clear from this figure that the CC(P;Q) energies of the X ^{3}B$_1$ and A ^{1}A$_1$ states [Fig. 3.6 (a) and (b)] and the corresponding $\Delta E_{\text{S-T}}$ values [Fig. 3.6 (c)] converge to the parent CCSDT data much faster than in the case of the uncorrected CC(P) computations. One can see the same by inspecting the numerical data shown in Table 3.5. In this context, it is worth commenting on the CC(P) and CC(P;Q) results obtained after 8000 MC iterations. In

that case, the CC(P;Q) calculations reduce the ∼2.4 millihartree errors relative to CCSDT characterizing the CC(P) energies of the X ^{3}B$_1$ and A ^{1}A$_1$ states to 0.1 millihartree or less, which is a desired behavior, but the CC(P;Q) $\Delta E_{\text{S-T}}$ value is less accurate than that obtained with the uncorrected CC(P). One should not read too much into this though. The fact that the CC(P;Q) calculations at 8000 MC iterations increase the very small 3 cm^{-1} error obtained with CC(P) to 13 cm^{-1} is a coincidence arising from the accidental cancellation of errors in the CC(P) total electronic energies obtained at this particular propagation time. Indeed, when the later stages of the i-FCIQMC propagations are considered, the differences between the CC(P) and CCSDT values of $\Delta E_{\text{S-T}}$ become increasingly negative, reaching −107 cm^{-1} at 50000 MC iterations, before eventually converging to the CCSDT limit, whereas the corresponding CC(P;Q) results display a smooth behavior, rapidly approaching CCSDT. In particular, they reduce the relatively large negative error value obtained for $\Delta E_{\text{S-T}}$ in the CC(P) calculations at 50000 MC iterations to a numerical 0 cm^{-1}. This highlights, once again, the ability of the CC(P;Q) corrections $\delta(P;Q)$ to offer a well-balanced description of the lowest singlet and triplet states in methylene, in addition to improving the individual state energies.

Table 3.5 Convergence of the CC(P) and CC(P;Q) energies of the X ^{3}B$_1$ and A ^{1}A$_1$ states of methylene, as described by the aug-cc-pVTZ basis set, and of the corresponding adiabatic singlet–triplet gaps toward their parent CCSDT values. The geometries of the X ^{3}B$_1$ and A ^{1}A$_1$ states, optimized in the FCI calculations using the TZ2P basis set, were taken from Ref. [181]. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-FCIQMC runs. The i-FCIQMC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the ROHF (X ^{3}B$_1$ state) and RHF (A ^{1}A$_1$ state) reference determinants and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbital was kept frozen and the spherical components of d and f orbitals were employed throughout. Adapted from Ref. [102].

	X ^{3}B$_1$			A ^{1}A$_1$			A ^{1}A$_1$ − X ^{3}B$_1$	
MC Iterations	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[c]	$(P;Q)$[c]
0	4.187[d]	0.177[e]	0	5.918[d]	0.656[e]	0	380[d]	105[e]
2000	3.948	0.162	1.8	5.361	0.549	3.0	310	85
4000	3.281	0.111	7.1	3.908	0.304	11.9	138	42
6000	2.749	0.072	12.4	2.993	0.190	19.7	53	26
8000	2.428	0.049	16.3	2.444	0.106	24.9	3	13
10000	2.192	0.038	19.0	2.093	0.080	28.7	−22	9
20000	1.703	0.018	26.3	1.358	0.025	37.7	−76	2
50000	1.133	0.005	39.1	0.644	0.004	54.8	−107	0
100000	0.532	0.000	59.5	0.171	0.000	76.5	−79	0
150000	0.218	0.000	76.8	0.037	0.000	90.7	−40	0
200000	0.076	0.000	88.7	0.006	0.000	97.2	−15	0
∞	−39.080575[f]			−39.065411[f]			3328[g]	

[a] Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

[b] The %T values are the percentages of triples captured during the i-FCIQMC propagations (the $S_z = 1$ triply excited determinants of the B$_1$ symmetry in the case of the X ^{3}B$_1$ state and the $S_z = 0$ triply excited determinants of the A$_1$ symmetry in the case of the A ^{1}A$_1$ state).

[c] Unless otherwise specified, the values of the singlet–triplet gap are reported as errors relative to CCSDT in cm^{-1}.

[d] Equivalent to CCSD.

[e] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_D$].

[f] Total CCSDT energy in hartree.

[g] The CCSDT singlet–triplet gap in cm^{-1}.

Table 3.6 The total numbers of walkers, reported as percentages of the total walker populations at 200000 MC iterations, characterizing the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. that were needed to generate the CC(P) and CC(P;Q) results for methylene reported in Table 3.5. Adapted from Ref. [102].

MC Iterations	X ^{3}B$_1$	A ^{1}A$_1$
0	0.02[a]	0.01[a]
2000	0.39	0.19
4000	1.00	0.51
6000	1.65	0.83
8000	2.17	1.11
10000	2.58	1.35
20000	4.11	2.17
50000	7.69	4.64
100000	18.59	13.13
150000	43.10	36.96
200000	100[b]	100[c]

[a] The initial walker population, meaning 1500 walkers on the ROHF (X ^{3}B$_1$ state) and RHF (A ^{1}A$_1$ state) reference determinants.
[b] The total number of walkers at 200000 MC iterations is 6118222.
[c] The total number of walkers at 200000 MC iterations is 14878766.

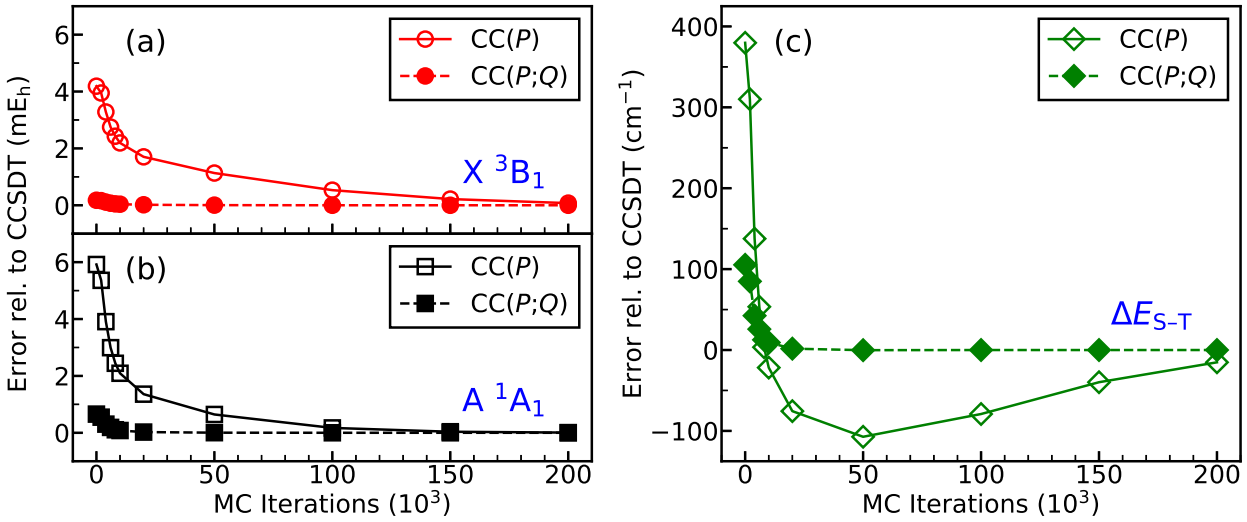

Figure 3.6 Convergence of the CC(P) and CC(P;Q) energies of the X ^{3}B$_1$ [panel (a)] and A ^{1}A$_1$ [panel (b)] states of methylene, as described by the aug-cc-pVTZ basis set, and of the corresponding adiabatic singlet–triplet gaps [panel (c)] toward their parent CCSDT values. The geometries of the X ^{3}B$_1$ and A ^{1}A$_1$ states, optimized in the FCI calculations using the TZ2P basis set, were taken from Ref. [181]. The P spaces consisted of all singles and doubles and subsets of triples identified during the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. and the Q spaces consisted of the triples not captured by i-FCIQMC. Adapted from Ref. [102].

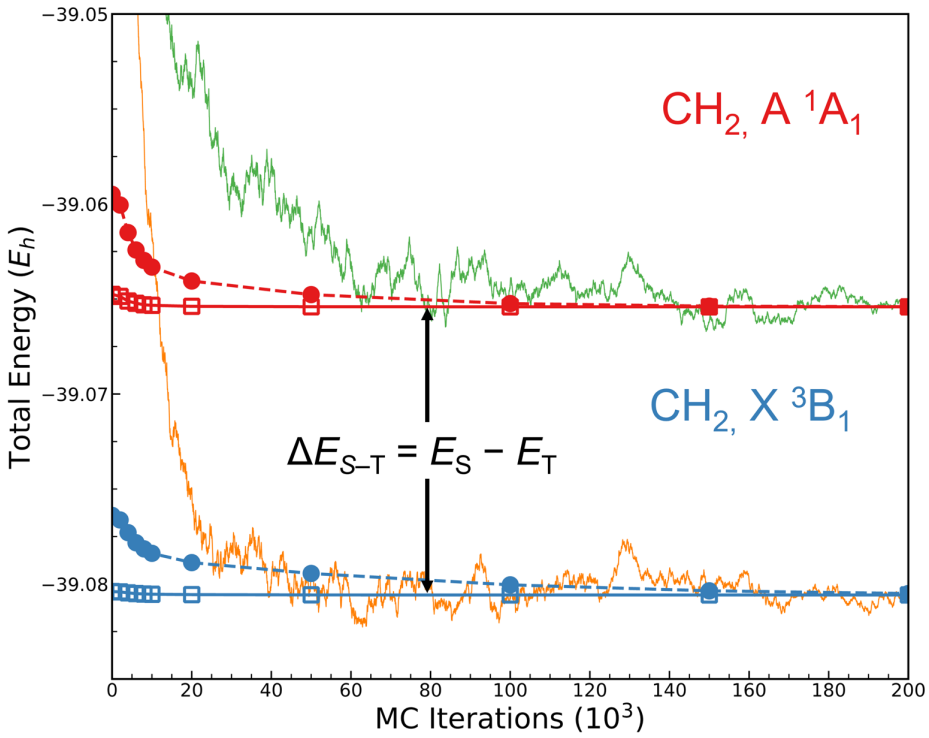

Figure 3.7 Comparison of convergences of the CC(P), CC(P;Q), and the underlying i-FCIQMC calculations toward their respective limits for the X $^3B_1$ and A $^1A_1$ states of the CH_2 molecule at their respective geometries optimized in the FCI calculations using the TZ2P basis set are taken from Ref. [181].

3.3.2 $(HFH)^-$

Our next example is the linear, $D_{\infty h}$-symmetric, $(HFH)^-$ anion, a prototype magnetic system in which unpaired spins of terminal hydrogen atoms couple to singlet and triplet states via a polarizable diamagnetic bridge of F^- [196]. The energies of the lowest two electronic states of the $(HFH)^-$ system, including the singlet ground state $X\,^1\Sigma_g^+$ and the first-excited triplet state $A\,^3\Sigma_u^+$, and the vertical gap between them, which is proportional to the magnetic exchange coupling constant J and which should approach zero as both H–F bonds are stretched to infinity, were used in the past to test various quantum chemistry approaches [66, 80, 84, 88, 153, 196–199]. Among them were methods developed in the Piecuch group, including CR-CC(2,3) [66, 197], CR-CC(2,4)[84], the DIP-EOMCC approaches with full and active-space treatments of $4h$-$2p$ correlations of top of CCSD [80, 153], and the active-orbital-based CC(t;3), CC(t,q;3), and CC(t,q;3,4) hierarchy [84, 88]. Here, we test the alternative to CC(t;3) offered by the semi-stochastic, FCIQMC-driven, CC(P;Q) algorithm aimed at the CCSDT energetics. As in our previous studies [66, 80, 84, 88, 153, 197], we used the 6-31G(d,p) basis set and several stretches of both H–F bonds, including $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å, where $R_{\text{H-F}}$ is the distance between the hydrogen and fluorine nuclei.

An accurate computation of the singlet–triplet gap in the $(HFH)^-$ system is complicated by the fact that, unlike the $A\,^3\Sigma_u^+$ state, which is weakly correlated and well represented by a single ROHF determinant, its ground-state counterpart $X\,^1\Sigma_g^+$ displays a substantial MR character that includes a significant contribution from the doubly excited $(HOMO)^2 \rightarrow (LUMO)^2$ determinant, in addition to the RHF configuration. The MR character of the $X\,^1\Sigma_g^+$ state, which is already noticeable at shorter H–F separations and which substantially strengthens as $R_{\text{H-F}}$ increases, can be illustrated by the ratio of the FCI expansion coefficients at the $(HOMO)^2 \rightarrow (LUMO)^2$ and RHF determinants or the equivalent T_2 cluster amplitude extracted from FCI, which increases, in absolute value, from 0.38 at $R_{\text{H-F}} = 1.50$ Å to 1.17 at $R_{\text{H-F}} = 4.00$ Å, when the 6-31G(d,p) basis is employed [66, 197] (the HOMO and LUMO have different symmetries, σ_g and σ_u, respectively, so that the

HOMO $\rightarrow$ LUMO T_1 amplitude is zero). As a result of all this, it is difficult to balance the lowest two states of the $(HFH)^-$ system in a single quantum chemistry calculation, especially when the SRCC framework using the RHF reference determinant for the X $^1\Sigma_g^+$ state and the ROHF reference for the A $^3\Sigma_u^+$ state is employed. Indeed, as shown in Refs. [66, 88, 197], the differences between the energies obtained in the CCSD/6-31G(d,p) computations and their FCI counterparts at $R_{\text{H-F}} = 1.50$ Å are 12.674 millihartree for the X $^1\Sigma_g^+$ state and only 2.628 millihartree when the A $^3\Sigma_u^+$ state is considered. The analogous differences at $R_{\text{H-F}} = 2.00$ Å are 19.398 and 2.068 millihartree, respectively. The observed large discrepancies between the errors in the CCSD energies for the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states translate into a poor description of the singlet–triplet gaps. One can see this by comparing the $\Delta E_{\text{S-T}}$ values resulting from the RHF/ROHF-based CCSD/6-31G(d,p) computations at $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å with the corresponding FCI data. CCSD/6-31G(d,p) gives -7320, -1838, 1656, 3605, and 230 cm^{-1}, respectively, as opposed to -9525, -4911, -2147, -277, and 0 cm^{-1} obtained with FCI [66, 88, 197]. If we are to improve the CCSD results within the SRCC framework, we must turn to higher-level theories, such as the CCSDT approach that interests us in this study [84, 88, 198], CCSDTQ [84], or the DIP-EOMCC methodology, especially after incorporating $4h$-$2p$ correlations [80, 153]. The CCSDT method is indeed very accurate, reducing the 2205, 3073, 3804, 3882, and 230 cm^{-1} errors relative to FCI in the $\Delta E_{\text{S-T}}$ values obtained with CCSD/6-31G(d,p) at $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å to 198, 270, 341, 420, and 58 cm^{-1}, respectively [84, 88, 198]. It also greatly improves the total electronic energies. Indeed, the differences between the CCSDT and FCI energies of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states in the entire $R_{\text{H-F}} = 1.50 - 4.00$ Å region obtained using the 6-31G(d,p) basis set do not exceed 2.276 and 0.389 millihartree, respectively [84, 88]. The analogous differences between the CCSD and FCI energies are as large as 20.546 millihartree for the former state and 2.628 millihartree when the latter state is considered (see Fig. 3.8 for a comparison of CCSD, CCSDT, and FCI energetics throughout the 1.500 Å–4.00Å range). One can reduce the remaining small errors in the CCSDT results even further or practi-

cally eliminate them by using CCSDTQ [84] or the DIP-EOMCC approaches with $4h$-$2p$ contributions, [80, 153] but the objective of this study is to assess the performance of our semi-stochastic CC(P;Q) methodology in converging the CCSDT data.

The results of our FCIQMC-driven CC(P;Q)/6-31G(d,p) computations for the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states of the linear $(HFH)^-$ system at the H–F distances $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å and the corresponding $\Delta E_{\text{S-T}}$ values, along with the underlying CC(P) data, are reported in Tables 3.7–3.9 and Fig. 3.9. In all of our CC(P) and CC(P;Q) computations and the underlying i-FCIQMC runs for the $D_{\infty h}$-symmetric $(HFH)^-$ system, we used the D_{2h} Abelian subgroup of $D_{\infty h}$. In particular, the i-FCIQMC calculations for the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states were set up to converge the lowest-energy states of the $^1A_g(D_{2h})$ and $^3B_{1u}$ (D_{2h}) symmetries. As a result, the subsets of triply excited determinants used to construct the P spaces for the subsequent CC(P) and CC(P;Q) computations for the X $^1\Sigma_g^+$ state at the various $R_{\text{H-F}}$ and τ values considered in this work were defined as the $S_z = 0$ triples of the A_g (D_{2h}) symmetry captured in the underlying i-FCIQMC propagations. Similarly, the subsets of triply excited determinants used to design the P spaces for the CC(P) and CC(P;Q) calculations for the A $^3\Sigma_u^+$ state were the $S_z = 1$ triples of the B_{1u} (D_{2h}) symmetry extracted from i-FCIQMC. In analogy to all other CC(P;Q) computations performed in this work, the Q spaces used to determine the $\delta(P;Q)$ corrections to the CC(P) energies were defined as the remaining triples not captured by the respective i-FCIQMC runs.

As shown in Table 3.7, and in line with our earlier CC(P;Q) work [88] and the above remarks, the CC(P) energies of the X $^1\Sigma_g^+$ state of $(HFH)^-$ obtained at $\tau = 0$, which are identical to those resulting from the conventional CCSD calculations reported in Refs. [66, 88, 197], are characterized by large errors relative to their $\tau = \infty$, *i.e.*, CCSDT, parents. Indeed, the differences between the $\tau = 0$ and $\tau = \infty$ CC(P) energies for the X $^1\Sigma_g^+$ state increase from 11.412 millihartree at $R_{\text{H-F}} = 1.50$ Å to more than 17 millihartree at $R_{\text{H-F}} =$ 2.00 and 2.50 Å. These differences become smaller at large H–F separations, represented in our calculations by $R_{\text{H-F}} = 4.00$ Å, where the $D_{\infty h}$-symmetric $(HFH)^-$ system is essentially

dissociated into the stretched hydrogen molecule, which has only two electrons, so that CCSD becomes exact, and the closed-shell fluoride ion, which has the electronic structure of the neon atom and which is characterized by small T_n correlations with $n > 2$, but they remain large when the $R_{\text{H-F}}$ values are smaller. This should be contrasted with the small, $\sim$1–2 millihartree, differences between the $\tau = 0$ and $\tau = \infty$ CC(P) energies obtained at all values of $R_{\text{H-F}}$ for the predominantly SR A $^3\Sigma_u^+$ state (see Table 3.8). As already alluded to above, and as shown in Table 3.9, this imbalance in the description of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states by the CCSD, *i.e.*, $\tau = 0$ CC(P), calculations gives rise to large errors in the resulting $\Delta E_{\text{S-T}}$ values relative to their $\tau = \infty$ (CCSDT) counterparts, which range from 2007 cm^{-1} to 3462 cm^{-1} in the $R_{\text{H-F}} = 1.50$–2.50 Å region. Once again, these errors become small at large H–F separations, such as $R_{\text{H-F}} = 4.00$ Å used in this work, where $(\text{HFH})^-$ is more or less equivalent to the stretched H_2 and F^-, resulting in the nearly degenerate singlet and triplet states and the 172 cm^{-1} difference between the $\tau = 0$ and $\tau = \infty$ CC(P) values of $\Delta E_{\text{S-T}}$, but at shorter H–F distances they are large and comparable to or even larger than the singlet–triplet gap values provided by CCSDT or FCI.

The situation dramatically changes, when the $\tau = 0$ CC(P) or CCSD energies are corrected for T_3 correlations with the help of the noniterative correction $\delta(P;Q)$, as in the $\tau = 0$ CC(P;Q) calculations, which are equivalent to the purely deterministic CR-CC(2,3) runs reported in Refs. [66, 84, 88, 197]. As shown in Tables 3.7–3.9, the $\tau = 0$ CC(P;Q), *i.e.*, CR-CC(2,3), energies of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states at the various H–F distances considered in this study and the gaps between them are substantially more accurate than their uncorrected CC(P) (*i.e.*, CCSD) counterparts. For example, the CR-CC(2,3) approach reduces the large, more than 17 millihartree, errors in the CCSD energies of the X $^1\Sigma_g^+$ state relative to their CCSDT [$\tau = \infty$ CC(P) or CC(P;Q)] parents at $R_{\text{H-F}} = 2.00$ and 2.50 Å to $\sim$1–3 millihartree. We see similarly significant improvements in the CCSD energies of the X $^1\Sigma_g^+$ state by CR-CC(2,3) at other H–F distances, even at the "easiest" $R_{\text{H-F}} = 4.00$ Å value, where the triples correction $\delta(P;Q)$ is capable of reducing the already small, 1.907 milli-

hartree, difference between the CCSD and CCSDT energies to the much smaller (in absolute value) 0.291 millihartree (see Table 3.7). Consistent with our earlier studies, [66, 84, 88, 197] the CR-CC(2,3) method performs even better when the weakly correlated A $^3\Sigma_u^+$ state is examined, reducing the $\sim$1–2 millihartree errors in the underlying CCSD energetics relative to our CCSDT target to about 0.2 millihartree (see Table 3.8). As a result of all of these accuracy improvements, the singlet–triplet gap values obtained using CR-CC(2,3) are much closer to their CCSDT parents than their CCSD counterparts, reducing the 2007, 2803, 3462, 3462, and 172 cm^{-1} errors relative to CCSDT obtained with CCSD at $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å, respectively, by factors ranging from 6 at $R_{\text{H-F}} = 2.50$ Å to 72 at $R_{\text{H-F}} = 1.50$ Å, but, as shown in Table 3.9 [see, also, Ref. [88], where one can find a comparison of the CCSD, CR-CC(2,3), and CCSDT $\Delta E_{\text{S-T}}$ data for additional H–F distances], the differences on the order of (-600)–(-300) cm^{-1} between the CR-CC(2,3) and CCSDT singlet–triplet separations in the intermediate $R_{\text{H-F}} = 2.00$–3.00 Å region remain. The question arises if one can refine the CR-CC(2,3) results by enriching the P spaces used in the CC(P;Q) calculations, which in CR-CC(2,3) consist of only singles and doubles, with the subsets of triply excited determinants identified by i-FCIQMC propagations.

As shown in Tables 3.7–3.9 and Fig. 3.9, once the leading triply excited determinants, captured using i-FCIQMC at $\tau > 0$, are included in the respective P spaces and the $\delta(P;Q)$ corrections due to the remaining T_3 correlations are added to the energies obtained in the CC(P) calculations, the resulting CC(P;Q) values of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ energies and vertical gaps between them display very fast convergence toward their CCSDT counterparts. This is already observed when the i-FCIQMC propagation times are short, engaging tiny walker populations that are orders of magnitude smaller than those required to converge the i-FCIQMC runs, and the fractions of the triply excited determinants captured by i-FCIQMC are small. For example, after as few as 2000 $\delta\tau = 0.0001$ a.u. MC iterations, where τ is only 0.2 a.u. and where, as shown in Table 3.10, the total walker populations characterizing the underlying i-FCIQMC runs are 0.01–0.11% of the respective numbers of walkers at $\tau = 20.0$

a.u. [the termination time for our i-FCIQMC propagations for $(HFH)^-$], the differences between the CC(P;Q) and CCSDT energies obtained for the strongly correlated X $^1\Sigma_g^+$ state are -0.035 millihartree for $R_{\text{H-F}} = 1.50$ Å, -0.056 millihartree for $R_{\text{H-F}} = 1.75$ Å, -0.110 millihartree for $R_{\text{H-F}} = 2.00$ Å, -0.583 millihartree for $R_{\text{H-F}} = 2.50$ Å, and -0.025 millihartree for $R_{\text{H-F}} = 4.00$ Å. In spite of using only about 10–30% of all triply excited determinants in the underlying P spaces, the FCIQMC-based CC(P;Q) energies of the X $^1\Sigma_g^+$ state obtained after 2000 MC iterations reduce the errors relative to CCSDT characterizing the CR-CC(2,3) [*i.e.*, $\tau = 0$ CC(P;Q)] computations in the $R_{\text{H-F}} = 1.50$–4.00 Å region by factors ranging from 5 to 13 (see Table 3.7). In fact, with an exception of $R_{\text{H-F}} = 2.00$ and 2.50 Å, they are much more accurate than the results produced by the purely deterministic CC(t;3) analog of the semi-stochastic CC(P;Q) methodology, reported in Refs. [84, 88]. One can observe even more dramatic improvements over CR-CC(2,3) offered by the FCIQMC-driven CC(P;Q) approach, when the propagation time τ increases. For example, after 4000 MC iterations, where the i-FCIQMC propagations are still far from being converged (*cf.* the total walker populations used by our i-FCIQMC runs relative to the termination time $\tau =$ 20.0 a.u. in Table 3.10) and the fractions of triples included in the stochastically determined P spaces, which range from 12% at $R_{\text{H-F}} = 4.00$ Å to 56% at $R_{\text{H-F}} = 1.50$ Å, remain relatively small, the differences between the CC(P;Q) and CCSDT energies obtained for the X $^1\Sigma_g^+$ state at $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å are -28, -9, -17, -50, and -4 microhartree, respectively, reducing the errors relative to CCSDT that characterize the corresponding CR-CC(2,3) calculations by factors ranging from 12 to 86 [2 to 61 when compared to the CC(t;3) results reported in Refs. [84, 88]]. As shown in Table 3.8, the performance of the FCIQMC-driven CC(P;Q) approach becomes even more impressive when the A $^3\Sigma_u^+$ state, which has a SR character, is examined. After 2000 $\delta\tau = 0.0001$ a.u. MC time steps, the errors in the CC(P;Q) energies relative to their CCSDT parents obtained for the A $^3\Sigma_u^+$ state at $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å are only -40, -24, -38, -29, and -14 microhartree, respectively. After 4000 MC iterations, they become -10, -10, -12,

−9, and −2 microhartree, respectively. Once again, these are considerable improvements compared to CR-CC(2,3) and CC(t;3) that both give errors on the order of −0.2 millihartree [84, 88], especially if we realize that the fractions of triples captured by the *i*-FCIQMC runs after 2000 and 4000 MC iterations are relatively small (5–28% and 5–49%, respectively) and, as shown in Table 3.10, the corresponding numbers of walkers represent only about 1–2% of the total numbers of walkers at $\tau = 20.0$ a.u., where we stopped our *i*-FCIQMC propagations.

As a consequence of the small errors in the CC(P;Q) total energies characterizing the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states in the early stages of the *i*-FCIQMC propagations, the resulting singlet–triplet gap values are very accurate as well. This is demonstrated in Table 3.9, where one can see that after 2000 $\delta\tau = 0.0001$ a.u. MC iterations, which is, as already explained, a very short propagation time engaging tiny walker populations and small fractions of triples, most of the differences between the CC(P;Q) and CCSDT $\Delta E_{\text{S-T}}$ values in the $R_{\text{H-F}} = 1.50$–4.00 Å region are on the order of a few reciprocal centimeter. The only exception is the semi-stochastic CC(P;Q) run at $R_{\text{H-F}} = 2.50$ Å, where the −122 cm^{-1} error relative to CCSDT characterizing the singlet–triplet gap obtained after 2000 MC time steps, while representing a five-fold error reduction compared to CR-CC(2,3), is comparable, in magnitude, to the CCSDT value of $\Delta E_{\text{S-T}}$. This happens because the CC(P;Q) energy of the strongly correlated X $^1\Sigma_g^+$ state obtained after 2000 MC iterations at $R_{\text{H-F}} = 2.50$ Å differs from its CCSDT parent by −0.583 millihartree, whereas the analogous difference between the CC(P;Q) and CCSDT energies for its weakly correlated A $^3\Sigma_u^+$ companion is only −29 microhartree. This is not a problem though, since by running *i*-FCIQMC a little longer and capturing about 20% of all triply excited determinants in the relevant P spaces, as is the case when 4000 $\delta\tau = 0.0001$ a.u. MC time steps are considered, one reduces the differences between the CC(P;Q) and CCSDT energies of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states to −50 and −9 microhartree, respectively (*cf.* Tables 3.7 and 3.8), so that the 122 cm^{-1} unsigned error in the CC(P;Q) value of $\Delta E_{\text{S-T}}$ relative to CCSDT obtained after 2000 MC iterations

decreases to less than 10 cm^{-1}. This is yet another illustration of the ability of the semi-stochastic CC(P;Q) methodology pursued in this work to balance the more MR singlet and weakly correlated triplet states of biradical systems in a single computation at the fraction of the cost of the parent high-level CC calculations. As shown in Table 3.9, at $\tau = 0.4$ a.u., where the i-FCIQMC propagations are still far from being converged, the FCIQMC-driven CC(P;Q) calculations recover the CCSDT values of the singlet–triplet gaps in $(HFH)^-$ at all H–F distances considered in this study to within a few reciprocal centimeter, reaching a 1–2 cm^{-1} or better accuracy after 6000 MC iterations.

Last, but not least, the results reported in Tables 3.7–3.9 and Fig. 3.9 also demonstrate the remarkable efficiency of the $\delta(P;Q)$ corrections in accelerating the convergence of the CC(P) energies of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states and the vertical gaps between them toward CCSDT, independent of the H–F distance considered. Let us, for example, compare the uncorrected CC(P) and corrected CC(P;Q) energies of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states of $(HFH)^-$ at the five H–F separations considered in this work obtained after 2000 MC iterations. In the case of the former, more MR, state, the CC(P;Q) corrections $\delta(P;Q)$ reduce the positive 2.601, 3.998, 3.511, 6.586, and 0.412 millihartree errors relative to CCSDT resulting from the CC(P) computations at $R_{\text{H-F}} = 1.50$, 1.75, 2.00, 2.50, and 4.00 Å to the much smaller negative error values of −0.035, −0.056, −0.110, −0.583, and −0.025 millihartree, respectively. When the latter state, which is characterized by much weaker correlations, is considered, the 0.995, 0.826, 0.834, 0.502, and 0.239 millihartree errors obtained with CC(P) are reduced to −40, −24, −38, −29, and −14 microhartree, respectively, when the CC(P;Q) approach is employed. It is interesting to notice that while the errors characterizing the CC(P) calculations for the A $^3\Sigma_u^+$ state are generally much smaller than their X $^1\Sigma_g^+$ counterparts, and the two states have a substantially different character, the error reductions offered by the CC(P;Q) corrections $\delta(P;Q)$, by at least one order of magnitude, apply to both states. As already alluded to above, and as shown in Table 3.9 and Fig. 3.9 (e) and (f), where we examine the convergence of the CC(P) and CC(P;Q) $\Delta E_{\text{S-T}}$ values toward their CCSDT

parents, the noniterative corrections $\delta(P;Q)$ are also very effective in improving the balance in the description of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states by the CC(P) approach and smoothing the convergence of the resulting singlet–triplet gaps toward their CCSDT limits. This can be illustrated by comparing the behavior of the error values relative to CCSDT characterizing the CC(P) calculations of $\Delta E_{\text{S-T}}$ at $R_{\text{H-F}} = 2.00$ Å with their CC(P;Q) counterparts, shown in Table 3.9. In the former case, the 3462 cm^{-1} error at $\tau = 0$ decreases, in absolute value, to 8 cm^{-1} at $\tau = 0.8$ a.u. (8000 MC iterations), to increase to 17 cm^{-1} at $\tau = 2.0$ a.u. (20000 MC iterations), to decrease again to a numerical 0 cm^{-1} at $\tau = 20.0$ a.u. (200000 MC iterations). Once the CC(P) energies of the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states are corrected using the $\delta(P;Q)$ corrections, the unsigned errors in the resulting CC(P;Q) values of $\Delta E_{\text{S-T}}$ relative to their CCSDT parent monotonically and rapidly decrease, from 282 cm^{-1} at $\tau = 0$ to a numerical 0 cm^{-1} at $\tau \geq 0.8$ a.u. It is clear from Tables 3.7–3.9 and Fig. 3.9 that while both the CC(P) and CC(P;Q) energies converge to the parent CCSDT limit, the latter energies and the gaps between them converge to CCSDT a lot faster.

Table 3.7 Convergence of the CC(P) and CC(P;Q) energies of the X $^1\Sigma_g^+$ state of $(HFH)^-$, as described by the 6-31G(d,p) basis set, at selected H–F distances $R_{H\text{-}F}$ toward their parent CCSDT values. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-FCIQMC runs. The i-FCIQMC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the RHF reference determinant and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbital was kept frozen and the spherical components of d orbitals were employed throughout. Adapted from Ref. [102].

	$R_{H\text{-}F} = 1.50$ Å			$R_{H\text{-}F} = 1.75$ Å			$R_{H\text{-}F} = 2.00$ Å			$R_{H\text{-}F} = 2.50$ Å			$R_{H\text{-}F} = 4.00$ Å		
MC Iters.	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]
0	11.412[c]	−0.343[d]	0	14.738[c]	−0.686[d]	0	17.453[c]	−1.455[d]	0	17.051[c]	−2.800[d]	0	1.907[c]	−0.291[d]	0
2000	2.601	−0.035	34.2	3.998	−0.056	30.5	3.511	−0.110	22.6	6.586	−0.583	15.2	0.412	−0.025	7.6
4000	0.843	−0.028	56.1	1.078	−0.009	49.6	1.979	−0.017	40.5	0.973	−0.050	25.6	0.141	−0.004	11.7
6000	0.595	−0.004	63.9	0.434	−0.003	58.1	0.432	−0.010	46.7	0.459	−0.012	30.2	0.076	−0.003	13.2
8000	0.225	−0.003	68.6	0.477	−0.007	61.4	0.187	−0.003	50.2	0.225	−0.003	33.9	0.037	−0.001	14.4
10000	0.258	−0.003	70.9	0.161	−0.002	63.3	0.136	−0.003	54.5	0.167	0.000	35.4	0.025	−0.001	15.3
20000	0.112	0.000	77.2	0.056	−0.001	71.0	0.079	−0.002	61.1	0.042	−0.001	41.8	0.026	−0.001	19.0
50000	0.017	0.000	88.4	0.019	0.000	85.8	0.005	0.000	77.5	0.009	0.000	58.8	0.002	−0.001	28.6
100000	0.002	0.000	97.7	0.001	0.000	96.3	0.001	0.000	94.4	0.000	0.000	81.8	0.000	0.000	54.8
150000	0.000	0.000	99.5	0.000	0.000	99.4	0.000	0.000	99.2	0.000	0.000	94.1	0.000	0.000	73.7
200000	0.000	0.000	99.9	0.000	0.000	100.0	0.000	0.000	99.9	0.000	0.000	99.2	0.000	0.000	86.9
∞	−100.588130[e]			−100.576056[e]			−100.561110[e]			−100.539783[e]			−100.525901[e]		

[a] Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

[b] The %T values are the percentages of triples captured during the i-FCIQMC propagations [the $S_z = 0$ triply excited determinants of the A_g (D_{2h}) symmetry].

[c] Equivalent to CCSD.

[d] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_D$].

[e] Total CCSDT energy in hartree.

Table 3.8 Convergence of the CC(P) and CC(P;Q) energies of the A $^3\Sigma_u^+$ state of $(HFH)^-$, as described by the 6-31G(d,p) basis set, at selected H–F distances $R_{H\text{-}F}$ toward their parent CCSDT values. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-FCIQMC runs. The i-FCIQMC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the ROHF reference determinant and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbital was kept frozen and the spherical components of d orbitals were employed throughout. Adapted from Ref. [102].

	$R_{H\text{-}F} = 1.50$ Å			$R_{H\text{-}F} = 1.75$ Å			$R_{H\text{-}F} = 2.00$ Å			$R_{H\text{-}F} = 2.50$ Å			$R_{H\text{-}F} = 4.00$ Å		
MC Iters.	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]
0	2.268[c]	−0.217[d]	0	1.967[c]	−0.181[d]	0	1.678[c]	−0.172[d]	0	1.277[c]	−0.167[d]	0	1.123[c]	−0.180[d]	0
2000	0.995	−0.040	27.8	0.826	−0.024	24.2	0.834	−0.038	19.1	0.502	−0.029	10.7	0.239	−0.014	4.5
4000	0.456	−0.010	49.4	0.477	−0.010	41.7	0.475	−0.012	33.7	0.236	−0.009	17.2	0.079	−0.002	5.4
6000	0.338	−0.005	56.4	0.266	−0.001	50.5	0.321	−0.005	41.2	0.174	−0.003	21.5	0.070	−0.003	5.9
8000	0.290	−0.003	60.1	0.254	−0.003	54.2	0.225	−0.003	44.7	0.195	−0.006	23.9	0.064	−0.002	6.0
10000	0.271	−0.003	61.1	0.267	−0.004	56.6	0.201	−0.002	45.7	0.064	−0.003	25.0	0.056	−0.002	6.4
20000	0.201	−0.002	67.9	0.151	−0.001	62.1	0.157	−0.002	52.2	0.078	−0.003	28.6	0.025	−0.001	7.4
50000	0.082	0.000	80.0	0.056	0.000	76.3	0.069	−0.001	66.1	0.049	−0.001	37.4	0.012	0.000	8.4
100000	0.021	0.000	91.8	0.016	0.000	89.4	0.015	0.000	82.8	0.014	0.000	52.9	0.002	0.000	11.7
150000	0.007	0.000	96.7	0.003	0.000	95.8	0.003	0.000	92.9	0.002	0.000	68.6	0.001	0.000	16.8
200000	0.001	0.000	98.8	0.001	0.000	98.4	0.001	0.000	97.1	0.000	0.000	81.8	0.000	0.000	23.8
∞	−100.545633[e]			−100.554908[e]			−100.552882[e]			−100.540435[e]			−100.526164[e]		

[a] Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

[b] The %T values are the percentages of triples captured during the i-FCIQMC propagations [the $S_z = 1$ triply excited determinants of the B_{1u} (D_{2h}) symmetry].

[c] Equivalent to CCSD.

[d] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_D$].

[e] Total CCSDT energy in hartree.

Table 3.9 Convergence of the CC(P) and CC(P;Q) singlet–triplet gaps of $(HFH)^-$, as described by the 6-31G(d,p) basis set, at selected H–F distances $R_{\text{H-F}}$ toward their parent CCSDT values. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-FCIQMC runs. The i-FCIQMC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the RHF ($X\ ^1\Sigma_g^+$ state) and ROHF ($A\ ^3\Sigma_u^+$ state) reference determinants and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbital was kept frozen and the spherical components of d orbitals were employed throughout. Adapted from Ref. [102].

	$R_{\text{H-F}} = 1.50$ Å		$R_{\text{H-F}} = 1.75$ Å		$R_{\text{H-F}} = 2.00$ Å		$R_{\text{H-F}} = 2.50$ Å		$R_{\text{H-F}} = 4.00$ Å	
MC Iters.	P[a]	$(P;Q)$[a]	P[a]	$(P;Q)$[a]	P[a]	$(P;Q)$[a]	P[a]	$(P;Q)$[a]	P[a]	$(P;Q)$[a]
0	2007[b]	−28[c]	2803[b]	−111[c]	3462[b]	−282[c]	3462[b]	−578[c]	172[b]	−24[c]
2000	353	1	696	−7	588	−16	1335	−122	38	−2
4000	85	−4	132	0	330	−1	162	−9	14	0
6000	56	0	37	0	24	−1	62	−2	1	0
8000	−14	0	49	−1	−8	0	7	1	−6	0
10000	−3	0	−23	0	−14	0	23	0	−7	0
20000	−20	0	−21	0	−17	0	−8	1	0	0
50000	−14	0	−8	0	−14	0	−9	0	−2	0
100000	−4	0	−3	0	−3	0	−3	0	−1	0
150000	−2	0	−1	0	−1	0	0	0	0	0
200000	0	0	0	0	0	0	0	0	0	0
∞	−9327[d]		−4641[d]		−1806[d]		143[d]		58[d]	

[a] Unless otherwise stated, all singlet–triplet gaps are reported as errors relative to CCSDT in cm^{-1}.

[b] Equivalent to CCSD.

[c] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_D$].

[d] The CCSDT singlet–triplet gap in cm^{-1}.

Table 3.10 The total numbers of walkers, reported as percentages of the total walker populations at 200000 MC iterations, characterizing the *i*-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. that were needed to generate the CC(P) and CC($P;Q$) results for the X $^1\Sigma_g^+$ and A $^3\Sigma_u^+$ states of $(HFH)^-$ reported in Tables 3.7 and 3.8. Adapted from Ref. [102].

	$R_{\text{H-F}} = 1.50$ Å		$R_{\text{H-F}} = 1.75$ Å		$R_{\text{H-F}} = 2.00$ Å		$R_{\text{H-F}} = 2.50$ Å		$R_{\text{H-F}} = 4.00$ Å	
MC Iters.	X $^1\Sigma_g^+$	A $^3\Sigma_u^+$	X $^1\Sigma_g^+$	A $^3\Sigma_u^+$	X $^1\Sigma_g^+$	A $^3\Sigma_u^+$	X $^1\Sigma_g^+$	A $^3\Sigma_u^+$	X $^1\Sigma_g^+$	A $^3\Sigma_u^+$
0	0.02[a]	0.09[a]	0.01[a]	0.10[a]	0.01[a]	0.13[a]	0.01[a]	0.24[a]	0.00[a]	0.59[a]
2000	0.11	0.59	0.08	0.64	0.06	0.73	0.03	0.95	0.01	1.74
4000	0.20	1.08	0.15	1.14	0.10	1.23	0.05	1.54	0.02	2.45
6000	0.27	1.36	0.19	1.46	0.13	1.59	0.07	1.86	0.02	2.97
8000	0.32	1.55	0.23	1.66	0.16	1.77	0.08	2.12	0.03	3.20
10000	0.37	1.68	0.26	1.79	0.18	1.91	0.09	2.29	0.03	3.29
20000	0.55	2.19	0.40	2.38	0.28	2.46	0.15	3.03	0.06	4.09
50000	1.42	4.19	1.10	4.59	0.81	4.81	0.47	5.58	0.21	6.83
100000	6.21	12.24	5.16	12.87	4.25	13.61	2.88	15.13	1.61	16.64
150000	25.44	35.16	23.37	36.28	21.01	37.71	17.11	38.88	12.45	41.09
200000	100[b]	100[c]	100[d]	100[e]	100[f]	100[g]	100[h]	100[i]	100[j]	100[k]

[a] The initial walker population, meaning 1500 walkers on the RHF (X $^1\Sigma_g^+$ state) and ROHF (A $^3\Sigma_u^+$ state) reference determinants.
[b] The total number of walkers at 200000 MC iterations is 9865967.
[c] The total number of walkers at 200000 MC iterations is 1749699.
[d] The total number of walkers at 200000 MC iterations is 12468454.
[e] The total number of walkers at 200000 MC iterations is 1431689.
[f] The total number of walkers at 200000 MC iterations is 15510033.
[g] The total number of walkers at 200000 MC iterations is 1123676.
[h] The total number of walkers at 200000 MC iterations is 24265207.
[i] The total number of walkers at 200000 MC iterations is 632102.
[j] The total number of walkers at 200000 MC iterations is 50189301.
[k] The total number of walkers at 200000 MC iterations is 254390.

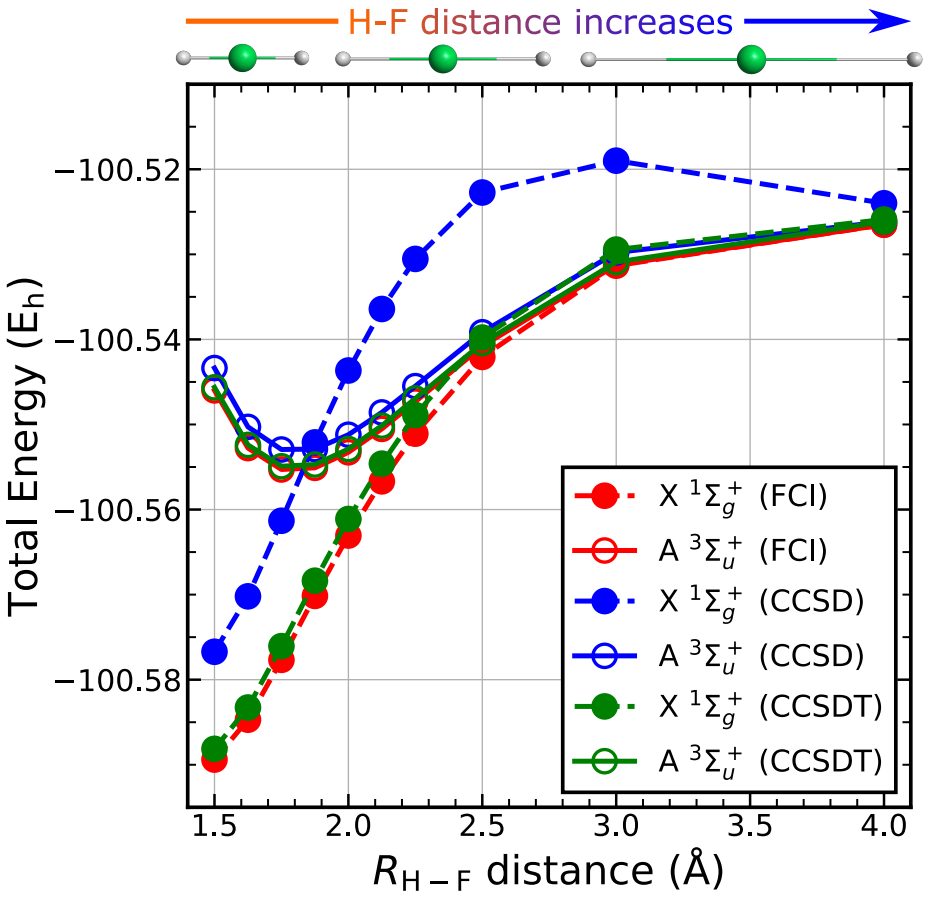

Figure 3.8 Total electronic energies of the X $^1\Sigma_g^+$ (open circles and solid line) and A $^3\Sigma_u^+$ (filled circles and dotted line) states of $(HFH)^-$ with increase in the H–F distance, from 1.5 Å to 4.0 Å, obtained from the FCI (red circles), CCSD (blue circles), and CCSDT (green circles) methods. Recreated from the data reported in Refs. [66, 84, 88, 197].

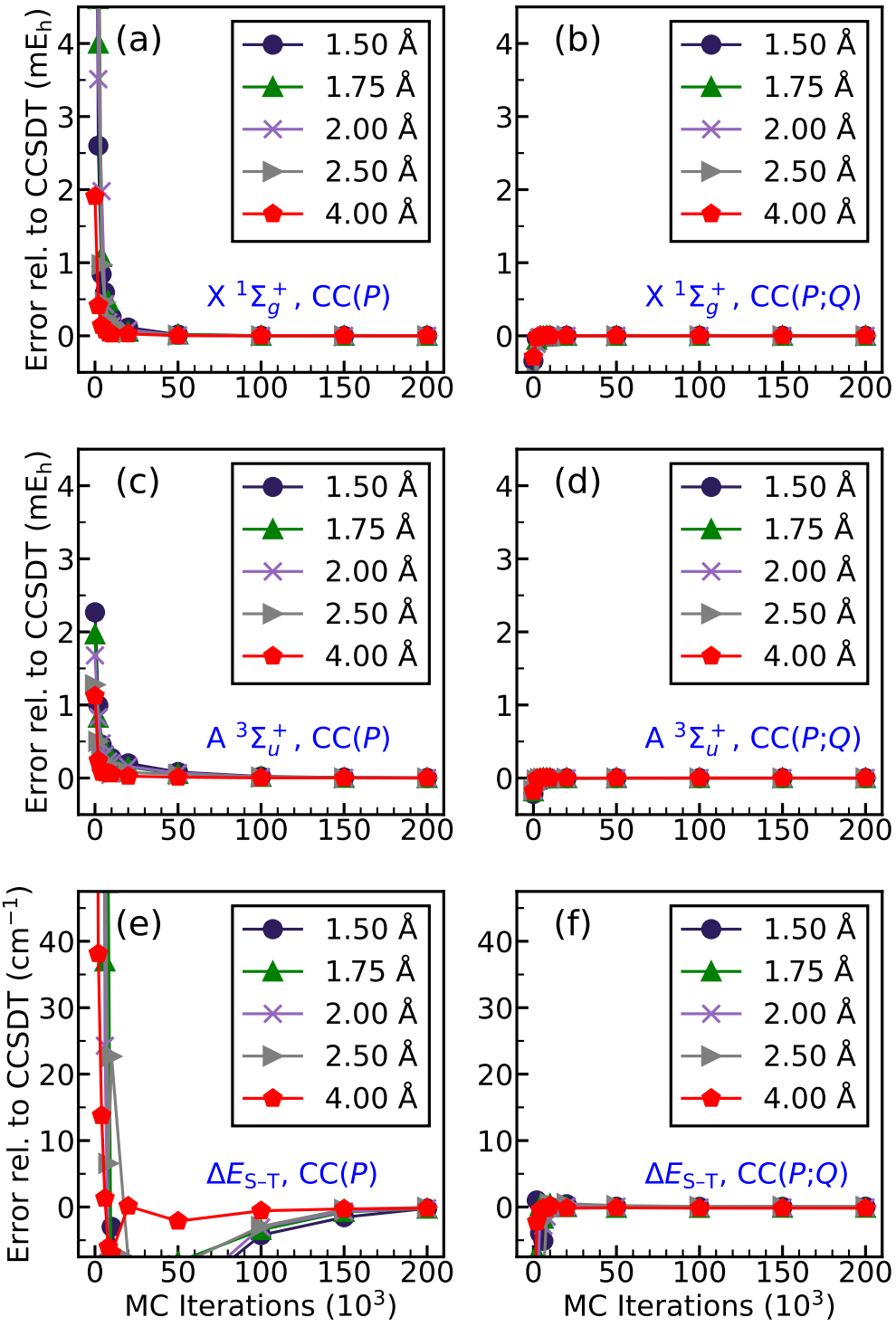

Figure 3.9 Convergence of the CC(P) and CC(P;Q) energies of the X $^1\Sigma_g^+$ [panels (a) and (b)] and A $^3\Sigma_u^+$ [panels (c) and (d)] states of $(HFH)^-$, as described by the 6-31G(d,p) basis set, and of the corresponding singlet–triplet gaps [panels (e) and (f)] toward their parent CCSDT values. The H–F distances $R_{H\text{-}F}$ used are 1.50 Å, 1.75 Å, 2.00 Å, 2.50 Å, and 4.00 Å. The P spaces consisted of all singles and doubles and subsets of triples identified during i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. and the Q spaces consisted of the triples not captured by i-FCIQMC. Adapted from Ref. [102].

3.3.3 Cyclobutadiene and cyclopentadienyl cation

We now proceed to the examination of the performance of the semi-stochastic CC(P;Q) algorithm in calculations involving medium-sized organic biradicals, starting from two prototypical anti-aromatic systems, cyclobutadiene and cyclopentadienyl cation, both described using the cc-pVDZ basis set. As in the rest of this chapter, we are mainly interested in how efficient the CIQMC-driven CC(P;Q) methodology is in recovering the CCSDT energies of the lowest singlet and triplet states and gaps between them. In the case of cyclobutadiene and cyclopentadienyl cation discussed in this subsection, we focus on examining vertical singlet–triplet gaps.

We begin with the FCIQMC-driven CC(P;Q) calculations for cyclobutadiene, in which we adopted the D_{4h}-symmetric geometry that represents the transition state for the automerization of cyclobutadiene proceeding on the lowest singlet potential, optimized with the MR average-quadratic CC (MR-AQCC) approach [200, 201] using the cc-pVDZ basis in Ref. [202]. We employed this geometry for two reasons. One of them is the fact that we used the same geometry in our earlier CIQMC- and CCMC-based [99, 101], CIPSI-driven [120], and active-orbital-based [87] CC(P;Q) calculations for cyclobutadiene, when examining its automerization. Because of this, we could verify the correctness of our FCIQMC-driven CC(P;Q) calculations for the lowest-energy singlet state, which is also the ground state of the system. Another is the observation that the D_{4h}-symmetric transition-state structure characterizing the automerization of cyclobutadiene is practically identical to the D_{4h}-symmetric minimum on the lowest triplet surface. Indeed, the MR-AQCC/cc-pVDZ C–C and C–H bond lengths defining the transition state on the ground-state singlet potential differ from those characterizing the triplet minimum optimized using unrestricted CCSD (UCCSD) in Ref. [203] by less than 0.009 and 0.001 Å, respectively.

At the D_{4h}-symmetric geometry used in our calculations, cyclobutadiene is characterized by the delocalization of four π electrons over four π MOs, which gives rise to the close-lying singlet and triplet states that require a highly accurate treatment of electron correlation

effects if we are to obtain a well-balanced description of the two states and the small energy separation between them. One can understand this by examining the valence π network of the D_{4h}-symmetric cyclobutadiene species, which consists of the doubly occupied nondegenerate a_{2u} orbital, the doubly degenerate e_g level, in which each component MO is occupied by a single electron, and the nondegenerate b_{1u} orbital, which in the zeroth-order description of the lowest singlet and triplet states remains empty. The two valence electrons in the degenerate e_g shell can couple to a singlet or a triplet, resulting in the open-shell singlet ground state, X $^1B_{1g}$, which has a substantial MR character, and the first excited triplet state, A $^3A_{2g}$, which is predominantly SR in nature (see Fig. 3.10 for an illustration of the π MO network of cyclobutadiene along with the triplet electronic configuration). In order to balance the substantial nondynamical correlation effects, needed for an accurate description of the low-spin X $^1B_{1g}$ state, with the dynamical correlations dominating its high-spin triplet A $^3A_{2g}$ companion within a conventional, particle-conserving, SRCC framework and produce reliable $\Delta E_{\text{S-T}}$ values for cyclobutadiene, which could compete with the high-accuracy *ab initio* data reported in Refs. [154, 155, 202–208], one has to consider robust treatments of the connected triply excited clusters, such as that offered by CCSDT [204, 208]. Indeed, full CCSDT, which is the target of this investigation, produces high-quality results for the lowest singlet and triplet states of the D_{4h}-symmetric cyclobutadiene system and the energy separation between them. For example, the $\Delta E_{\text{S-T}}$ value obtained in the CCSDT/cc-pVDZ calculations at the transition-state geometry used in the present study, of -4.8 kcal/mol, is practically identical to the results of the state-of-the-art DEA-EOMCC computations including the high-rank $4p$-$2h$ correlations on top of CCSD, reported in Refs. [154, 155], which give -5.0 kcal/mol when the cc-pVDZ basis set is employed (for similar recent observations regarding the reliability of full CCSDT in generating virtually exact singlet–triplet gap values for cyclobutadiene, see Ref. [208]). It is, therefore, interesting to explore if the semi-stochastic CC(P;Q) methodology investigated in this work is capable of converging the CCSDT results for the X $^1B_{1g}$ and A $^3A_{2g}$ states of cyclobutadiene and vertical gap between them out of the

early stages of CIQMC propagations.

The results of our FCIQMC-driven CC(P) and CC(P;Q) computations for cyclobutadiene are summarized in Table 3.11 and Fig. 3.11. In all of our calculations, starting with the stochastic i-FCIQMC steps and ending with the deterministic CC(P;Q) and CCSDT runs, we used the D_{2h} Abelian subgroup of the D_{4h} point group characterizing the cyclobutadiene's geometry adopted in this work. Consequently, the i-FCIQMC propagations for the X $^1B_{1g}$ and A $^3A_{2g}$ states were set up to converge the lowest states of the ^1A_g (D_{2h}) and $^3B_{1g}$ (D_{2h}) symmetries. Consistent with the CC(P) and CC(P;Q) runs that follow the i-FCIQMC steps and the accompanying CCSD, CR-CC(2,3), and CCSDT computations, the reference determinants used to initiate our i-FCIQMC propagations were the closed-shell, D_{2h}-adapted, RHF function obtained by placing two electrons on one of the e_g valence orbitals for the lowest-energy ^1A_g (D_{2h}) state and the high-spin ROHF determinant, adapted to D_{2h} as well, for the lowest $^3B_{1g}$ (D_{2h}) state. As a result, the lists of triply excited determinants extracted from the i-FCIQMC runs at the various propagation times $\tau > 0$, needed to define the P spaces for the CC(P) and CC(P;Q) computations, consisted of the $S_z = 0$ triples of the A_g (D_{2h}) symmetry for the X $^1B_{1g}$ state and the $S_z = 1$ triples of the B_{1g} (D_{2h}) symmetry in the case of the A $^3A_{2g}$ state. Given our interest in converging the CCSDT energetics, the Q spaces used to construct the $\delta(P;Q)$ corrections consisted of the remaining triply excited determinants, absent in the i-FCIQMC wave functions of the X $^1B_{1g}$ and A $^3A_{2g}$ states at a given τ.

The results shown in Table 3.11 and Fig. 3.11 display several similarities with the previously discussed methylene and $(HFH)^-$ cases. One cannot, for example, obtain an accurate description of the more MR singlet ground state and the energy separation between the X $^1B_{1g}$ and A $^3A_{2g}$ states without incorporating the leading triply excited determinants in the P space. Indeed, when the P space consists of only singly and doubly excited determinants, as in the $\tau = 0$ CC(P) (*i.e.*, CCSD) and CC(P;Q) [*i.e.*, CR-CC(2,3)] calculations, one ends up with the enormous errors in the energies of the X $^1B_{1g}$ state relative to their

CCSDT parent, which are 47.979 millihartree in the former case and 14.636 millihartree when the latter computation is considered. The $\tau = 0$ CC(P;Q) energy of the A ^{3}A$_{2g}$ state is a lot more accurate, reducing the large, 23.884 millihartree, error relative to CCSDT obtained in the underlying CC(P) calculation to -60 microhartree, but this does not help too much. The corresponding CR-CC(2,3) triples correction to CCSD, which neglects the coupling of the low-order T_1 and T_2 clusters with their higher-order T_3 counterpart, is incapable of offering a balanced description of the X ^{1}B$_{1g}$ and A ^{3}A$_{2g}$ states, so that the resulting singlet–triplet gap is very poor. The 9.2 kcal/mol difference between the $\Delta E_{\text{S-T}}$ values obtained in the $\tau = 0$ CC(P;Q) or CR-CC(2,3) and CCSDT calculations is so large that the X ^{1}B$_{1g}$ $-$ A ^{3}A$_{2g}$ separation predicted by CR-CC(2,3) has a wrong sign compared to its -4.8 kcal/mol CCSDT counterpart, while being nearly identical in magnitude. This difference becomes even larger when the uncorrected $\tau = 0$ CC(P), meaning CCSD, calculations are considered (15.1 kcal/mol).

As shown in Table 3.11 and Fig. 3.11, the situation dramatically changes when the P spaces used in the CC(P) and CC(P;Q) calculations are enriched with the subsets of triply excited determinants captured by the i-FCIQMC propagations. The convergence of the CC(P;Q) energies of the X ^{1}B$_{1g}$ and A ^{3}A$_{2g}$ states, especially the former ones, and the vertical separations between them is particularly impressive. For example, after as few as 6000 $\delta\tau = 0.0001$ a.u. MC time steps and i-FCIQMC capturing less than 30% of all triples in the P space, where, as demonstrated in Table 3.12, the walker population characterizing the i-FCIQMC run for the X ^{1}B$_{1g}$ state is only 0.02% of the total number of walkers at $\tau = 8.0$ a.u. (the termination time for our i-FCIQMC propagations for cyclobutadiene), the CC(P;Q) approach reduces the 14.636 millihartree difference between the CR-CC(2,3) and CCSDT energies of the strongly correlated singlet ground state to 2.223 millihartree. While the CR-CC(2,3) description of the A ^{3}A$_{2g}$ state, which has a largely SR character, is already excellent, the CC(P;Q) calculation performed after 6000 MC iterations, which uses only 26% of triples in the P space and a tiny walker population that amounts to 0.04% of all walkers at

$\tau = 8.0$ a.u. in the underlying i-FCIQMC propagation, improves it too, reducing the small, 60 microhartree, unsigned difference between the CR-CC(2,3) and CCSDT energies to an even smaller 51 microhartree. As a consequence of the above improvements, especially for the X $^1B_{1g}$ state, the error relative to CCSDT characterizing the $\Delta E_{\text{S-T}}$ value obtained in the FCIQMC-driven CC(P;Q) calculations after 6000 $\delta\tau = 0.0001$ a.u. MC time steps, where the underlying i-FCIQMC propagations are still in their early stages, is only 1.4 kcal/mol, as opposed to 9.2 kcal/mol obtained at $\tau = 0$ with CR-CC(2,3). The resulting X $^1B_{1g}$ – A $^3A_{2g}$ energy separation, of −3.4 kcal/mol, has not only the correct sign, but is also very close to the −4.8 kcal/mol value obtained with CCSDT. If we wait a little longer, by executing the extra 2000 MC iterations, so that the i-FCIQMC propagations can capture 34%–39% of all triply excited determinants, we can reduce the already small 2.223 millihartree, 51 microhartree, and 1.4 kcal/mol errors in the CC(P;Q) energies of the X $^1B_{1g}$ and A $^3A_{2g}$ states and separation between them relative to CCSDT, obtained after 6000 MC time steps, to 0.835 millihartree, 31 microhartree, and 0.5 kcal/mol, respectively. It is clear from Table 3.11 and Fig. 3.11 that the convergence of the semi-stochastic CC(P;Q) results for the lowest-energy singlet and triplet states of cyclobutadiene, especially the X $^1B_{1g}$ energies and the X $^1B_{1g}$ – A $^3A_{2g}$ gap values, which the $\tau = 0$ CC(P;Q) or CR-CC(2,3) calculations describe poorly, toward CCSDT is very fast, even when the underlying i-FCIQMC propagations are far from convergence. It is also apparent from our calculations that the noniterative corrections $\delta(P;Q)$ play a significant role in accelerating convergence of the corresponding CC(P) energetics toward CCSDT. As shown, for example, in Table 3.11, the relatively large differences between the uncorrected CC(P) energies of the X $^1B_{1g}$ and A $^3A_{2g}$ states and vertical gap between them obtained at $\tau = 0.8$ a.u., *i.e.*, after 8000 $\delta\tau = 0.0001$ a.u. MC iterations, and the corresponding CCSDT data, which exceed 11 and 7 millihartree and 3 kcal/mol, respectively, are reduced to 0.835 millihartree, 31 microhartree, and 0.5 kcal/mol, when the CC(P;Q) approach is employed. We can see similar improvements in the CC(P) energies at other τ values.

Table 3.11 Convergence of the CC(P) and CC(P;Q) energies of the X ^{1}B$_{1g}$ and A ^{3}A$_{2g}$ states of cyclobutadiene, as described by the cc-pVDZ basis set, and of the corresponding vertical singlet–triplet gaps toward their parent CCSDT values. All calculations were performed at the D$_{4h}$-symmetric transition-state geometry of the X ^{1}B$_{1g}$ state optimized in the MR-AQCC calculations reported in Ref. [202]. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-FCIQMC runs. The i-FCIQMC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the RHF (X ^{1}B$_{1g}$ state) and ROHF (A ^{3}A$_{2g}$ state) reference determinants and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the four lowest core orbitals were kept frozen and the spherical components of d orbitals were employed throughout. Adapted from Ref. [102].

	X ^{1}B$_{1g}$			A ^{3}A$_{2g}$			X ^{1}B$_{1g}$ − A ^{3}A$_{2g}$	
MC Iterations	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[c]	$(P;Q)$[c]
0	47.979[d]	14.636[e]	0	23.884[d]	-0.060[e]	0	15.1[d]	9.2[e]
2000	40.663	11.059	3.5	21.004	0.031	3.0	12.3	6.9
4000	27.235	5.921	16.6	14.317	0.068	14.2	8.1	3.7
6000	17.188	2.223	29.5	10.016	0.051	25.5	4.6	1.4
8000	11.207	0.835	39.2	7.463	0.031	34.3	3.3	0.5
10000	8.299	0.429	46.6	5.865	0.020	41.0	1.5	0.3
20000	2.030	0.013	70.0	2.461	0.005	62.8	-0.3	0.0
50000	0.049	0.000	96.9	0.166	0.000	94.2	-0.1	0.0
80000	0.001	0.000	99.9	0.009	0.000	99.6	0.0	0.0
∞	-154.232002[f]			-154.224380[f]			-4.8[g]	

[a] Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

[b] The %T values are the percentages of triples captured during the i-FCIQMC propagations [the $S_z = 0$ triply excited determinants of the A$_g$(D$_{2h}$) symmetry in the case of the X ^{1}B$_{1g}$ state and the $S_z = 1$ triply excited determinants of the B$_{1g}$ (D$_{2h}$) symmetry in the case of the A ^{3}A$_{2g}$ state].

[c] Unless otherwise specified, the values of the singlet–triplet gaps are reported as errors relative to CCSDT in kcal/mol.

[d] Equivalent to CCSD.

[e] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_{\rm D}$].

[f] Total CCSDT energy in hartree.

[g] The CCSDT singlet–triplet gap in kcal/mol.

Table 3.12 The total numbers of walkers, reported as percentages of the total walker populations at 80000 MC iterations, characterizing the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. that were needed to generate the CC(P) and CC(P;Q) results for cyclobutadiene reported in Table 3.11. Adapted from Ref. [102].

MC Iterations	X $^1B_{1g}$	A $^3A_{2g}$
0	0.00[a]	0.00[a]
2000	0.00	0.00
4000	0.01	0.02
6000	0.02	0.04
8000	0.03	0.07
10000	0.05	0.09
20000	0.16	0.28
50000	3.81	4.93
80000	100[b]	100[c]

[a] The initial walker population, meaning 1500 walkers on the RHF (X $^1B_{1g}$ state) and ROHF (A $^3A_{2g}$ state) reference determinants.
[b] The total number of walkers at 80000 MC iterations is 8457504823.
[c] The total number of walkers at 80000 MC iterations is 4067481034.

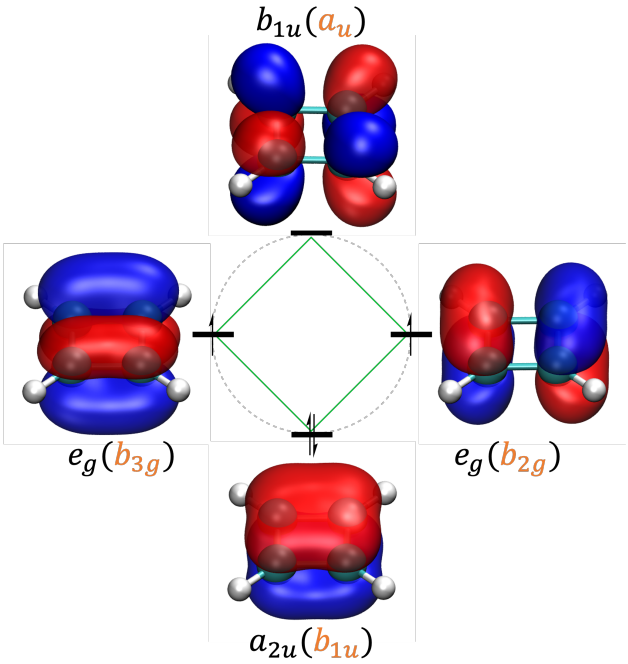

Figure 3.10 π molecular orbital network of the cyclobutadiene molecule, obtained at the HF/cc-pVDZ level, at the D_{4h}-symmetric transition-state geometry of the X $^1B_{1g}$ state optimized in the MR-AQCC calculations in Ref. [202]. The orbital irreducible representations in the D_{4h} symmetry are shown in black and the corresponding labels in the C_{2v} symmetry are shown in the parenthesis in orange. This electronic configuration refers to the triplet state.

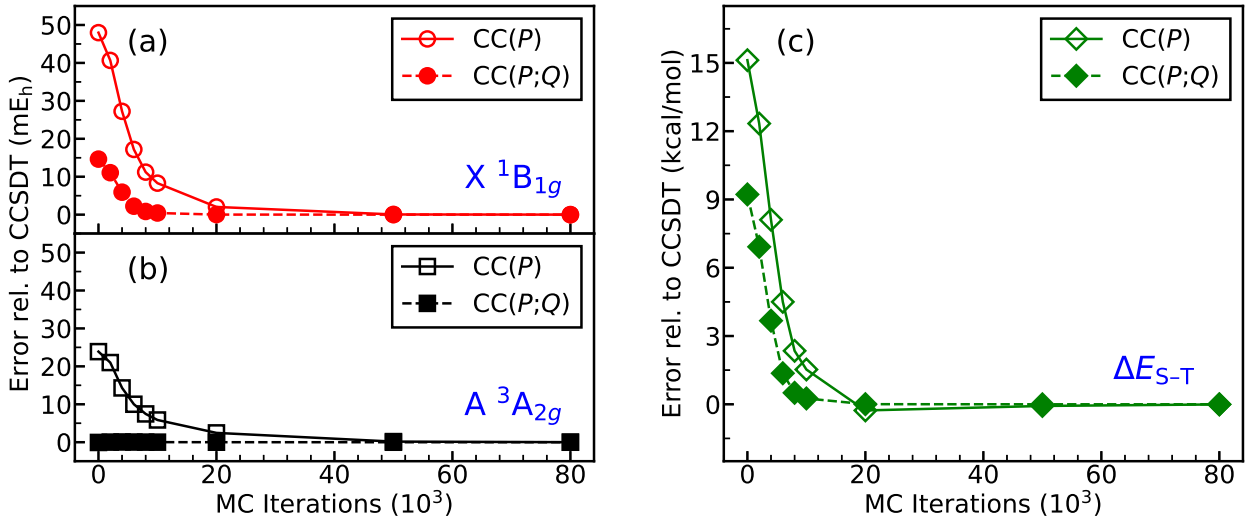

Figure 3.11 Convergence of the CC(P) and CC(P;Q) energies of the X ^{1}B$_{1g}$ [panel (a)] and A ^{3}A$_{2g}$ [panel (b)] states of cyclobutadiene, as described by the cc-pVDZ basis set, and of the corresponding vertical singlet–triplet gaps [panel (c)] toward their parent CCSDT values. All calculations were performed at the D$_{4h}$-symmetric transition-state geometry of the X ^{1}B$_{1g}$ state optimized in the MR-AQCC calculations in Ref. [202]. The P spaces consisted of all singles and doubles and subsets of triples identified during the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. and the Q spaces consisted of the triples not captured by i-FCIQMC. Adapted from Ref. [102].

Most of the observations regarding the performance of the semi-stochastic CC(P;Q) methodology and its CC(P) counterpart remain valid when the larger cyclopentadienyl cation, which is also the largest molecular system considered in our CC(P)/CC(P;Q) work to date, is examined. Following the previous DEA-EOMCC studies of cyclopentadienyl cation from the Piecuch group [154, 155], where the effect of high-order 4p-2h correlations on the singlet–triplet gap was investigated, we used the D$_{5h}$-symmetric geometry corresponding to a minimum on the lowest triplet surface obtained in the UCCSD/cc-pVDZ optimization in Ref. [203]. At this geometry, cyclopentadienyl cation is characterized by the delocalization of four π electrons over five π MOs, resulting in the doubly occupied nondegenerate a_2'' orbital, the doubly degenerate e_1'' shell, in which each component MO is occupied by a single electron, and the doubly degenerate e_2'' shell, which in the zeroth-order description of the lowest-energy singlet and triplet states remains empty. In analogy to the previously discussed cyclobutadiene system, the two electrons in the degenerate e_1'' MOs can couple to a singlet

or triplet, but compared to cyclobutadiene, where the lowest-energy singlet state is also a ground state, the state ordering in cyclopentadienyl cation is reversed, so that the lowest triplet, designated as $\mathrm{X}\,^3\mathrm{A}_2'$, is the ground state and the lowest-energy singlet, denoted as $\mathrm{A}\,^1\mathrm{E}_2'$, is the first excited state (see Fig. 3.12 for an illustration of the π MOs). Similar to all other examples considered in this work, in order to obtain a well-balanced description of the $\mathrm{X}\,^3\mathrm{A}_2'$ state, which has a SR character dominated by dynamical correlations, and its $\mathrm{A}\,^1\mathrm{E}_2'$ companion, which is an open-shell singlet characterized by significant MR correlations, and obtain an accurate value of $\Delta E_{\mathrm{S\text{-}T}}$ within a conventional SRCC framework, one must turn to higher-level theories that can offer a robust treatment of T_n clusters with $n > 2$. Otherwise, as shown in Ref. [203], and as confirmed in our calculations, the results can be very poor. For example, the $\mathrm{A}\,^1\mathrm{E}_2' - \mathrm{X}\,^3\mathrm{A}_2'$ separation in cyclopentadienyl cation resulting from the restricted CCSD calculations using the cc-pVDZ basis, which are equivalent to our $\tau = 0$ CC(P) computations, is about 23 kcal/mol. This is in large disagreement with the most accurate *ab initio* calculations of the singlet–triplet gap in cyclopentadienyl cation performed to date using the DEA-EOMCC formalism including $3p$-$1h$ as well as $4p$-$2h$ correlations on top of the CCSD treatment of the underlying closed-shell core, which give about 14 kcal/mol when the cc-pVDZ basis set is employed [154, 155] (for the examples of other high-level SRCC and MRCC calculations of the singlet–triplet gap in cyclopentadienyl cation, see Ref. [203]; Ref. [155] also provides the well-converged MR perturbation theory data, which agree with the state-of-the-art DEA-EOMCC computations reported in Refs. [154, 155]). The restricted CCSDT approach, which is the target SRCC method in this study, provides a much better description, reducing the approximately 9 kcal/mol error relative to the most accurate DEA-EOMCC calculations with up to $4p$-$2h$ excitations reported in Refs. [154, 155] obtained with restricted CCSD to less than 3 kcal/mol, when the cc-pVDZ basis set is employed. It would certainly be interesting to examine if the inclusion of higher–than–triply excited clusters, such as T_4, could further improve the CCSDT description of the singlet–triplet gap in cyclopentadienyl cation, but in this work we focus on the ability of the

semi-stochastic, CIQMC-based, CC(P;Q) methodology to improve the CR-CC(2,3) $\Delta E_{\text{S-T}}$ values and converge the results of CCSDT computations. We hope to return to the topic of the role of T_4 clusters in describing the singlet–triplet gap in cyclopentadienyl cation in one of our future studies. It may be worth pointing out that the A $^1\text{E}_2' - \text{X}\ ^3\text{A}_2'$ gap obtained in the restricted CCSDT/cc-pVDZ calculations, which give $\Delta E_{\text{S-T}} = 16.7$ kcal/mol, is in very good agreement with the 16.1 kcal/mol resulting from the DEA-EOMCC/cc-pVDZ computations truncated at 3p-1h excitations [154, 155].

The results of our CIQMC-driven CC(P) and CC(P;Q) computations for cyclopentadienyl cation are reported in Table 3.13 and Fig. 3.13. As already alluded to above, to reduce the computational costs of the CIQMC propagations preceding the CC(P) and CC(P;Q) steps, especially in the later stages of the CIQMC runs that are included in Table 3.13 and Fig. 3.13 for the completeness of our presentation, we replaced the i-FCIQMC algorithm, which we exploited in our calculations for methylene, $(\text{HFH})^-$, and cyclobutadiene, by its truncated i-CISDTQ-MC counterpart. It has been established in Ref. [101] that the replacement of i-FCIQMC by i-CISDTQ-MC, when identifying the leading higher–than–doubly excited determinants for the inclusion in the P spaces used in the semi-stochastic CC(P) and CC(P;Q) runs, has virtually no effect on the rate at which these runs converge the parent SRCC energetics. In analogy to cyclobutadiene, all of our i-CISDTQ-MC, semi-stochastic CC(P) and CC(P;Q), and deterministic CCSD, CR-CC(2,3), and CCSDT computations utilized the largest Abelian subgroup of the D_{5h} point group characterizing the cyclopentadienyl cation's structure examined in the present study, which is C_{2v}. This means that in setting up our calculations for the X $^3\text{A}_2'$ state, we treated it as the lowest state of the $^3\text{B}_2$ (C_{2v}) symmetry, whereas the doubly degenerate A $^1\text{E}_2'$ state was represented by its $^1\text{A}_1$ (C_{2v}) component. Similar to cyclobutadiene, and to remain consistent with the CC(P), CC(P;Q), and other SRCC runs for cyclopentadienyl cation carried out in this study, the reference determinant used to initiate the i-CISDTQ-MC propagation for the lowest-energy $^3\text{B}_2$ (C_{2v}) state was the triplet ROHF determinant. In the case of the $^1\text{A}_1$ (C_{2v}) component of the

A $^1E_2'$ state, we used the RHF determinant obtained by pairing the two valence electrons in one of the e_1'' MOs to initiate the corresponding i-CISDTQ-MC run. Consistent with the above description, the subsets of triply excited determinants used to construct the P spaces for the semi-stochastic CC(P) and CC(P;Q) computations for the X $^3A_2'$ state were the $S_z = 1$ triples of the B_2 (C_{2v}) symmetry captured by i-CISDTQ-MC. In the case of the A $^1E_2'$ state, represented, as explained above, by its $^1A_1$ (C_{2v}) component, we used the $S_z = 0$ triples of the A_1 (C_{2v}) symmetry identified by the i-CISDTQ-MC propagation set up to converge the lowest A_1 (C_{2v}) state. As usual, the corresponding Q spaces were spanned by the remaining triply excited determinants that were not captured by the i-CISDTQ-MC runs when the lists of P-space triples were created.

Our calculations for cyclopentadienyl cation, summarized in Table 3.13 and Fig. 3.13, demonstrate that the CC(P;Q) energies of the X $^3A_2'$ and A $^1E_2'$ states and vertical gaps between them display fast convergence toward the respective CCSDT values with the propagation time τ. This is particularly apparent in the case of the CC(P;Q) energies of the more MR A $^1E_2'$ state and the A $^1E_2'$ – X $^3A_2'$ separation, which cannot be accurately described if the underlying P spaces contain only singly and doubly excited determinants. Indeed, the CR-CC(2,3) energy of the A $^1E_2'$ state, which is equivalent to the $\tau = 0$ CC(P;Q) value, is much more accurate than the result of the associated CC(P) or CCSD calculation, which produces the enormous error relative to CCSDT exceeding 38 millihartree, but the substantial, > 6 millihartree, difference with the CCSDT energy remains. The situation for the SR X $^3A_2'$ state, where the CR-CC(2,3) approach reduces the nearly 29 millihartree error relative to CCSDT obtained in the CCSD calculations to $\sim$0.2 millihartree, is a lot better, but this does not help the resulting $\Delta E_{\text{S-T}}$ value, which differs from its CCSDT counterpart by almost 4 kcal/mol (almost a quarter of the CCSDT value of $\Delta E_{\text{S-T}}$). The discrepancy between the errors in the CR-CC(2,3) energies of the X $^3A_2'$ and A $^1E_2'$ states is simply too large. Clearly, one needs to incorporate some triples in the corresponding P spaces, especially in the case of the more challenging A $^1E_2'$ state.

Once the $\tau = 0$ P spaces are augmented with the leading triply excited determinants identified by the i-CISDTQ-MC propagations and the noniterative corrections $\delta(P;Q)$ are added to the CC(P) energies to estimate the effects of the remaining T_3 correlations, we observe smooth convergence of the resulting CC(P;Q) energetics toward their respective CCSDT limits. This includes significant improvements in the poor description of the A $^1\mathrm{E}'_2$ state and the A $^1\mathrm{E}'_2$ – X $^3\mathrm{A}'_2$ separation by CR-CC(2,3). As shown in Table 3.13, already after 10000 $\delta\tau = 0.0001$ a.u. MC time steps, where the i-CISDTQ-MC propagations are still in their infancy, capturing only 25–30% of all triples and using tiny walker populations, on the order of 0.1–0.2% of the total numbers of walkers at $\tau = 8.0$ a.u. (see Table 3.14), the 6.245 millihartree and 3.8 kcal/mol errors in the energy of the A $^1\mathrm{E}'_2$ state and the $\Delta E_{\text{S-T}}$ value relative to CCSDT obtained with CR-CC(2,3) reduce in the CC(P;Q) calculations to 2.248 millihartree and 1.3 kcal/mol, respectively. By running i-CISDTQ-MC a little longer and capturing about 50–60% of all triples in the relevant P spaces, as is the case after 20000 MC iterations, where the walker populations compared to $\tau = 8.0$ a.u. are still tiny, the errors in the CC(P;Q) values of the A $^1\mathrm{E}'_2$ energy and $\Delta E_{\text{S-T}}$ relative to their CCSDT parents drop down by an order of magnitude compared to 10000 MC iterations, to 0.217 millihartree and 0.1 kcal/mol, respectively. Although the excellent description of the predominantly SR X $^3\mathrm{A}'_2$ state by the CR-CC(2,3) approach hardly needs any improvement, the i-CISDTQ-MC-driven CC(P;Q) calculations are helping here too, reducing the 0.245 millihartree difference between the CR-CC(2,3) and CCSDT energies to 0.108 millihartree after 10000 MC iterations (26 microhartree when the number of MC iterations is increased to 20000). As anticipated, the uncorrected CC(P) energies of the X $^3\mathrm{A}'_2$ and A $^1\mathrm{E}'_2$ states converge to the respective CCSDT limits too, but they do it at a much slower pace than their CC(P;Q) counterparts. A comparison of the results of the CC(P) and CC(P;Q) calculations for the A $^1\mathrm{E}'_2$ – X $^3\mathrm{A}'_2$ gap shown in Table 3.13 and Fig. 3.13 (c) may create an impression as if the noniterative corrections $\delta(P;Q)$ offer very little, but this would be misleading. The relatively fast convergence of the CC(P) values of $\Delta E_{\text{S-T}}$ toward their CCSDT parent in

the early stages of the underlying i-CISDTQ-MC propagations, which compares well with that observed in the corresponding CC(P;Q) computations, is a result of the fortuitous cancellation of large errors characterizing the CC(P) energies of the X ^{3}A$_2'$ and A ^{1}E$_2'$ states. Since no other system examined in this study displays similar error cancellations, and since costs of computing corrections $\delta(P;Q)$, which offer major error reductions in the individual CC(P) energies, while accelerating their convergence toward the SRCC target, are low, we recommend using the $\delta(P;Q)$-corrected CC(P;Q) energetics.

Table 3.13 Convergence of the CC(P) and CC(P;Q) energies of the X ^{3}A$_2'$ and A ^{1}E$_2'$ states of cyclopentadienyl cation, as described by the cc-pVDZ basis set, and of the corresponding vertical singlet–triplet gaps toward their parent CCSDT values. All calculations were performed at the D_{5h}-symmetric geometry of the X ^{3}A$_2'$ state optimized using the unrestricted CCSD/cc-pVDZ approach reported in Ref. [203]. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-CISDTQ-MC runs. The i-CISDTQ-MC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the ROHF (X ^{3}A$_2'$ state) and RHF (A ^{1}E$_2'$ state) reference determinants and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the five lowest core orbitals were kept frozen and the spherical components of d orbitals were employed throughout. Adapted from Ref. [102].

	X ^{3}A$_2'$			A ^{1}E$_2'$			A ^{1}E$_2'$ − X ^{3}A$_2'$	
MC Iterations	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[c]	$(P;Q)$[c]
0	28.840[d]	0.245[e]	0	38.572[d]	6.245[e]	0	6.1[d]	3.8[e]
2000	27.396	0.272	0.8	35.598	5.948	1.0	5.1	3.6
4000	22.253	0.267	5.1	27.946	5.078	6.5	3.6	3.0
6000	17.394	0.212	11.6	21.124	3.971	14.7	2.3	2.4
8000	13.743	0.152	18.3	16.042	2.756	23.0	1.4	1.6
10000	11.027	0.108	24.8	12.947	2.248	30.9	1.2	1.3
20000	4.250	0.026	52.1	3.964	0.217	61.4	-0.2	0.1
50000	0.155	0.001	95.3	0.060	0.001	98.3	-0.1	0.0
80000	0.007	0.000	99.8	0.001	0.000	100.0	0.0	0.0
∞	−192.615924[f]			−192.589235[f]			16.7[g]	

[a] Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

[b] The %T values are the percentages of triples captured during the i-CISDTQ-MC propagations [the $S_z = 1$ triply excited determinants of the B_2 (C_{2v}) symmetry in the case of the X ^{3}A$_2'$ state and the $S_z = 0$ triply excited determinants of the A_1 (C_{2v}) symmetry in the case of the A ^{1}E$_2'$ state].

[c] Unless otherwise specified, the values of the singlet–triplet gaps are reported as errors relative to CCSDT in kcal/mol.

[d] Equivalent to CCSD.

[e] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_D$].

[f] Total CCSDT energy in hartree.

[g] The CCSDT singlet–triplet gap in kcal/mol.

Table 3.14 The total numbers of walkers, reported as percentages of the total walker populations at 80000 MC iterations, characterizing the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. that were needed to generate the CC(P) and CC(P;Q) results for the cyclopentadienyl cation reported in Table 3.13. Adapted from Ref. [102].

MC Iterations	X ^{3}A$'_2$	A ^{1}E$'_2$
0	0.00[a]	0.00[a]
2000	0.01	0.00
4000	0.03	0.02
6000	0.06	0.05
8000	0.11	0.09
10000	0.16	0.13
20000	0.62	0.54
50000	13.20	15.73
80000	100[b]	100[c]

[a] The initial walker population, meaning 1500 walkers on the ROHF (X ^{3}A$'_2$ state) and RHF (A ^{1}E$'_2$ state) reference determinants.
[b] The total number of walkers at 80000 MC iterations is 7867091953.
[c] The total number of walkers at 80000 MC iterations is 11371381724.

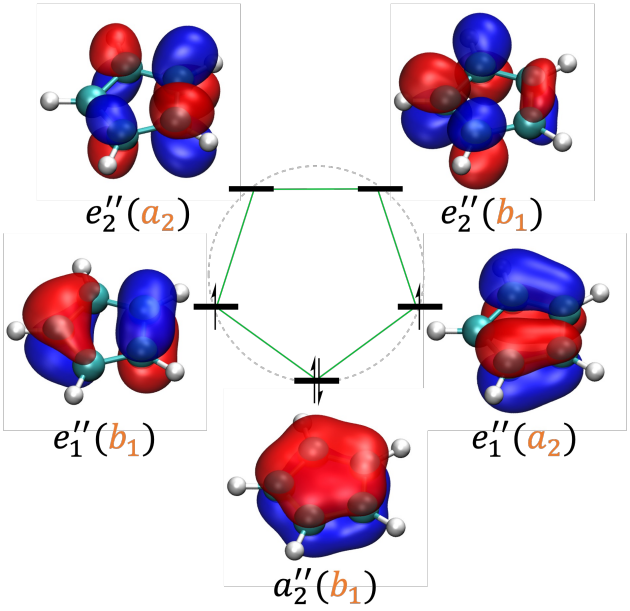

Figure 3.12 π molecular orbital network of the cyclopentadienyl cation molecule, obtained at the HF/cc-pVDZ level, at the D_{5h}-symmetric geometry of the X ^{3}A$'_2$ state optimized using the unrestricted CCSD/cc-pVDZ approach in Ref. [203]. The orbital irreducible representations in the D_{5h} symmetry are shown in black and the corresponding labels in the C_{2v} symmetry are shown in the parenthesis in orange. This electronic configuration refers to the triplet state.

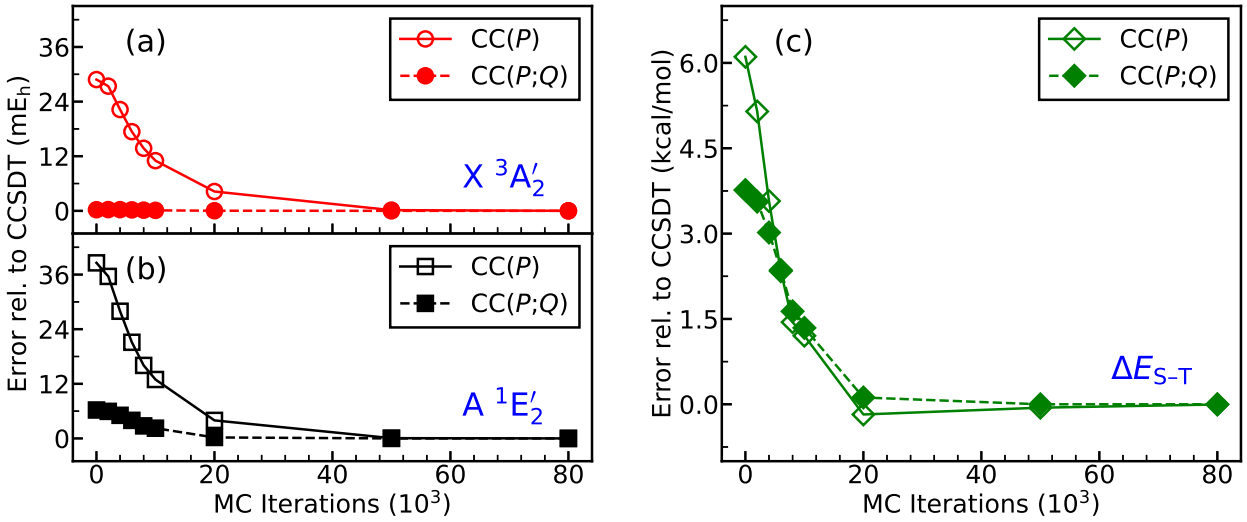

Figure 3.13 Convergence of the CC(P) and CC(P;Q) energies of the X ^{3}A$'_2$ [panel (a)] and A ^{1}E$'_2$ [panel (b)] states of cyclopentadienyl cation, as described by the cc-pVDZ basis set, and of the corresponding vertical singlet–triplet gaps [panel (c)] toward their parent CCSDT values. All calculations were performed at the D$_{5h}$-symmetric geometry of the X ^{3}A$'_2$ state optimized using the unrestricted CCSD/cc-pVDZ approach in Ref. [203]. The P spaces consisted of all singles and doubles and subsets of triples identified during the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. and the Q spaces consisted of the triples not captured by i-CISDTQ-MC. Adapted from Ref. [102].

3.3.4 Trimethylenemethane

Our final example is trimethylenemethane, a fascinating non-Kekulé hydrocarbon examined as early as in 1948 [209] and 1950 [210], in which four valence π electrons are delocalized over four closely spaced π-type orbitals. Assuming the D$_{3h}$ symmetry, which is the symmetry of the minimum-energy structure on the ground-state triplet surface of trimethylenemethane, the four MOs of this system's valence π network consist of the nondegenerate $1a''_2$ orbital, the doubly degenerate $1e''$ shell, and the nondegenerate $2a''_2$ orbital. If one adopts the C$_{2v}$ symmetry, relevant to the low-lying singlet states, which is also the largest Abelian subgroup of D$_{3h}$ exploited in our CCSD, CR-CC(2,3), CCSDT, and CIQMC-driven CC(P) and CC(P;Q) computations, the nondegenerate $1a''_2$ and $2a''_2$ orbitals in a D$_{3h}$ description become the $1b_1$ and $3b_1$ MOs, respectively, whereas the degenerate $1e''$ shell splits into the $1a_2$ and $2b_1$ components (see Fig. 3.14 for an illustration of the π MOs).

The first experimental identification of trimethylenemethane dates back to 1966 [211],

a definitive experimental verification, using electron paramagnetic resonance, of its triplet ground state was accomplished already in 1976 [212], and the electronic structure of trimethylenemethane has been well understood for decades (*cf.*, *e.g.*, Ref. [213] and references therein), but an accurate characterization of its triplet ground state and low-lying singlet states and energy separations between them continues to present a significant challenge to quantum chemistry approaches [80, 88, 153, 154, 214–239]. The D_{3h}-symmetric triplet ground state, designated as $X\,^3A_2'$ (in a C_{2v} description adopted in this study, $X\,^3B_2$), which is dominated by the $|\{\text{core}\}(1a_2'')^2(1e_1'')^1(1e_2'')^1|$ configuration (in C_{2v}, $|\{\text{core}\}(1b_1)^2(1a_2)^1(2b_1)^1|$), is relatively easy to describe, but the next two states, which are the nearly degenerate singlets stabilized by the Jahn–Teller distortion that lifts their exact degeneracy in a D_{3h} description, are not. The lower of the two singlets, which is characterized by a C_s-symmetric minimum that can be approximated by a twisted C_{2v} structure and which is, therefore, usually designated as the $A\,^1B_1$ state, is an open-shell singlet that emerges from the $|\{\text{core}\}(1b_1)^2(1a_2)^1(2b_1)^1|$ configuration. The second singlet, labeled as the $B\,^1A_1$ state, is a C_{2v}-symmetric multiconfigurational state dominated by the $|\{\text{core}\}(1b_1)^2(1a_2)^2|$ and $|\{\text{core}\}(1b_1)^2(2b_1)^2|$ closed-shell determinants. The $A\,^1B_1$ state, although lower in energy compared to its $B\,^1A_1$ counterpart, has not been observed experimentally due to unfavorable Franck–Condon factors [229, 240], so we do not consider it in this work. However, the second singlet, $B\,^1A_1$, has been detected in photoelectron spectroscopy experiments reported in Refs. [240, 241], which located it at 16.1 ± 0.1 kcal/mol above the $X\,^3A_2'$ ground state. Thus, following our previous deterministic, active-orbital-based, CC(P;Q) work [88] and the state-of-the-art DEA- and DIP-EOMCC computations with up to $4p$-$2h$ and $4h$-$2p$ excitations reported in Refs. [80, 153, 154], in carrying out the CIQMC-driven CC(P) and CC(P;Q) calculations discussed in this subsection and executing the accompanying CCSD, CR-CC(2,3), and CCSDT runs, we focused on the D_{3h}-symmetric triplet ground state, $X\,^3A_2'$, the C_{2v}-symmetric $B\,^1A_1$ singlet, and the adiabatic gap between them, adopting the geometries of the two states optimized using the spin-flip density functional theory (SF-DFT) and the 6-31G(d) basis in

Ref. [231]. In analogy to other organic biradicals discussed in this chapter, we employed the cc-pVDZ basis set, so that the parent CCSDT computations, needed to judge the performance of our semi-stochastic CC(P) and CC(P;Q) methods, and the more expensive CC(P) and CC(P;Q) calculations employing large, near-100%, fractions of triples in the relevant P spaces (captured in the later stages of the underlying CIQMC propagations) were not too difficult to execute on the computers available to us. As shown in our earlier deterministic CC(P;Q) work [88], in which we tested the active-orbital-based CC(t;3) method, which recovers the CCSDT energetics to within small fractions of kilocalorie per mole, and as confirmed by the authors of Ref. [237], who managed to perform the CCSDT/cc-pVTZ calculations, the use of a larger cc-pVTZ basis changes the adiabatic $\mathrm{B}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{A}_2'$ gap by about 0.5–1 kcal/mol, *i.e.*, the use the cc-pVDZ basis is sufficient to draw meaningful conclusions regarding the performance of the semi-stochastic CC(P) and CC(P;Q) approaches.

While the main goal of this study is to examine the efficiency of the CIQMC-driven CC(P;Q) approaches in converging the CCSDT energetics, it is worth pointing out that the parent CCSDT calculations using the ROHF reference determinant for the $\mathrm{X}\,^3\mathrm{A}_2'$ state and the RHF reference for the more strongly correlated $\mathrm{B}\,^1\mathrm{A}_1$ state, in spite of their SR character, are capable of producing a reasonably accurate description of the adiabatic $\mathrm{B}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{A}_2'$ separation in trimethylenemethane. Indeed, the purely electronic $\mathrm{B}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{A}_2'$ gap, designated, in analogy to other singlet–triplet gaps considered in this work, as $\Delta E_{\mathrm{S\text{-}T}}$, resulting from the ROHF/RHF-based CCSDT/cc-pVDZ computations using the SF-DFT/6-31G(d) geometries of the $\mathrm{X}\,^3\mathrm{A}_2'$ and $\mathrm{B}\,^1\mathrm{A}_1$ states optimized in Ref. [231] is 21.7 kcal/mol [88] (*cf.* Table 3.15). The corresponding experimentally derived result, obtained by subtracting the zero-point vibrational energy correction ΔZPVE resulting from the SF-DFT/6-31G(d) calculations reported in Ref. [231] from the experimental $\mathrm{B}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{A}_2'$ gap determined in Refs. [240, 241], is 18.1 kcal/mol. The CCSDT/cc-pVDZ value of $\Delta E_{\mathrm{S\text{-}T}}$ is not as accurate as the electronic $\mathrm{B}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{A}_2'$ gaps generated in the high-level DEA- and DIP-EOMCC calculations with the explicit inclusion of $4p$-$2h$ and $4h$-$2p$ correlations on top of CCSD,

which produce 18–19 kcal/mol [80, 153, 154], but it is certainly much better than 46.1, 24.4, and 29.8 kcal/mol obtained with the ROHF/RHF-based CCSD, CCSD(T), and CR-CC(2,3) methods, respectively, when the cc-pVDZ basis set is employed [88] [as demonstrated in Ref. [88], the use of a larger cc-pVTZ basis makes the CCSD, CCSD(T), and CR-CC(2,3) results even worse; the CCSD/cc-pVDZ and CR-CC(2,3)/cc-pVDZ values of $\Delta E_{\text{S-T}}$ are included in Table 3.15 as the $\tau = 0$ CC(P) and CC(P;Q) data, respectively]. While much of the 3.6 kcal/mol difference between the electronic $\text{B}\,^1\text{A}_1 - \text{X}\,^3\text{A}_2'$ separation obtained in the ROHF/RHF-based CCSDT/cc-pVDZ calculations and its experimentally derived estimate of 18.1 kcal/mol determined in Ref. [231] is, most likely, a consequence of the neglect of T_4 clusters in the CCSDT approach, we should keep in mind that the latter estimate depends on the source of the information about the ΔZPVE correction. For example, if one replaces the ΔZPVE value obtained in the SF-DFT/6-31G(d) calculations reported in Ref. [231] by its CCSD(T)/6-311++G(2d,2p) estimate and accounts for the core polarization effects determined with the help of the CCSD(T)/cc-pCVQZ computations, combining the resulting information with the experimental $\text{B}\,^1\text{A}_1 - \text{X}\,^3\text{A}_2'$ separation determined in Refs. [240, 241], the purely electronic, experimentally derived, adiabatic $\Delta E_{\text{S-T}}$ gap increases to 19.4 kcal/mol [237], which differs from our CCSDT/cc-pVDZ result by 2.3 kcal/mol. On the other hand, as shown in Ref. [237], the CCSDT value of the adiabatic $\text{B}\,^1\text{A}_1 - \text{X}\,^3\text{A}_2'$ gap increases with the basis set too, to 23.1 kcal/mol when the cc-pVTZ basis is employed, which reinforces our view that without accounting for T_4 correlations one cannot bring the results of conventional SRCC computations to a close agreement with the experimentally derived data. While the examination of the role of T_4 clusters, basis set, geometries of the $\text{X}\,^3\text{A}_2'$ and $\text{B}\,^1\text{A}_1$ states employed in the calculations, ΔZPVE corrections, *etc.*, would certainly be interesting, it would also be outside the scope of the present study. Thus, in the remainder of this subsection, we return to the analysis of the performance of the CIQMC-driven CC(P;Q) approach and its CC(P) counterpart, especially their ability to converge the parent CCSDT energetics when the cc-pVDZ basis is employed.

The results of our semi-stochastic CC(P)/cc-pVDZ and CC(P;Q)/cc-pVDZ computations for the X ^{3}A$_2^\prime$ and B ^{1}A$_1$ states of trimethylenemethane and the adiabatic gap between them, along with the associated CCSD, CR-CC(2,3), and CCSDT data, are summarized in Table 3.15 and Fig. 3.15. As in the case of cyclopentadienyl cation, to reduce the computational costs of the underlying CIQMC propagations, especially in their later stages, we resorted to the truncated i-CISDTQ-MC approach. In analogy to cyclobutadiene and cyclopentadienyl cation, we terminated our i-CISDTQ-MC propagations after 80000 $\delta\tau = 0.0001$ a.u. MC time steps, where the differences between the CC(P;Q) and CCSDT energies of the X ^{3}A$_2^\prime$ and B ^{1}A$_1$ states fall below 1 microhartree. Consistent with the CC(P), CC(P;Q), and other SRCC calculations for trimethylenemethane reported in Table 3.15 and Fig. 3.15, we used the ROHF determinant to initiate the i-CISDTQ-MC propagation for the D$_{3h}$-symmetric X ^{3}A$_2^\prime$ (in C$_{2v}$, X ^{3}B$_2$) state and the RHF determinant to initiate the i-CISDTQ-MC run for the C$_{2v}$-symmetric B ^{1}A$_1$ state. The lists of triply excited determinants captured by the i-CISDTQ-MC runs at the various times $\tau > 0$, needed to construct the P spaces for the CC(P) and CC(P;Q) computations, were the $S_z = 1$ triples of the B$_2$ (C$_{2v}$) symmetry in the case of the X ^{3}A$_2^\prime$ state and the $S_z = 0$ triples of the A$_1$ (C$_{2v}$) symmetry when considering the B ^{1}A$_1$ state. The remaining triples not captured by i-CISDTQ-MC defined the corresponding Q spaces.

It is clear from the results presented in Table 3.15 and Fig. 3.15 that the semi-stochastic CC(P;Q) approach is very effective in converging the parent CCSDT energetics characterizing the X ^{3}A$_2^\prime$ and B ^{1}A$_1$ states of trimethylenemethane and the adiabatic gap between them. It offers substantial improvements in the results of the CR-CC(2,3) calculations in the early stages of the underlying i-CISDTQ-MC propagations, especially when the multiconfigurational B ^{1}A$_1$ state and the adiabatic B ^{1}A$_1$ $-$ X ^{3}A$_2^\prime$ separation $\Delta E_{\text{S-T}}$, which are poorly described by CR-CC(2,3), are examined, while greatly accelerating the convergence of the CC(P) energies toward CCSDT. Indeed, after 6000 $\delta\tau = 0.0001$ a.u. MC iterations, which is a very short propagation time engaging only $\sim$0.1% of the total walker popula-

tions at $\tau = 8.0$ a.u., where we terminated our i-CISDTQ-MC runs (*cf.* Table 3.16), and i-CISDTQ-MC capturing as little as 14–17% of all triply excited determinants, the semi-stochastic CC(P;Q) methodology reduces the 13.370 millihartree difference between the CR-CC(2,3) and CCSDT energies of the B ^{1}A$_1$ state and the 8.1 kcal/mol error in the CR-CC(2,3) value of the B ^{1}A$_1$ $-$ X ^{3}A$_2^\prime$ gap relative to CCSDT to 1.260 millihartree and 0.6 kcal/mol, respectively, which is a chemical accuracy regime. Interestingly, the i-CISDTQ-MC-based CC(P;Q) value of $\Delta E_{\text{S-T}}$ obtained after 6000 MC iterations matches the quality of the B ^{1}A$_1$ $-$ X ^{3}A$_2^\prime$ gap resulting from the fully deterministic CC(P;Q) calculations using the CC(t;3) approach, which give a 0.5 kcal/mol error relative to CCSDT when the cc-pVDZ basis set is employed [88]. After the additional 4000 MC time steps, where the i-CISDTQ-MC propagations for the X ^{3}A$_2^\prime$ and B ^{1}A$_1$ states are still very far from convergence and where the fractions of triples captured by i-CISDTQ-MC increase to about 30%, the small errors relative to CCSDT characterizing the i-CISDTQ-MC-based CC(P;Q) values of the energy of the B ^{1}A$_1$ state and $\Delta E_{\text{S-T}}$ at $\tau = 0.6$ a.u. drop down by factors of 4–6, to 0.314 millihartree and 0.1 kcal/mol, respectively, illustrating how rapid the convergence of the CIQMC-driven CC(P;Q) calculations toward the parent SRCC data can be. While the CR-CC(2,3) description of the X ^{3}A$_2^\prime$ state, which has a SR character, is much better than in the case of its strongly correlated B ^{1}A$_1$ counterpart, the semi-stochastic CC(P;Q) computations offer great improvements in this case too. They are, for example, capable of reducing the $\sim$0.4 millihartree difference between the CR-CC(2,3) and CCSDT energies to a 0.1 millihartree level after 10000 MC iterations and i-CISDTQ-MC capturing less than 30% of all triples. In analogy to all other molecular examples considered in this chapter, the uncorrected CC(P) values of the energies of the X ^{3}A$_2^\prime$ and B ^{1}A$_1$ states and separation between them converge to their CCSDT limits too, but it is clear from Table 3.15 and Fig. 3.15 that they do it at a much slower rate than their CC(P;Q) counterparts. This can be illustrated by comparing the errors relative to CCSDT characterizing the CC(P) and CC(P;Q) energies of the X ^{3}A$_2^\prime$ and B ^{1}A$_1$ states and separation between them obtained after 6000 MC iterations. They

are more than 11 millihartree, about 21 millihartree, and almost 6 kcal/mol, respectively, in the former case and only 0.253 millihartree, 1.260 millihartree, and 0.6 kcal/mol, when the CC(P) energies are corrected for the remaining T_3 correlations using the CC(P;Q) approach. As explained in Section 3.1, the CC(P) energies converge to CCSDT more slowly than their $\delta(P;Q)$-corrected CC(P;Q) counterparts, since the initial, $\tau = 0$, CC(P) calculation for a given electronic state is equivalent to CCSD, where $T_3 = 0$. The CIQMC-driven CC(P;Q) calculations start from CR-CC(2,3), which provides information about T_3 clusters via non-iterative corrections to CCSD. This once again emphasizes the benefits of using corrections $\delta(P;Q)$ in the context of the semi-stochastic CC(P;Q) work.

Table 3.15 Convergence of the CC(P) and CC(P;Q) energies of the X ^{3}A$_2'$ and B ^{1}A$_1$ states of trimethylenemethane, as described by the cc-pVDZ basis set, and of the corresponding adiabatic singlet–triplet gaps toward their parent CCSDT values. The D_{3h}- and C_{2v}-symmetric geometries of the X ^{3}A$_2'$ and B ^{1}A$_1$ states, respectively, optimized in the SF-DFT/6-31G(d) calculations, were taken from Ref. [231]. The P spaces used in the CC(P) and CC(P;Q) calculations were defined as all singly and doubly excited determinants and subsets of triply excited determinants extracted from the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. The Q spaces used to determine the CC(P;Q) corrections consisted of the triply excited determinants not captured by the corresponding i-CISDTQ-MC runs. The i-CISDTQ-MC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 1500 walkers on the ROHF (X ^{3}A$_2'$ state) and RHF (B ^{1}A$_1$ state) reference determinants and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the four lowest core orbitals were kept frozen and the spherical components of d orbitals were employed throughout. Adapted from Ref. [102].

	X ^{3}A$_2'$			B ^{1}A$_1$			B ^{1}A$_1$ − X ^{3}A$_2'$	
MC Iterations	P[a]	$(P;Q)$[a]	%T[b]	P[a]	$(P;Q)$[a]	%T[b]	P[c]	$(P;Q)$[c]
0	19.202[d]	0.418[e]	0	58.051[d]	13.370[e]	0	24.4[d]	8.1[e]
2000	17.975	0.422	1.1	50.012	9.362	1.2	20.1	5.6
4000	14.462	0.357	6.6	32.925	3.236	7.7	11.6	1.8
6000	11.319	0.253	14.1	20.628	1.260	16.8	5.8	0.6
8000	9.066	0.173	21.3	14.601	0.649	25.5	3.5	0.3
10000	7.429	0.123	27.9	10.680	0.314	33.1	2.0	0.1
20000	3.294	0.031	52.3	2.675	0.028	61.1	-0.4	0.0
50000	0.213	0.001	92.8	0.061	0.000	97.1	-0.1	0.0
80000	0.012	0.000	99.5	0.002	0.000	99.9	0.0	0.0
∞	−155.466242[f]			−155.431596[f]			21.7[g]	

[a] Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

[b] The %T values are the percentages of triples captured during the i-CISDTQ-MC propagations [the $S_z = 1$ triply excited determinants of the B_2 (C_{2v}) symmetry in the case of the X ^{3}A$_2'$ state and the $S_z = 0$ triply excited determinants of the A_1 symmetry in the case of the B ^{1}A$_1$ state].

[c] Unless otherwise specified, the values of the singlet–triplet gaps are reported as errors relative to CCSDT in kcal/mol.

[d] Equivalent to CCSD.

[e] Equivalent to CR-CC(2,3) [the most complete variant of CR-CC(2,3) abbreviated sometimes as CR-CC(2,3),D or CR-CC(2,3)$_D$].

[f] Total CCSDT energy in hartree.

[g] The CCSDT singlet–triplet gap in kcal/mol.

Table 3.16 The total numbers of walkers, reported as percentages of the total walker populations at 80000 MC iterations, characterizing the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. that were needed to generate the CC(P) and CC(P;Q) results for trimethylenemethane reported in Table 3.15. Adapted from Ref. [102].

MC Iterations	X ^{3}A$_2'$	B ^{1}A$_1$
0	0.00[a]	0.00[a]
2000	0.01	0.01
4000	0.06	0.05
6000	0.14	0.11
8000	0.24	0.18
10000	0.34	0.26
20000	1.09	0.93
50000	14.60	15.05
80000	100[b]	100[c]

[a] The initial walker population, meaning 1500 walkers on the ROHF (X ^{3}A$_2'$ state) and RHF (B ^{1}A$_1$ state) reference determinants.
[b] The total number of walkers at 80000 MC iterations is 2363904677.
[c] The total number of walkers at 80000 MC iterations is 3543757954.

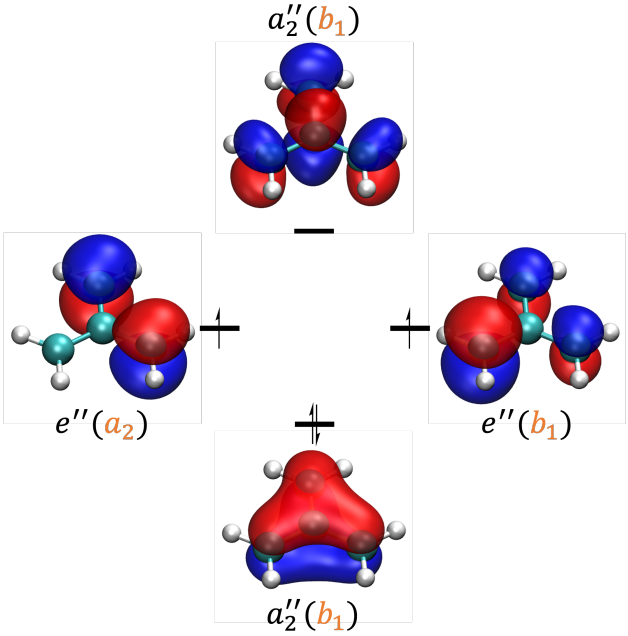

Figure 3.14 π molecular orbital network of the trimethylenemethane molecule, obtained at the HF/cc-pVDZ level, at the D$_{3h}$-symmetric geometry of the X ^{3}A$_2'$ state optimized in the SF-DFT/6-31G(d) calculations and taken from Ref. [231] The orbital irreducible representations in the D$_{3h}$ symmetry are shown in black and the corresponding labels in the C$_{2v}$ symmetry are shown in the parenthesis in orange. This electronic configuration refers to the triplet state.

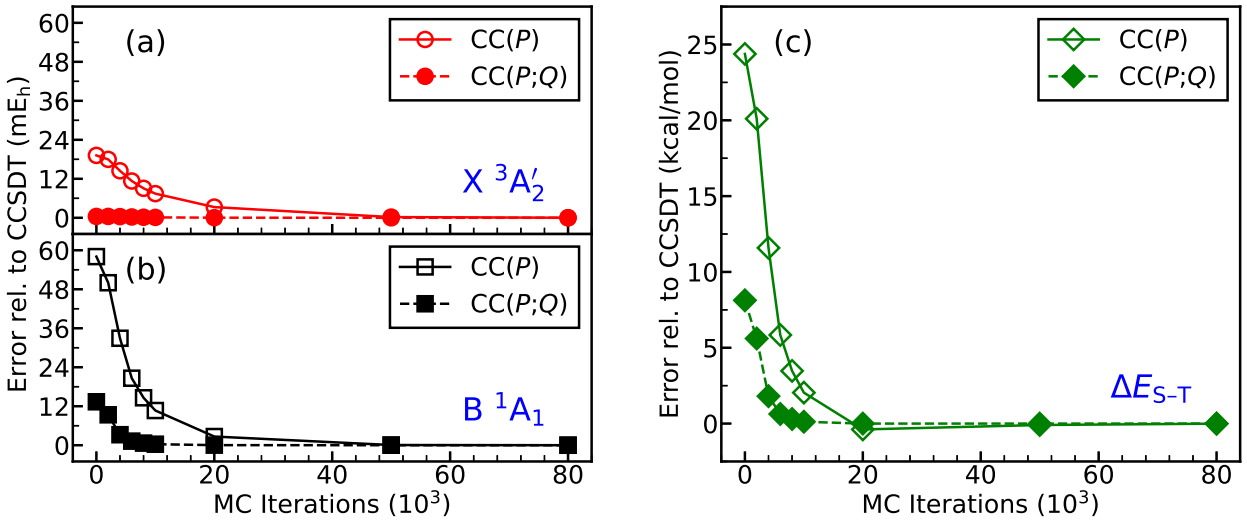

Figure 3.15 Convergence of the CC(P) and CC(P;Q) energies of the X ^{3}A$_2'$ [panel (a)] and B ^{1}A$_1$ [panel (b)] states of trimethylenemethane, as described by the cc-pVDZ basis set, and of the corresponding adiabatic singlet–triplet gaps [panel (c)] toward their parent CCSDT values. The geometries of the X ^{3}A$_2'$ and B ^{1}A$_1$ states, optimized in the SF-DFT/6-31G(d) calculations, were taken from Ref. [231]. The P spaces consisted of all singles and doubles and subsets of triples identified during the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. and the Q spaces consisted of the triples not captured by i-CISDTQ-MC. Adapted from Ref. [102].

CHAPTER 4

THE SEMI-STOCHASTIC EXTENSIONS OF PARTICLE NONCONSERVING EQUATION-OF-MOTION COUPLED-CLUSTER THEORIES

This chapter is based on the method development and programming work, followed by benchmark computations, aimed at extending the semi-stochastic approaches to particle nonconserving EOMCC schemes, which is subject to the manuscript in preparation [242], in which I have played a lead role in every aspect of the work other than the proposal for developing such ideas by my advisor Professor Piotr Piecuch and oversight from him.

As described in the Introduction, despite remarkable strides in computer hardware and efficient software advancements, providing an accurate and reliable description of MR problems and open-shell systems remain a challenge. This includes excited states dominated by two-electron transitions and electronic spectra of radical and biradical species, to name a few. In the previous chapter, we have already discussed how we can use the semi-stochastic CC($P;Q$) approach to tackle these challenging situations, and in this chapter, we describe the how the semi-stochastic ideas can be combined with the particle nonconserving EOMCC formalisms to achieve highly accurate results for open-shell systems.

As previously discussed, the particle nonconserving EOMCC approaches offer a simple and elegant way of describing the electronic spectra of open-shell systems by formally adding electrons to (electron attachment) or by removing electrons from (electron ionization) the nearest closed-shell reference core. Within this family of EOMCC methods, we find the EA and and IP EOMCC approaches [44, 78, 79, 138–146, 243–248], where the reference and the target systems differ by one electron ($|N_{\rm target} - N_{\rm reference}| = 1$), the DEA and DIP EOMCC methodologies [80, 147–154, 157, 249, 250], where $|N_{\rm target} - N_{\rm reference}| = 2$, and their higher order extensions. These approaches prove particularly advantageous in the study of electronic spectra of open-shell systems because the operation of adding electrons to or removing electrons from a closed-shell reference to attain a target system automatically creates an appropriate multi-reference model space specific to the system of interest while

relaxing the remaining electrons. In contrast to employing genuine MRCC approaches, these methods, being formally single reference in nature, are considerably much simpler to use. Furthermore, compared to traditional spin-integrated, spin-orbital implementations of particle-conserving CC/EOMCC treatments employing unrestricted or restricted open-shell reference determinants they offer distinct advantages, such as rigorous spin and symmetry adaptation of the computed states and the capability to balance both high-spin and low-spin states in an accurate and equally balanced footing.

Within this area, the EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p) methods [78, 140–142, 162, 163, 251–253], have demonstrated notable success in accurately describing open-shell systems featuring one unpaired electron outside a closed-shell core *i.e.*, they can effectively treat the (1,0) and (0,1) sectors of the Fock space. Additionally, the DEA-EOMCC(4p-2h) and DIP-EOMCC(4h-2p) methods [80, 153–156, 254–256] enable quantitative descriptions of the (2,0) and (0,2) sectors of the Fock space, thereby facilitating the precise determination of singlet–triplet gaps in biradical systems. To facilitate the application of these approaches without compromising on the accuracy, the Piecuch group has previously leveraged the active space ideas based on active orbitals to select the leading components in the respective electron attaching or removing operators, which resulted in methods [78, 80, 140–142, 153–156, 162, 163, 251–256] designated as EA-EOMCCSDt, IP-EOMCCSDt, DEA-EOMCC(3p-1h,4p-2h)$\{N_u\}$, DEA-EOMCC(4p-2h)$\{N_u\}$, and DIP-EOMCC(4h-2p)$\{N_o\}$ (N_o and N_u indicate the numbers of active occupied and unoccupied orbitals used to define the active-space). While the aforementioned methodologies have proven to be very reliable in describing the electronic spectra of radicals and determining singlet–triplet gaps in biradical systems, the selection of an appropriate active-space relies heavily on the specific system under study and the chemical intuition of the investigator. Notably, the active-space selection can pose a significant challenge in large and complex molecular systems. In this work, we propose an alternative to using active orbitals to select the dominant higher-order excitations. Drawing inspiration from our group's prior semi-stochastic CC(P)/EOMCC(P) and

CC(P;Q) work [99–102, 137], where the CIQMC wave function propagations [91–94, 257] were fused with deterministic CC/EOMCC computations to provide high-level CC/EOMCC energetics at a reduced computational cost and in a black-box manner. To be more specific, we extend the semi-stochastic particle-conserving EOMCC(P) approach [100, 137] to the particle nonconserving regime, resulting in the EA-EOMCC(P), IP-EOMCC(P), DEA-EOMCC(P), and DIP-EOMCC(P) methodologies [242]. These methods utilize CIQMC wave function propagations to stochastically select lists of dominant 3p-2h/3h-2p/4p-2h/4h-2p components in the EOMCC electron-attaching and electron-removing operators in an automated fashion and subsequently solve the particle-nonconserving EOMCC equations based on those stochastically determined lists. To validate the efficiency of these semi-stochastic approaches, we examine the C_2N, CNC, N_3, and NCO radicals and methylene and TMM biradicals.

4.1 Theory

Consistent with the previously developed semi-stochastic CC(P) and EOMCC(P) schemes [99–102, 137], the CIQMC-driven EA/IP/DEA/DIP-EOMCC methods, abbreviated as EA-EOMCC(P), IP-EOMCC(P), DEA-EOMCC(P), and DIP-EOMCC(P), respectively, utilize stochastic QMC wave function propagations to select the dominant high-order correlations in an automated black-box manner. The key algorithmic details of the semi-stochastic EA/IP/DEA/DIP-EOMCC calculations are as follows [242]:

1. First, perform a CCSD calculation, for the $(N \mp 1)$- or $(N \mp 2)$-electron closed-shell system in the case of the EA/IP-EOMCC, or DEA/DIP-EOMCC, respectively, to obtain the cluster amplitudes and to construct the similarity transformed Hamiltonian.

2. Next, start a CIQMC propagation by placing a certain number of walkers on the reference determinant pertaining to the N-electron target system, while using the one- and two-electron integrals from the closed-shell reference.

3. Afterwards extract a list of the most important determinants relevant to the EA/IP

and DEA/DIP EOMCC theory of interest [*e.g.*, 3*p*-2*h* excited determinants relative to $|\Phi^{(N-1)}\rangle$ for EA-EOMCC(3*p*-2*h*), 3*h*-2*p* excited determinants relative to $|\Phi^{(N+1)}\rangle$ for IP-EOMCC(3*h*-2*p*), 4*p*-2*h* excited determinants relative to $|\Phi^{(N-2)}\rangle$ for DEA-EOMCC(4*p*-2*h*), and 4*h*-2*p* excited determinants relative to $|\Phi^{(N+2)}\rangle$ for DIP-EOMCC(4*h*-2*p*)] from the CIQMC wave function at a given time τ to define the P spaces for the EA/IP-EOMCC(P) and DEA/DIP-EOMCC(P) calculations. If the target approach is EA-EOMCC(3*p*-2*h*), the P space ($\mathscr{H}^{(P)}$) is defined as all 1*p* ($\mathcal{G}^{1p}$), all 2*p*-1*h* ($\mathcal{G}^{2p\text{-}1h}$), and a subset of 3*p*-2*h* ($g^{3p\text{-}2h}$) excited determinants having at least n_P (e.g., one) positive

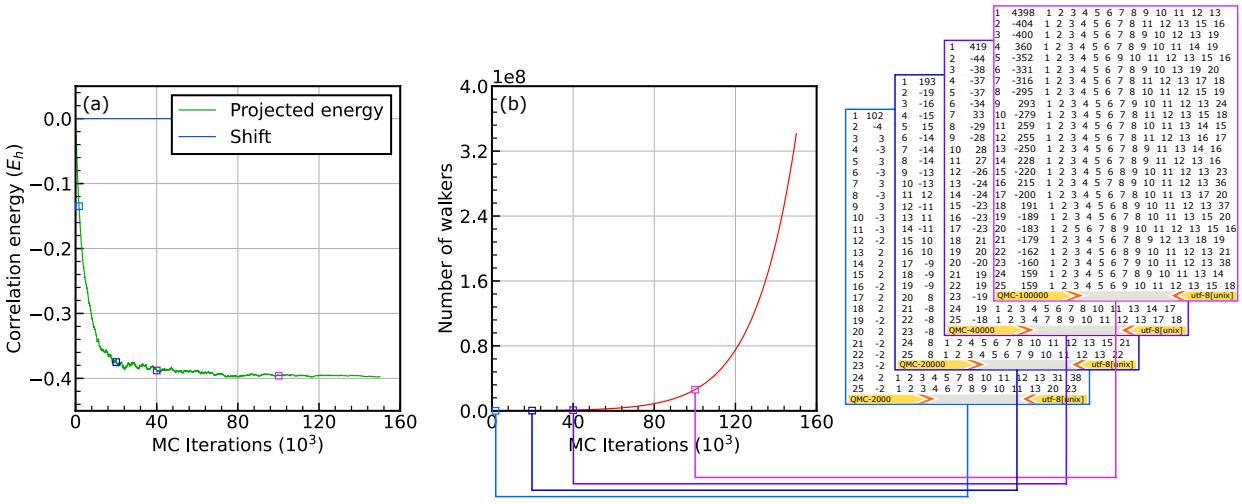

Figure 4.1 A schematic illustration depicting the construction of P-spaces in the EA/IP/DEA/DIP-EOMCC(P) computations. Panel (a) showcases the stabilization of correlation energy (green line) and the corresponding increase in the total number of walkers is shown in panel (b) (red line). On the right four snapshots from a QMC calculation are presented, featuring the lists of determinants picked up by the QMC algorithm at various time steps (light blue for 2000, dark blue for 20000, violet for 40000, and magenta for 100000 QMC iterations). It is evident that QMC deems some determinants more important than others by placing more walkers on them.

or negative walker, such that $\mathscr{H}^{(P)} = \mathcal{G}^{1p} \oplus \mathcal{G}^{2p\text{-}1h} \oplus g^{3p\text{-}2h}$. Similarly, define $\mathscr{H}^{(P)} = \mathcal{G}^{1h} \oplus \mathcal{G}^{2h\text{-}1p} \oplus g^{3h\text{-}2p}$ for IP-EOMCC(3*h*-2*p*), $\mathscr{H}^{(P)} = \mathcal{G}^{2p} \oplus \mathcal{G}^{3p\text{-}1h} \oplus g^{4p\text{-}2h}$ for DEA-EOMCC(4*p*-2*h*), and in case of DIP-EOMCC(4*h*-2*p*) use $\mathscr{H}^{(P)} = \mathcal{G}^{2h} \oplus \mathcal{G}^{3h\text{-}1p} \oplus g^{4h\text{-}2p}$ (see Fig. 4.1 for an illustration).

4. Then solve the semi-stochastic EA/IP-EOMCC and DEA/DIP-EOMCC equations in

the stochastically determined P spaces, *i.e.*, use $R_\mu^{(+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h}^{(\mathrm{MC})}$ for EA-EOMCC(P), $R_\mu^{(-1)} = R_{\mu,1h} + R_{\mu,2h\text{-}1p} + R_{\mu,3h\text{-}2p}^{(\mathrm{MC})}$ for IP-EOMCC(P), $R_\mu^{(+2)} = R_{\mu,2p} + R_{\mu,3p\text{-}1h} + R_{\mu,4p\text{-}2h}^{(\mathrm{MC})}$ for DEA-EOMCC(P), and $R_\mu^{(-2)} = R_{\mu,2h} + R_{\mu,3h\text{-}1p} + R_{\mu,4h\text{-}2p}^{(\mathrm{MC})}$ for DIP-EOMCC(P).

5. Finally, check the convergence of the resulting energies by repeating the steps 3 and 4 at some later CIQMC propagation time $\tau' > \tau$. If the energies do not change within a given convergence threshold, we can stop the calculation. One can also stop the calculation if τ in steps 3 and 4 is chosen such that the stochastically determined P spaces contain sufficiently large fraction of higher order excitations relevant to the target EOMCC level.

Before going to the next section, we must discuss an interesting aspect of the semi-stochastic particle nonconserving EOMCC(P) approaches. At $\tau = 0$, the EA-EOMCC(P), IP-EOMCC(P), DEA-EOMCC(P), and DIP-EOMCC(P) energies are identical to the energies obtained in the EA-EOMCC(2p-1h), IP-EOMCC(2h-1p), DEA-EOMCC(3p-1h), and DIP-EOMCC(3h-1p) calculations. This is because at $\tau = 0$ the P spaces in the respective calculations do not contain any 3p-2h, 3h-2p, 4p-2h, or 4h-2p determinants. On the other hand, when $\tau = \infty$, the P spaces in the EA-EOMCC(P), IP-EOMCC(P), DEA-EOMCC(P), and DIP-EOMCC(P) calculations contain all the 3p-2h, 3h-2p, 4p-2h, and 4h-2p determinants, respectively, and, as a result, the semi-stochastic EA/IP/DEA/DIP-EOMCC calculations provide energetics identical to that of the EA-EOMCC(3p-2h), IP-EOMCC(3h-2p), DEA-EOMCC(4p-2h), and DIP-EOMCC(4h-2p) approaches. This relationship between the EA-EOMCC(P), IP-EOMCC(P), DEA-EOMCC(P), and DIP-EOMCC(P) methods and the fully deterministic EA-EOMCC(3p-2h), IP-EOMCC(3h-2p), DEA-EOMCC(4p-2h), and DIP-EOMCC(4h-2p) methodologies were helpful in examining the correctness of our codes. They also point to the ability of these semi-stochastic particle nonconserving EOMCC approaches driven by the information extracted from CIQMC to offer a systematically improv-

able description as τ goes from 0 to ∞. In the next section, we examine the performance of the semi-stochastic EA-EOMCC(P) and IP-EOMCC(P) approaches in converging the high-level EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p) methods and the section after that illustrates the capability of their double electron attachment [DEA-EOMCC(P)] and double ionization potential [DIP-EOMCC(P)] extensions to converge the DEA-EOMCC(4p-2h) and DIP-EOMCC(4h-2p) methodologies.

4.2 Adiabatic Excitations in C_2N, CNC, N_3, and NCO

In order to demonstrate the ability of the semi-stochastic EA-EOMCC(P) and IP-EOMCC(P) approaches in converging their high-level, fully-deterministic parents, EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p), we carried out benchmark calculations for the ground and a few lowest-lying doublet states along with the corresponding adiabatic excitation energies of the C_2N [Table 4.1 and Fig. 4.2], CNC [Table 4.2 and Fig. 4.3], N_3 [Table 4.3 and Fig. 4.4], and NCO [Table 4.4 and Fig. 4.5] radicals. All the systems utilized the DZP[4s2p1d] basis set [167, 168] and employed geometries optimized with the SAC-CI SDT-R/PS method combined with the DZP[4s2p1d] basis set, as reported in Ref. [162]. Following our previous semi-stochastic work [99–101, 137], we used the HANDE software package [170, 171] to execute all our i-FCIQMC calculations needed to generate the lists of 3p-2h and 3h-2p determinants entering the P spaces for the EA-EOMCC(P) and IP-EOMCC(P) calculations. Our standalone CC/EOMCC codes, interfaced with the RHF, ROHF, and integral transformation routines available in the GAMESS software package [172–174], were used to carry out the EA-EOMCC(P), IP-EOMCC(P), and the fully deterministic CCSD, EA-EOMCC(2p-1h), EA-EOMCC(3p-2h), IP-EOMCC(2h-1p), and IP-EOMCC(3h-2p) computations. Each i-FCIQMC run was initiated by placing 500 walkers on the relevant reference function (see Tables 4.1–4.4 for the details), the initiator parameter n_a was set at 3, and all of the i-FCIQMC propagations used a time step of $\tau = 0.0001$ a.u. As the true point group of the system of interest was not Abelian for all four molecules, namely C_2N, CNC, N_3, and NCO, we utilized their largest Abelian subgroups in the calculations. This choice was necessary,

since all of our CC/EOMCC codes interfaced with GAMESS and the CIQMC routines in HANDE can only handle Abelian symmetries. In all the post-HF computations, the core electrons corresponding to the 1s shells of the carbon, nitrogen, and oxygen atoms were kept frozen.

4.2.1 Application of the EA-EOMCC(P) approach to C_2N and CNC radicals

In this subsection, we investigate the performance of the semi-stochastic EA-EOMCC(P) approach in converging the EA-EOMCC(3p-2h) total electronic and adiabatic excitation energies of a few low-lying doublet states of C_2N and CNC radicals. The reference wave functions are obtained by performing CCSD calculations on the nearest closed-shell C_2N^+ and CNC^+ cations using the DZP[4s2p1d] basis set [167, 168] and, as mentioned above, the nuclear geometries of the ground and excited states of C_2N and CNC, optimized using EA SAC-CI-SDT-R/PS and DZP[4s2p1d] basis set, are taken from Ref. [162].

4.2.1.1 C_2N

We begin our discussion by using the C_2N radical, where we study the X $^2\Pi$ ground state and three low-lying valence doublet excited states, A $^2\Delta$, B $^2\Sigma^-$, and C $^2\Sigma^+$. Following Ref. [162], we used the EA SAC-CI-SDT-R/PS with DZP[4s2p1d] optimized geometries, which are R_{CC} = 1.400 Å and R_{CN} = 1.185 Å for the X $^2\Pi$ state, R_{CC} = 1.315 Å and R_{CN} = 1.207 Å for the A $^2\Delta$ state, R_{CC} = 1.302 Å and R_{CN} = 1.223 Å for the B $^2\Sigma^-$ state, and R_{CC} = 1.311 Å and R_{CN} = 1.214 Å for the C $^2\Sigma^+$ state. At these geometries, all the three valence excited states studied here, are characterized by two-electron transitions as well as significant 2p-1h components (the B $^2\Sigma^-$ also has non-negligible 3p-2h contributions) in the wave functions relative to the reference C_2N^+ ion. Therefore, to properly describe these states, an accurate treatment of 3p-2h excitations is required and this is evident from Table 1 of Ref. [162] or Table IV of Ref. [68] (*cf.*, also, Table 4.1 of this work), where the basic EA-EOMCCSD or EA-EOMCC(2p-1h) approach predicts an incorrect state ordering, *i.e.*, suggesting the B $^2\Sigma^-$ state to be higher in energy than the C $^2\Sigma^+$ state. Furthermore, the A $^2\Delta$ − X $^2\Pi$, B $^2\Sigma^-$ − X $^2\Pi$, and C $^2\Sigma^+$ − X $^2\Pi$ adiabatic excitation energies obtained

from EA-EOMCC(2p-1h) (6.190 eV, 7.860 eV, and 6.722 eV) are ∼3–5 eV away from the experimental results (2.636 eV, 2.779 eV, and 3.306 eV) reported in Ref. [258]. The higher order EA-EOMCC(3p-2h) method that accounts for a complete treatment of 3p-2h terms in the electron attaching operator $R_\mu^{(+1)}$, reduces these errors to ∼1 eV or lower (3.055 eV, 3.678 eV, and 3.809 eV), necessitating the need to use at least the EA-EOMCC(3p-2h) level of theory, if a realistic description is wanted. However, as mentioned before, EA-EOMCC(3p-2h) has a scaling of $\mathscr{N}^7$ and, hence, could become very expensive very quickly. It would, therefore, be interesting to see if the semi-stochastic EA-EOMCC(P) approach is capable of recovering the EA-EOMCC(3p-2h)-quality results in this challenging case of C_2N with a fraction of 3p-2h terms selected in a automated manner using the i-FCIQMC wave function propagations.

Following the semi-stochastic EA-EOMCC(P) algorithm, as described above, and our interest in converging the EA-EOMCC(3p-2h) energetics of the ground and three lowest lying valence excited states of C_2N, the cluster operator and the electron attaching operator were approximated as $T = T_1 + T_2$ and $R_\mu^{(+1)} = R_{\mu,1p} + R_{\mu,2p\text{-}1h} + R_{\mu,3p\text{-}2h}^{(\mathrm{MC})}$, where the lists of 3$p-2h$ determinants defining the $R_{\mu,3p\text{-}2h}^{(\mathrm{MC})}$ components at a given time τ were obtained from the corresponding i-FCIQMC propagations at the same value of τ, meaning the B_1 (C_{2v}) component of $^2\Pi$ for the X $^2\Pi$ state, the A_1 (C_{2v}) component of $^2\Delta$ for the A $^2\Delta$ state, A_2 (C_{2v}) for the B $^2\Sigma^-$ state, and the A_1 (C_{2v}) for the C $^2\Sigma^+$ state (C_{2v} is the largest Abelian subgroup of the true point group of C_2N, C∞v). The results of our EA-EOMCC(P) computations are reported in Table 4.1 and Fig. 4.2.

As can be seen from Table 4.1, at $\tau = 0.0$, the the EA-EOMCC(P) results are far from the EA-EOMCC(3p-2h) energetics, which is expected since, at this stage, the P spaces lack any 3p-2h determinants, rendering these results equivalent to those obtained from the basic EA-EOMCC(2p-1h) method. The ground state of the C_2N radical, X $^2\Pi$, is dominated by 1p excitations relative to the C_2N^+ reference and, as a result, it is described reasonably well by EA-EOMCCSD [EA-EOMCC($P|_{\tau=0}$)]. Still errors of ∼5 millihartree remain relative to

EA-EOMCC(3p-2h). This error rapidly decreases to the 1–3 millihartree range as soon as we start incorporating ∼20–30% 3p-2h determinants in the P space. Indeed, at 20000 MC iterations and with about 30% of 3p-2h determinants the error in EA-EOMCC(P) energy is only 1.421 millihartree. The next state denoted as A $^2\Delta$, which is also the first excited state, has significant 2p-1h contributions relative to the reference determinant and, therefore, the effect of adding 3p-2h determinants in the P space is more prominent. For this state, the EA-EOMCC($P|_{\tau=0}$) approach (EA-EOMCCSD) has, massive, 120 millihartree errors relative to the target EA-EOMCC(3p-2h) method. This huge error is brought down to less than 20 millihartree with 40% of 3p-2h components and at 20000 QMC iterations. If we let QMC run a little longer until 50000 QMC iterations, it recovers ∼70% of 3p-2h determinants and the errors in the total energies are reduced to less than 1 millihartree relative to EA-EOMCC(3p-2h). If we look at the adiabatic excitation energies, they are converged to 0.2 eV or better at a much earlier QMC time step. For example, already at 20000 Monte Carlo time steps the A $^2\Delta$ − X $^2\Pi$ excitation energy is less than 0.5 eV in error and after 50000 iterations the error is only 0.013 eV. The next state of interest is B $^2\Sigma^-$ and this is also the most difficult to describe state in this system. It is characterized by significant 3p-2h contributions in addition to dominant 2p-1h correlations, and as a result, the EA-EOMCC(2p-1h) approach fails miserably to describe this state, producing ∼160 millihartree errors relative to EA-EOMCC(3p-2h). As we start including 3p-2h determinants in the P space with the help of CIQMC wave function propagations, the errors decrease at an unbelievable rate (see Fig. 4.2). For example, just after 4000 MC iterations and with only 4% of 3p-2h determinants in the P space, the huge > 150 millihartree errors decrease to 12.217 millihartree. After 30000 MC iterations, the CIQMC wave function samplings select 22.8% of 3p-2h determinants in the P space and the EA-EOMCC(P) energy at this stage is only ∼1 millihartree from EA-EOMCC(3p-2h). This extremely fast convergence in the total electronic energy is also reflected in the rapid convergence of the B $^2\Sigma^-$ − X $^2\Pi$ gap. At 4000 MC time steps this gap is already within 0.2 eV from the EA-EOMCC(3p-2h) gap and as soon as we reach 30000

MC iterations this B $^2\Sigma^-$ − X $^2\Pi$ gap is < 0.01 eV from the target. For the final C $^2\Sigma^+$ state for this system in our study, the errors in the EA-EOMCC($2p$-$1h$) energy is > 110 millihartree relative to EA-EOMCC($3p$-$2h$). For this state also, with incorporation of $3p$-$2h$ determinants in the P space, the errors decrease rapidly. At 10000 MC iterations and with 18% of $3p$-$2h$ determinants in the P space the errors in total electronic energies are reduced to half the errors in EA-EOMCCSD. As we incorporate more and more $3p$-$2h$ terms in the P space, the errors keep decreasing further. In terms of the C $^2\Sigma^+$ − X $^2\Pi$ gap, the results are even more favorable. For example, the 2.913 eV error at $\tau = 0$ reduces to 0.432 eV at 30000 MC iterations and this error decreases even further to less than 0.1 eV if we let QMC run an additional 10000 iterations.

Table 4.1 Convergence of the EA-EOMCC(P) energies [abbreviated as EA(P)] of the X $^2\Pi$, A $^2\Delta$, B $^2\Sigma^-$, and C $^2\Sigma^+$ states of C_2N, as described by the DZP[4s2p1d] basis set of Refs. [167, 168], and of the corresponding adiabatic excitation energies toward their parent EA-EOMCC(3p-2h) values. The geometries of the X $^2\Pi$, A $^2\Delta$, B $^2\Sigma^-$, and C $^2\Sigma^+$ states, optimized in the SAC-CI SDT-R/PS calculations using the same basis set, were taken from Ref. [162]. The P spaces used in the EA-EOMCC(P) calculations were defined as all 1p and 2p-1h determinants and subsets of 3p-2h determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The i-FCIQMC calculations preceding the EA-EOMCC(P) steps were initiated by placing 500 walkers on the ROHF reference determinants of the corresponding states and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbitals of the carbon and nitrogen atoms were kept frozen.

MC	X $^2\Pi$		A $^2\Delta$		B $^2\Sigma^-$		C $^2\Sigma^+$		Ad. Excit. Energy		
Iters.	EA(P)[a]	%(3p-2h)[b]	EA(P)[a]	%(3p-2h)[b]	EA(P)[a]	%(3p-2h)[b]	EA(P)[a]	%(3p-2h)[b]	ΔE_1[c]	ΔE_2[c]	ΔE_3[c]
0	4.696[d]	0	119.913[d]	0	158.397[d]	0	111.733[d]	0	3.135[d]	4.182[d]	2.913[d]
4000	5.634	8.1	104.423	8.1	12.217	4.3	98.803	7.1	2.688	0.179	2.535
10000	2.686	18.8	56.413	21.6	5.812	9.3	66.416	17.7	1.462	0.085	1.734
20000	1.421	30.2	18.608	38.3	2.407	15.8	33.336	31.0	0.468	0.027	0.868
30000	0.883	39.8	6.901	50.1	1.212	22.8	16.752	42.7	0.164	0.009	0.432
40000	0.679	48.1	4.193	60.5	0.757	30.4	9.944	52.6	0.096	0.002	0.252
50000	0.428	57.1	0.907	68.0	0.306	38.2	4.400	60.9	0.013	−0.003	0.108
60000	0.253	64.8	0.582	74.1	0.405	47.2	3.028	67.7	0.009	0.004	0.076
100000	0.071	82.4	0.011	90.1	0.045	76.3	0.180	85.1	−0.002	−0.001	0.003
∞	−130.404919[e]		−130.292647[e]		−130.269764[e]		−130.264924[e]		3.055[f]	3.678[f]	3.809[f]

[a] Unless otherwise stated, all energies are reported as errors relative to EA-EOMCC(3p-2h) in millihartree.

[b] The %(3p-2h) values are the percentages of 3p-2h determinants captured during the i-FCIQMC propagations (the $S_z = 1/2$ 3p-2h determinants of the B_1 symmetry in the case of the X $^2\Pi$ state, the A_1 symmetry in the case of the A $^2\Delta$ state, the A_2 symmetry in the case of the B $^2\Sigma^-$ state, and the A_1 symmetry in the case of the C $^2\Sigma^+$ state).

[c] Unless otherwise specified, the adiabatic excitation energies are reported as errors relative to EA-EOMCC(3p-2h) in eV; ΔE_1 = A $^2\Delta$ − X $^2\Pi$, ΔE_2 = B $^2\Sigma^-$ − X $^2\Pi$, and ΔE_3 = C $^2\Sigma^+$ − X $^2\Pi$.

[d] Equivalent to EA-EOMCC(2p-1h).

[e] Total EA-EOMCC(3p-2h) energy in hartree.

[f] The EA-EOMCC(3p-2h) adiabatic excitation energy in eV.

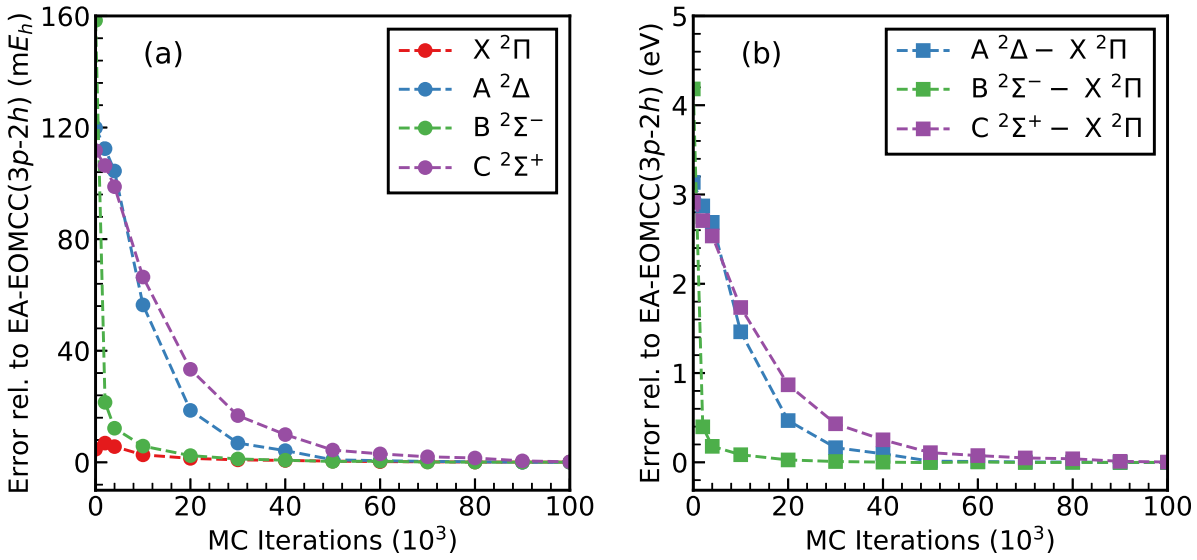

Figure 4.2 Convergence of (a) the EA-EOMCC(P) energies of the X $^2\Pi$, A $^2\Delta$, B $^2\Sigma^-$, and C $^2\Sigma^+$ states of C_2N, as described by the DZP[4s2p1d] basis set, and (b) the corresponding adiabatic excitation energies toward their parent EA-EOMCC(3p-2h) values.

4.2.1.2 CNC

The next example studied using the semi-stochastic EA-EOMCC(P) approach is the linear, $D_{\infty h}$ symmetric, CNC radical, where we considered the ground X $^2\Pi_g$ and the two low-lying A $^2\Delta_u$ and B $^2\Sigma_u^+$ doublet excited electronic states and the corresponding adiabatic excitation energies, A $^2\Delta_u$ − X $^2\Pi_g$ and B $^2\Sigma_u^+$ − X $^2\Pi_g$. Following Ref. [162], we used the EA SAC-CI-SDT-R/PS with DZP[4s2p1d] optimized geometries, which are $R_{\mathrm{CN}} = 1.253$ Å for the X $^2\Pi_g$ state, $R_{\mathrm{CN}} = 1.256$ Å for the A $^2\Delta_u$, and $R_{\mathrm{CN}} = 1.259$ Å for the B $^2\Sigma_u^+$ state. At these geometries, the X $^2\Pi_g$ state is dominated by $1p$ excitations out of the closed-shell reference state CNC^+, whereas the A $^2\Delta_u$ and B $^2\Sigma_u^+$ states are characterized by significant two-electron excitations, resulting in large $2p$-$1h$ contributions in the corresponding wave functions. Therefore, it is of no surprise that the effect of incorporating $3p$-$2h$ correlations is more important in the for A $^2\Delta_u$ and B $^2\Sigma_u^+$ states, but not so much in case of the X $^2\Pi_g$ state. So, in order to accurately describe the total electronic energies of the A $^2\Delta_u$ and B $^2\Sigma_u^+$ states, we need at least the EA-EOMCC($3p$-$2h$) level of theory and this is reflected in the A $^2\Delta_u$ − X $^2\Pi_g$ and B $^2\Sigma_u^+$ − X $^2\Pi_g$ adiabatic excitation energies. It can be seen from Table 1 of Ref. [162] or Table IV of Ref. [68] (*cf.*, also, Table 4.2 of this work) that the EA-EOMCCSD predicted A $^2\Delta_u$ − X $^2\Pi_g$ and B $^2\Sigma_u^+$ − X $^2\Pi_g$ adiabatic excitation energies are 7.206 eV and 7.639 eV, respectively, which are as much as 3.3–3.5 eV away from the experimentally obtained values of 3.761 eV and 4.315 eV [258]. When the EA-EOMCC calculations include all possible $3p$-$2h$ correlations, the resulting EA-EOMCC($3p$-$2h$) approach predicts these excitation energies to be 4.105 eV and 4.718 eV, respectively, which deviate from the experimental values by only 0.3–0.4 eV. The active-space EA-EOMCCSDt method also perform very well, not only improving the EA-EOMCCSD results, but also yielding results as good as the target EA-EOMCC($3p$-$2h$) methodology. Our target in this work is, therefore, to reproduce the EA-EOMCC($3p$-$2h$) energetics using the semi-stochastic, i-FCIQMC-driven, EA-EOMCC(P) approach. In all of our i-FCIQMC calculations for CNC, we used the D_{2h} Abelian subgroup of its true point group $D_{\infty h}$. In particular, following the computational

protocol described earlier, the underlying i-FCIQMC calculations required to select the lists of $3p$-$2h$ determinants defining the $R^{(\mathrm{MC})}_{3p\text{-}2h}$ components in the EA-EOMCC(P) calculations for CNC were set to converge the lowest energy states of B_{3g}, A_u, and B_{1u} symmetry for the X $^2\Pi_g$, A $^2\Delta_u$, and B $^2\Sigma^+_u$ states at their respective equilibrium geometries. The results of our semi-stochastic EA-EOMCC(P) approach in recovering EA-EOMCC($3p$-$2h$)-quality energetics for the X $^2\Pi_g$, A $^2\Delta_u$, and B $^2\Sigma^+_u$ states of CNC and the corresponding A $^2\Delta_u$ − X $^2\Pi_g$ and B $^2\Sigma^+_u$ − X $^2\Pi_g$ excitation energies are reported in Table 4.2 and Fig. 4.3 of this subsection.

As previously mentioned, the first row of Table 4.2 refers to EA-EOMCC(P) results for $\tau = 0$, which are equivalent to EA-EOMCCSD and as already discussed above, it provides a very poor description of the lowest-lying doublet states of CNC. As the QMC wave function evolves in time, which in this context refers to accumulating more and more $3p$-$2h$ determinants that can be used to enrich the P spaces in the semi-stochastic calculations, EA-EOMCC(P) systematically converges to our target EA-EOMCC($3p$-$2h$). The X $^2\Pi_g$ state, that is 4.881 millihartree away from EA-EOMCC($3p$-$2h$) at $\tau = 0$, narrows down to within ∼1 millihartree of EA-EOMCC($3p$-$2h$) with only 16% of $3p$-$2h$ determinants in the P space and this happens only at 20000 MC iterations. If the i-FCIQMC propagation is allowed to run an additional 10000 MC iteration, the EA-EOMCC(P) energies are less than 1 millihartree away from target and the percentage of $3p$-$2h$ determinants in this case is only 21%. For the remaining two states, A $^2\Delta_u$ and B $^2\Sigma^+_u$, the gradual increase in the $3p$-$2h$ determinants in the P space, as dictated by the underlying i-FCIQMC wave function propagation, is more pronounced. For the A $^2\Delta_u$ state, only at 4000 MC iterations and with only 5.5% of $3p$-$2h$ determinants in the P space, the ∼119 millihartree error in the EA-EOMCC($P|_{\tau=0}$) [EA-EOMCC($2p$-$1h$] energy, relative to EA-EOMCC($3p$-$2h$) steeply decreases to 12 millihartree and this happens only at 4000 MC iterations. If we let the i-FCIQMC propagations to run longer, the EA-EOMCC(P) energies converge to the target EA-EOMCC($3p$-$2h$) in a very rapid pace. For example, at 10000 MC iteration and with 12% of $3p$-$2h$ determinants

in the P space, the errors reduce to less than 5 millihartree and by the time 20%–30% 3p-2h determinants are captured, EA-EOMCC(P) is only 1–2 millihartree away from the target. In case of the A $^2\Delta_u$ − X $^2\Pi_g$ adiabatic excitation energy the situation is much more favorable. Already at 4000 MC iteration, the errors are less than 0.2 eV relative to EA-EOMCC(3p-2h) and after 10000 MC iterations the errors are always much lower than 0.1 eV. For the final, and most difficult to describe, B $^2\Sigma_u^+$ state, at 10000 MC iterations and with 18% of 3p-2h determinants, the errors in the EA-EOMCC(P) energies are reduced by almost a factor of 3 compared to EA-EOMCC($P|_{\tau=0}$), which is ∼112 millihartree away from EA-EOMCC(3p-2h). This error is further reduced to ∼5 millihartree or better when we include about 40% or more 3p-2h determinants in the P space. Interestingly, the B $^2\Sigma_u^+$ − X $^2\Pi_g$ adiabatic excitation energies are much more well behaved, and only at 20000 MC iterations and with 30% of 3p-2h determinants in the P space corresponding to the EA-EOMCC(P) calculation pertaining to the B $^2\Sigma_u^+$ state, this energy difference is converged to within 0.3 eV relative to EA-EOMCC(3p-2h).

The two above examples demonstrate that the semi-stochastic EA-EOMCC(P) method is capable of rapidly converging the EA-EOMCC(3p-2h) energetics for the individual electronic states and the corresponding adiabatic excitation energies of radicals with a very small percentage of 3p-2h determinants in the respective P spaces.

Table 4.2 Convergence of the EA-EOMCC(P) energies [abbreviated as EA(P)] of the X $^2\Pi_g$, A $^2\Delta_u$, and B $^2\Sigma_u^+$ states of CNC, as described by the DZP[4s2p1d] basis set of Refs. [167, 168], and of the corresponding adiabatic excitation energies toward their parent EA-EOMCC(3p-2h) values. The geometries of the X $^2\Pi_g$, A $^2\Delta_u$, and B $^2\Sigma_u^+$ states, optimized in the SAC-CI SDT-R/PS calculations using the same basis set, were taken from Ref. [162]. The P spaces used in the EA-EOMCC(P) calculations were defined as all 1p and 2p-1h determinants and subsets of 3p-2h determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The i-FCIQMC calculations preceding the EA-EOMCC(P) steps were initiated by placing 500 walkers on the ROHF reference determinants of the corresponding states and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbitals of the carbon and nitrogen atoms were kept frozen.

MC	X $^2\Pi_g$		A $^2\Delta_u$		B $^2\Sigma_u^+$		Ad. Excit. Energy	
Iters.	EA(P)[a]	%(3p-2h)[b]	EA(P)[a]	%(3p-2h)[b]	EA(P)[a]	%(3p-2h)[b]	ΔE_1[c]	ΔE_2[c]
0	4.881[d]	0	118.827[d]	0	112.241[d]	0	3.101[d]	2.921[d]
4000	4.591	5.1	11.668	5.5	93.500	8.1	0.193	2.419
10000	2.419	11.1	4.816	12.0	45.183	18.0	0.065	1.164
20000	1.370	16.0	1.851	20.6	12.348	31.3	0.013	0.299
30000	0.817	20.9	0.865	27.5	5.369	42.2	0.001	0.124
40000	0.572	24.5	0.407	35.4	2.221	51.1	−0.004	0.045
50000	0.369	28.8	0.194	43.4	1.707	59.1	−0.005	0.036
60000	0.264	32.0	0.096	52.6	0.228	67.0	−0.005	−0.001
100000	0.040	40.9	0.015	79.7	0.006	86.9	−0.001	−0.001
∞	−130.411530[e]		−130.260673[e]		−130.238150[e]		4.105[f]	4.718[f]

[a] Unless otherwise stated, all energies are reported as errors relative to EA-EOMCC(3p-2h) in millihartree.

[b] The %(3p-2h) values are the percentages of 3p-2h determinants captured during the i-FCIQMC propagations (the $S_z = 1/2$ 3p-2h determinants of the B_{3g} symmetry in the case of the X $^2\Pi_g$ state, the A_u symmetry in the case of the A $^2\Delta_u$ state, the B_{1u} symmetry in the case of the B $^2\Sigma_u^+$ state).

[c] Unless otherwise specified, the adiabatic excitation energies are reported as errors relative to EA-EOMCC(3p-2h) in eV; $\Delta E_1 = \text{A}\ ^2\Delta_u - \text{X}\ ^2\Pi_g$ and $\Delta E_2 = \text{B}\ ^2\Sigma_u^+ - \text{X}\ ^2\Pi_g$.

[d] Equivalent to EA-EOMCC(2p-1h).

[e] Total EA-EOMCC(3p-2h) energy in hartree.

[f] The EA-EOMCC(3p-2h) adiabatic excitation energy in eV.

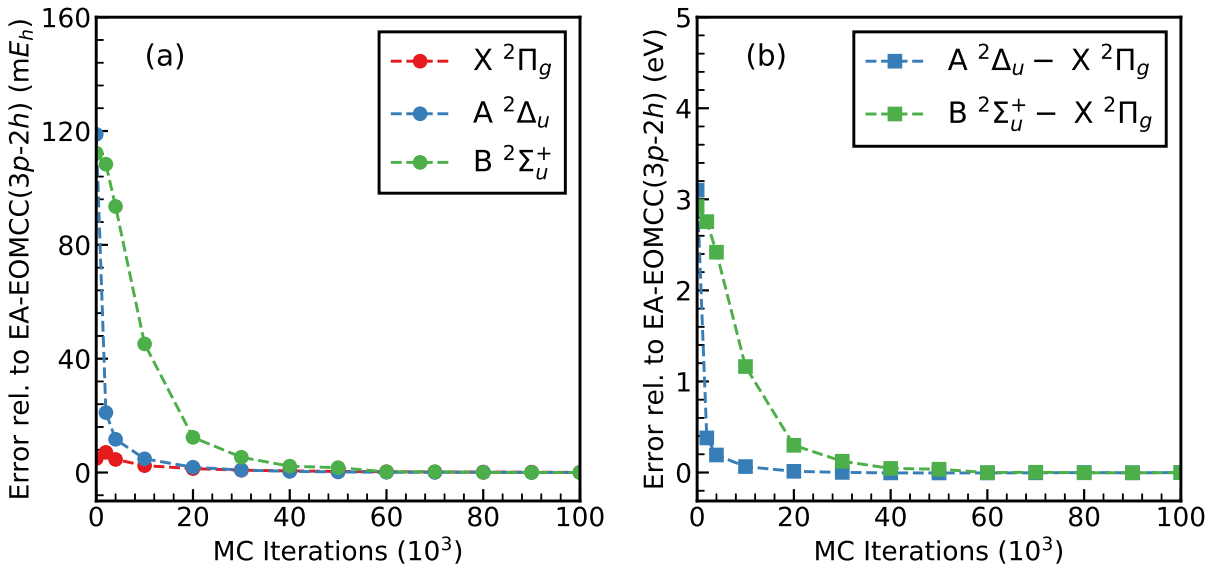

Figure 4.3 Convergence of (a) the EA-EOMCC(P) energies of the X $^2\Pi_g$, A $^2\Delta_u$, and B $^2\Sigma_u^+$ states of CNC, as described by the DZP[4s2p1d] basis set, and (b) the corresponding adiabatic excitation energies toward their parent EA-EOMCC(3p-2h) values.

4.2.2 Application of the IP-EOMCC(P) approach to N_3 and NCO radicals

After examining the performance of the EA-EOMCC(P) methodology in converging the high-level EA-EOMCC(3p-2h) energetics, we proceed to assess the capability of the IP-EOMCC(P) approach in recovering the IP-EOMCC(3h-2h) energetics by applying the IP-EOMCC(P) approach in computing the total electronic energies of the ground and a few excited doublet states of N_3 and NCO radicals and the corresponding adiabatic excitation energies. The reference wave functions were obtained by performing CCSD calculations on the N_3^- and NCO^- anions using the DZP basis set [167, 168] and geometries obtained from IP SAC-CI-SDT-R/PS / DZP optimizations as reported in Ref. [162].

4.2.2.1 N_3

Our first system under investigation is the N_3 radical, where we studied the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states, along with the corresponding B $^2\Sigma_u^+$– X $^2\Pi_g$ adiabatic excitation energy. For this, we utilized the IP SAC-CI-SDT-R/PS (in combination with the DZP[4s2p1d] basis set) optimized geometries obtained from Ref. [162]. These geometries are $R_{\mathrm{N-N}} = 1.188$ Å for the X $^2\Pi_g$ state and $R_{\mathrm{N-N}} = 1.185$ Å for the B $^2\Sigma_u^+$ state. At these geometries both the states are characterized by predominant 1h excitations from the reference wave function $N_3{}^-$, with the B $^2\Sigma_u^+$ state having some 2h-1p contributions. Consequently, the effect of including higher-level correlations, such as 3h-2p, is less pronounced than in C_2N and CNC. Nevertheless, there are still noticeable improvements when going from IP-EOMCC(2h-1p) to IP-EOMCC(3h-2p). As evidenced in Table 1 of Ref. [162] and Table V of Ref. [68] (*cf.*, also Table 4.3 of this work), the lower-order IP-EOMCCSD approach estimates the B $^2\Sigma_u^+$– X $^2\Pi_g$ energy to be 4.640 eV, which exhibits a discrepancy of about 0.1 eV compared to the experimental value of 4.555 eV reported in Ref. [258]. The IP-EOMCC(3h-2p) method, that includes a complete treatment of 3h-2p components of $R_\mu^{(+1)}$, reduces this error to 0.04 eV, marking a significant improvement. Therefore, it would be interesting to see how effective the i-FCIQMC driven IP-EOMCC(P) approaches are in recovering IP-EOMCC(3h-2p) energetics. In all the i-FCIQMC calculations the D_{2h} Abelian subgroup of the true point

group of the molecule $D_{\infty h}$ was utilized. The underlying i-FCIQMC propagations were set to converge the lowest energy states of B_{3g} (D_{2h}) symmetry for the X $^2\Pi_g$ and the B_{1u} (D_{2h}) symmetry in case of the B $^2\Sigma_u^+$ state at their respective equilibrium geometries. The results of our semi-stochastic IP-EOMCC(P) approach in recovering total electronic energies of the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states of N_3 and the corresponding X $^2\Pi_g$ − B $^2\Sigma_u^+$ adiabatic excitation energy is reported in Table 4.3 and Fig. 4.4.

From Table 4.3, it can be observed that at $\tau = 0$, the IP-EOMCC(P) calculated total electronic energies of the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states are approximately 13-15 millihartree away from the IP-EOMCC(3h-2p) energies. As soon as the accumulation of 3h-2p determinants defining the $R^{(\mathrm{MC})}_{\mu,3h\text{-}2p}$ component of the electron ionizing operator $R^{(-1)}_{\mu}$ begins, the IP-EOMCC(P) energies quickly start converging towards the IP-EOMCC(3h-2p) energetics. For the X $^2\Pi_g$ state, at just 4000 MC iterations and with 28.7% of 3h-2p determinants in the P space, the IP-EOMCC(P) approach reduces the 13 millihartree error in the IP-EOMCCSD method to 2.627 millihartree. By 10000 MC iterations, IP-EOMCC(P) is already within ∼1 millihartree relative to IP-EOMCC(3h-2p), with the underlying P space containing about 40% of 3h-2p terms. The situation is similar in case of the B $^2\Sigma_u^+$ state, where the ∼15 millihartree error in the IP-EOMCC(2h-1p) is reduced to 5 millihartree after only 4000 MC iterations and with 32% of 3h-2p determinants. At 10000 MC iterations, this error is further decreased to 1.340 millihartree, and at this state the underlying P space is populated by 48.1% 3h-2p determinants. In case of the X $^2\Pi_g$ − B $^2\Sigma_u^+$ gap, convergence is much easier due to favorable cancellation of errors. Here, IP-EOMCC(2h-1p) already has a small error of 0.042 eV relative to IP-EOMCC(3h-2p), but the benefit of using the IP-EOMCC(P) approach is still evident. At only 10000 MC iterations, this already small energy gap of 0.042 eV is further decreased to 0.006 eV, demonstrating the ability of the semi-stochastic IP-EOMCC(P) approach to improve the IP-EOMCCSD energetics, even when the scope of improvement is not very large.

Table 4.3 Convergence of the IP-EOMCC(P) energies [abbreviated as IP(P)] of the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states of N_3, as described by the DZP[4s2p1d] basis set of Refs. [167, 168], and of the corresponding adiabatic excitation energy toward their parent IP-EOMCC(3h-2p) values. The geometries of the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states, optimized in the SAC-CI SDT-R/PS calculations using the same basis set, were taken from Ref. [162]. The P spaces used in the IP-EOMCC(P) calculations were defined as all 1h and 2h-1p determinants and subsets of 3h-2p determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The i-FCIQMC calculations preceding the IP-EOMCC(P) steps were initiated by placing 500 walkers on the ROHF reference determinants of the corresponding states and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbitals of the nitrogen atoms were kept frozen.

MC	X $^2\Pi_g$		B $^2\Sigma_u^+$		Ad. Excit. Energy
Iters.	IP(P)[a]	%(3h-2p)[b]	IP(P)[a]	%(3h-2p)[b]	ΔE[c]
0	13.078[d]	0	14.623[d]	0	0.042[d]
4000	2.627	28.7	5.035	31.9	0.066
10000	1.127	41.4	1.340	48.1	0.006
20000	0.488	53.8	0.343	61.1	−0.004
30000	0.271	61.3	0.134	68.2	−0.004
40000	0.189	68.4	0.036	73.9	−0.004
50000	0.104	72.5	0.012	77.2	−0.003
60000	0.082	75.9	0.004	80.1	−0.002
100000	0.052	84.3	0.000	85.7	−0.001
∞	−163.729333[e]		−163.560374[e]		4.598[f]

[a] Unless otherwise stated, all energies are reported as errors relative to IP-EOMCC(3h-2p) in millihartree.
[b] The %(3h-2p) values are the percentages of 3h-2p determinants captured during the i-FCIQMC propagations (the $S_z = 1/2$ 3h-2p determinants of the B_{3g} symmetry in the case of the X $^2\Pi_g$ state and the B_{1u} symmetry in the case of the B $^2\Sigma_u^+$ state).
[c] Unless otherwise specified, the adiabatic excitation energies are reported as errors relative to IP-EOMCC(3h-2p) in eV; ΔE = B $^2\Sigma_u^+$ − X $^2\Pi_g$.
[d] Equivalent to IP-EOMCC(2h-1p).
[e] Total IP-EOMCC(3h-2p) energy in hartree.
[f] The IP-EOMCC(3h-2p) adiabatic excitation energy in eV.

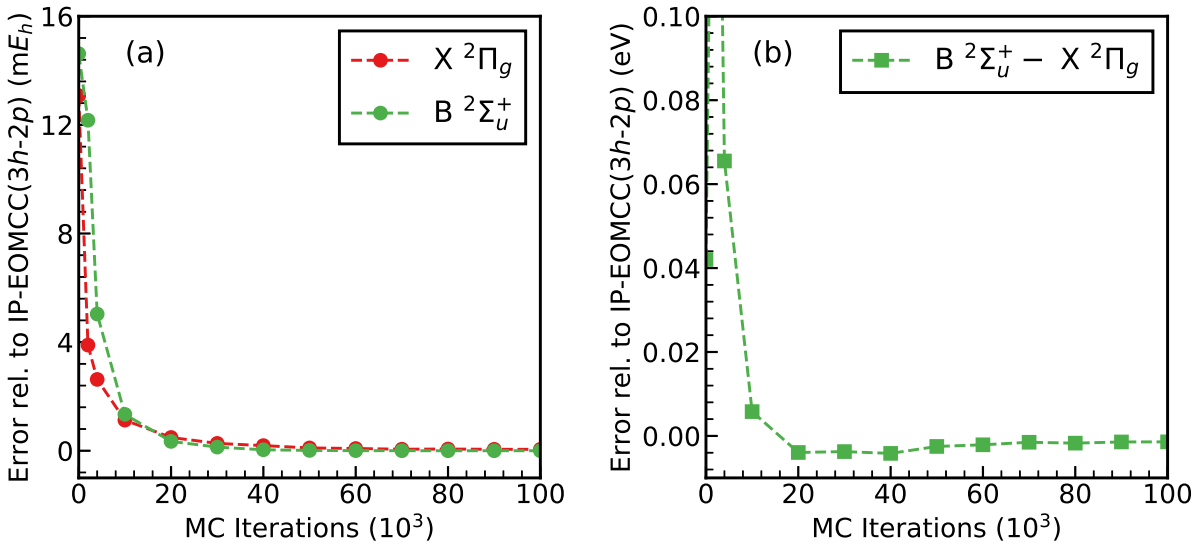

Figure 4.4 Convergence of (a) the IP-EOMCC(P) energies of the X $^2\Pi_g$ and B $^2\Sigma_u^+$ states of N_3, as described by the DZP[4s2p1d] basis set, and (b) the corresponding adiabatic excitation energy toward their parent IP-EOMCC(3h-2p) values.

4.2.2.2 NCO

Our final example is the NCO radical, where we investigated the ground X $^2\Pi$ state and two low-lying valence excited doublet states A $^2\Delta$ and B $^2\Pi$ and the corresponding A $^2\Delta$ $-$ X $^2\Pi$ and B $^2\Pi$ $-$ X $^2\Pi$ excitation energies. Following Ref. [162], we employed the IP SAC-CI-SDT-R/PS / DZP[4s2p1d] optimized geometries, which are $R_{\rm NC}$ = 1.230 Å and $R_{\rm CO}$ = 1.193 Å for the X $^2\Pi$ state, $R_{\rm NC}$ = 1.191 Å and $R_{\rm CO}$ = 1.190 Å for the A $^2\Delta$ state, and $R_{\rm NC}$ = 1.220 Å and $R_{\rm CO}$ = 1.309 Å for the B $^2\Pi$ state. At these geometries, all the low-lying states mentioned above are dominated by 1h excitations from the closed-shell reference ion NCO^- with some 2h-1p contributions being important for the B $^2\Pi$ state. So, the basic IP-EOMCCSD approach is fairly reasonable in this case, however, similar to the case of N_3, the effect of incorporating 3h-2p correlations using the $R_{\mu,3h\text{-}2p}$ component of the electron ionizing operator $R_\mu^{(-1)}$ offers non-negligible improvements. This is evident from Table 1 of Ref. [162] and Table V of Ref. [68] (see also, Table 4.4 of this work), where the IP-EOMCCSD method predicts the A $^2\Delta$ $-$ X $^2\Pi$ and B $^2\Pi$ $-$ X $^2\Pi$ excitation energies to be 2.900 eV and 4.199 eV, respectively, which are $\sim$0.1–0.2 eV from the experimental results of 2.821 eV and 3.937 eV reported in Ref. [258]. The IP-EOMCC(3h-2p) approach further improves these results bringing them to within 0.04 eV or better relative to the experimental values. Now, it would be interesting to see how the semi-stochastic, i-FCIQMC-driven, IP-EOMCC(P) method performs in this case. All the results pertaining to the IP-EOMCC(P) approach are reported in Table 4.4 and Fig. 4.5. In all the i-FCIQMC calculations we used the C_{2v} Abelian subgroup of the true point group of NCO, $C_{\infty v}$. In particular the underlying i-FCIQMC calculations were set to converge the lowest energy states of B_2, A_1, and B_2 symmetry for the X $^2\Pi$, A $^2\Delta$, and B $^2\Pi$ states at the respective equilibrium geometries reported above.

The first row of Table 4.4 contains IP-EOMCC(P) results for $\tau = 0$, which is the start of i-FCIQMC wave function propagations. At this stage, the P spaces utilized in the IP-EOMCC(P) calculations do not contain any 3h-2p determinants, and, as described earlier,

this is equivalent to the IP-EOMCCSD method and for the X $^2\Pi$ state, this energy is about 10 millihartree away from the high-level IP-EOMCC(3h-2p) energy. This error is readily decreased to 3.503 millihartree as soon as the i-FCIQMC propagations complete 4000 steps and recover 23% of 3h-2p determinants. When the i-FCIQMC propagations reach 10000 MC iterations, this error is further reduced to 1.862 millihartree and with an additional 10000 MC iterations the errors are in the sub-millihartree regime. In case of the A $^2\Delta$ state, IP-EOMCC(2h-1p) has an error of 11 millihartree relative to IP-EOMCC(3h-2p). At 4000 MC iterations and with $\sim$24% of 3h-2p determinants in the P space, the IP-EOMCC(P) approach brings this error down to 3 millihartree and after 10000 MC iterations the IP-EOMCC(P) energy is in the sub-millihartree region relative to IP-EOMCC(3h-2p). At this stage, the P space contains 37.4% of 3h-2p determinants pertaining to the A $^2\Delta$ state. In case of the A $^2\Delta$ $-$ X $^2\Pi$ energy gap, the convergence is even faster. At the start of the i-FCIQMC propagation, when the P spaces do not contain any 3h-2p determinants, the error in the adiabatic excitation energy relative to IP-EOMCC(3h-2p) is only 0.036 eV. However, as soon as we start including 3h-2p determinants in the P spaces pertaining to the IP-EOMCC(P) computations, this error rapidly decreases as can be seen from Table 4.4 and Fig. 4.5. In case of the B $^2\Pi$ state, which is the most difficult to describe between the three states of NCO investigated here, the IP-EOMCCSD method produces an error of $>$ 20 millihartree relative to the IP-EOMCC(3h-2p) method. At 4000 MC iterations and with 26% of 3h-2p determinants in the P space, the IP-EOMCC(P) energies reach within $\sim$5 millihartree to the IP-EOMCC(3h-2p) energetics. With an additional 6000 MC iterations, this is further reduced to 1.964 millihartree and at this stage the P spaces contain $\sim$42% of 3h-2p determinants. The B $^2\Pi$ $-$ X $^2\Pi$ adiabatic excitation energies are much easier to converge to the IP-EOMCC(3h-2p) results. The IP-EOMCCSD predicted energy gap, which has an error of 0.288 eV is already within 0.05 eV at 4000 MC iterations and at 1000 MC iterations this error is a mere 0.003 eV.

The two above examples demonstrate that the semi-stochastic IP-EOMCC(P) method is

capable of rapidly converging the IP-EOMCC($3h$-$2p$) energetics for the individual electronic states as well as the corresponding adiabatic excitation energies of radicals with very small percentages of $3h$-$2p$ determinants in the respective P spaces.

Table 4.4 Convergence of the IP-EOMCC(P) energies [abbreviated as IP(P)] of the X $^2\Pi$, A $^2\Delta$, and B $^2\Pi$ states of NCO, as described by the DZP[4s2p1d] basis set of Refs. [167, 168], and the corresponding adiabatic excitation energies toward their parent IP-EOMCC(3h-2p) values. The geometries of the X $^2\Pi$, A $^2\Delta_u$, and B $^2\Pi$ states, optimized in the SAC-CI SDT-R/PS calculations using the same basis set, were taken from Ref. [162]. The P spaces used in the IP-EOMCC(P) calculations were defined as all 1h and 2h-1p determinants and subsets of 3h-2p determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The i-FCIQMC calculations preceding the IP-EOMCC(P) steps were initiated by placing 500 walkers on the ROHF reference determinants of the corresponding states and the n_a parameter of the initiator algorithm was set at 3. In all post-Hartree–Fock calculations, the lowest core orbitals of the carbon, nitrogen, and oxygen atoms were kept frozen.

MC	X $^2\Pi$		A $^2\Delta$		B $^2\Pi$		Ad. Excit. Energy	
Iters.	IP(P)[a]	%(3h-2p)[b]	IP(P)[a]	%(3h-2p)[b]	IP(P)[a]	%(3h-2p)[b]	ΔE_1[c]	ΔE_2[c]
0	9.587[d]	0	10.921[d]	0	20.154[d]	0	0.036[d]	0.288[d]
4000	3.503	23.2	3.198	23.6	5.302	25.6	−0.008	0.049
10000	1.862	34.9	0.991	37.4	1.964	41.5	−0.024	0.003
20000	0.973	46.6	0.359	52.5	0.943	54.9	−0.017	−0.001
30000	0.428	55.0	0.153	61.4	0.428	63.9	−0.008	0.000
40000	0.329	62.5	0.102	68.0	-0.106	70.2	−0.006	−0.012
50000	0.164	68.1	0.032	72.6	0.003	74.8	−0.004	−0.004
60000	0.097	72.4	0.015	76.1	0.133	77.9	−0.002	0.001
100000	0.011	81.8	0.000	84.1	0.002	85.2	0.000	0.000
∞	−167.591596[e]		−167.486358[e]		−167.447865[e]		2.864[f]	3.911[f]

[a] Unless otherwise stated, all energies are reported as errors relative to IP-EOMCC(3h-2p) in millihartree.

[b] The %(3h-2p) values are the percentages of 3h-2p determinants captured during the i-FCIQMC propagations (the $S_z = 1/2$ 3h-2p determinants of the B_2 symmetry in the case of the X $^2\Pi$ state, the A_1 symmetry in the case of the A $^2\Delta$ state, and the B_2 symmetry in the case of the B $^2\Pi$ state).

[c] Unless otherwise specified, the adiabatic excitation energies are reported as errors relative to IP-EOMCC(3h-2p) in eV; ΔE_1 = A $^2\Delta$ − X $^2\Pi$ and ΔE_2 = B $^2\Pi$ − X $^2\Pi$.

[d] Equivalent to IP-EOMCC(2h-1p).

[e] Total IP-EOMCC(3h-2p) energy in hartree.

[f] The IP-EOMCC(3h-2p) adiabatic excitation energy in eV.

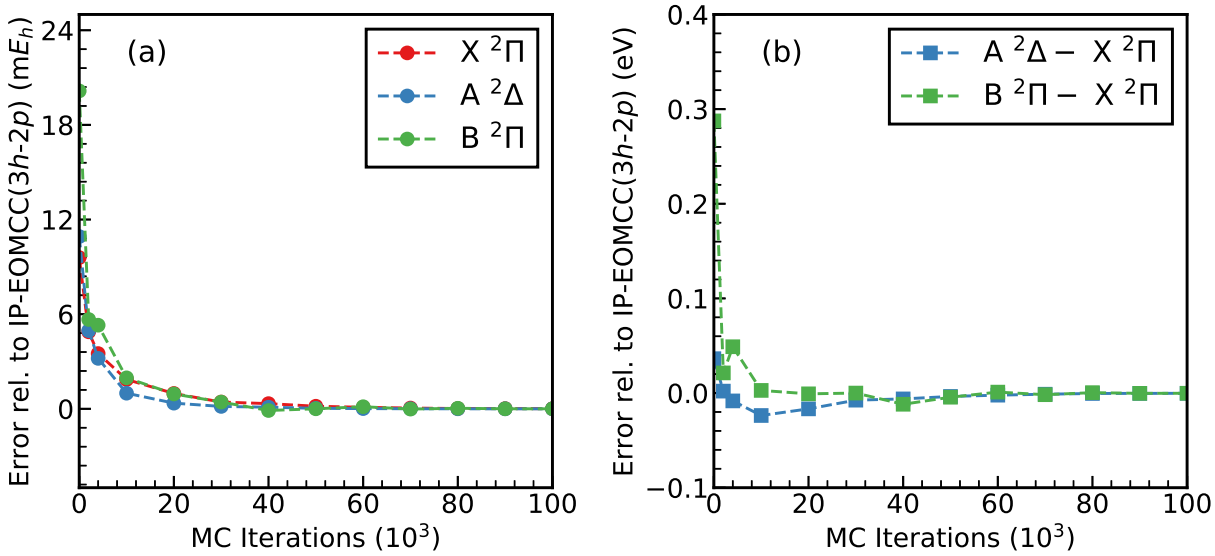

Figure 4.5 Convergence of (a) the IP-EOMCC(P) energies of the X $^2\Pi$, A $^2\Delta$, and B $^2\Pi$ states of NCO, as described by the DZP[4s2p1d] basis set, and (b) the corresponding adiabatic excitation energies toward their parent IP-EOMCC(3h-2p) values.

4.3 Singlet–Triplet Gaps in Methylene and Trimethylenemethane

After exploring the performance of the semi-stochastic EA-EOMCC(P) and IP-EOMCC(P) methods, which describe the (1,0) and (0,1) sectors of the Fock space, we move on to their extensions to the (2,0) and (0,2) sectors of the Fock space, which we refer to as the semi-stochastic DEA- [DEA-EOMCC(P)] and DIP-EOMCC [DIP-EOMCC(P)] approaches. These methods are particularly suitable for studying singlet and triplet states of biradical systems and the corresponding singlet–triplet gaps. To explore the capability of the semi-stochastic DEA-EOMCC(P) and DIP-EOMCC(P) approaches in converging their high-level fully-deterministic parent DEA-EOMCC(4p-2h) and DIP-EOMCC(4h-2p) methods, we carried out benchmark calculations for the lowest singlet and triplet states along with the corresponding singlet–triplet gaps ($\Delta E_{\text{S-T}} = E_{\text{singlet}} - E_{\text{triplet}}$) in the CH_2 [Tables 4.5 and 4.6 and Figs. 4.6 and 4.7] and trimethylenemethane [Table 4.7 and Figs. 4.9 and 4.10]. We utilized the TZ2P basis set [180] and FCI/TZ2P geometries as reported in Ref. [181] for the CH_2 molecule and for trimethylenemethane, we employed the 6-31G(d) basis set [178, 179] with the SF-DFT/6-31G(d) geometries reported in Ref. [231]. For the smaller CH_2 molecule we used i-FCIQMC, while for the larger trimethylenemethane system we exploited the i-CISDTQ-MC propagations to generate the lists of 4p-2h and 4h-2p determinants entering the P spaces for the DEA-EOMCC(P) and DIP-EOMCC(P) calculations. Again, following our previous semi-stochastic work [99–101, 137], we used the HANDE software package [170, 171] to execute all our QMC calculations. Our standalone CC/EOMCC codes, interfaced with the RHF, ROHF, and integral transformation routines available in the GAMESS software package [172–174], were used to carry out the DEA-EOMCC(P), DIP-EOMCC(P), and the fully deterministic CCSD, DEA-EOMCC(3p-1h), DEA-EOMCC(4p-2h), DIP-EOMCC(3h-1p), and DIP-EOMCC(4h-2p) computations. Each i-FCIQMC and i-CISDTQ-MC run was initiated by placing 500 walkers on the relevant reference function (see Tables 4.5–4.7 for the details), the initiator parameter n_a was set at 3, and all of the i-FCIQMC and i-CISDTQMC propagations used a time step of $\tau = 0.0001$ a.u. If the true point group of the system

of interest was not Abelian, which was the case for trimethylenemethane, we utilized the largest Abelian subgroups in the calculations. This choice was necessary, since all of our CC/EOMCC codes interfaced with GAMESS and the CIQMC routines in HANDE can only handle Abelian symmetries. In all the post-HF computations, the core electrons corresponding to the 1s shells of the carbon atom was kept frozen.

4.3.1 Methylene

The discussions for the semi-stochastic DEA- and DIP-EOMCC methods, DEA-EOMCC(P) and DIP-EOMCC(P), begin with the results for total electronic energies of the X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states of methylene and the corresponding singlet–triplet gaps. In the DEA-EOMCC calculations we used the CH_2^{2+} dication as the reference and for the DIP-EOMCC approach the CH_2^{2-} dianion was used as the reference. The geometries used were the FCI/TZ2P geometries as reported in Ref. [181] for the CH_2 molecule and throughout this work we utilized the TZ2P basis set [180]. All the results of the DEA-EOMCC(P) and DIP-EOMCC(P) calculations are reported in Tables 4.5 and 4.6 and Figs. 4.6 and 4.7. The ground X ^{3}B$_1$ and the second excited state B ^{1}B$_1$ are characterized as having a SR character that can be well represented by high-spin triplet and open-shell singlet configurations of $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^1(1b_1)^1$ type. The first and the third excited states, A ^{1}A$_1$ and C ^{1}A$_1$, have a significant MR character originating due to the mixing of the two configurations $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2$ and $(1a_1)^2(2a_1)^2(1b_2)^2(1b_1)^2$, which makes it very challenging for many electronic structure methods to provide an accurate description. As a result, in order to obtain accurate results for the A ^{1}A$_1$ − X ^{3}B$_1$, B ^{1}B$_1$ − X ^{3}B$_1$, and C ^{1}A$_1$ − X ^{3}B$_1$ singlet–triplet gaps, one needs to use methods that can offer well balanced treatments of both dynamical and nondynamical correlations. This makes CH_2 a very popular example to test the performance of newly developed electronic structure methods. The most basic DEA-EOMCCSD [DEA-EOMCC(3p-1h)] and DIP-EOMCCSD [DIP-EOMCC(3h-1p)] fail to provide an accurate description of the triplet ground and the low-lying singlet states of methylene and the corresponding singlet–triplet gaps, and one needs to incorporate 4p-2h

and $4h$-$2p$ correlations for an accurate description. From Table 1 of Ref. [153] one can compare the FCI, DEA-EOMCC($3p$-$1h$), DEA-EOMCC($4p$-$2h$), DIP-EOMCC($3h$-$1p$), and DIP-EOMCC($4h$-$2p$) calculated singlet–triplet gaps $\mathrm{A}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{B}_1$, $\mathrm{B}\,^1\mathrm{B}_1 - \mathrm{X}\,^3\mathrm{B}_1$, and $\mathrm{C}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{B}_1$. It is evident that the DEA-EOMCC($3p$-$1h$) predicted gaps are -0.11, -1.89, and -3.64 kcal/mol in error compared to the FCI singlet–triplet gaps of 11.14, 35.59, and 61.67 kcal/mol, respectively. The complete incorporation of $4p$-$2h$ correlations via the DEA-EOMCC($4p$-$2h$) methodology improves these results, especially for the second and third gaps, bringing down the errors to 0.38, -0.02, and 0.21 kcal/mol compared to FCI. On the other hand, the DIP-EOMCC($3h$-$1p$) method predicts these singlet–triplet gaps to be -4.53, -3.22, and -4.63 kcal/mol away from FCI. We see a significant improvement by using the DIP-EOMCC($4h$-$2p$) approach, which reduces these errors to only -0.44, -0.51, and -0.48 kcal/mol relative to FCI. So, it would be interesting to see how our DEA-EOMCC(P) and DIP-EOMCC(P) approaches perform in this interesting case of methylene. We used the i-FCIQMC method to extract the lists of $4p$-$2h$ determinants, in case of DEA-EOMCC(P), and the $4h$-$2p$ determinants, in case of the DIP-EOMCC(P) approach. The i-FCIQMC runs were initiated on the ROHF determinant of B_1 symmetry for the $\mathrm{X}\,^3\mathrm{B}_1$ state and the RHF determinants of A_1 symmetries for the $\mathrm{A}\,^1\mathrm{A}_1$ and $\mathrm{C}\,^1\mathrm{A}_1$ states, and the A_2 symmetry $\mathrm{B}\,^1\mathrm{B}_1$, and with the one- and two-body integrals obtained from the reference ionic systems (CH_2^{2+} in case of DEA-EOMCC(P) and CH_2^{2-} for the DIP-EOMCC(P) case). In the next paragraph, we explore the performance of the semi-stochastic, i-FCIQMC-driven, DEA-EOMCC(P) approach and in the following paragraph we investigate the performance of the DIP-EOMCC(P) approach.

From Table 4.5, one can notice that at $\tau = 0$, the DEA-EOMCC(P) [equivalent to the DEA-EOMCC($3p$-$1h$)] energies are 14 millihartree in error compared to DEA-EOMCC($4p$-$2h$) for the $\mathrm{X}\,^3\mathrm{B}_1$ state. As soon as we start incorporating $4p$-$2h$ determinants in the P spaces defining the $R^{(\mathrm{MC})}_{\mu,4p\text{-}2h}$ components of $R^{(+2)}_{\mu}$ operator in the DEA-EOMCC(P) calculations, the energies rapidly converge to DEA-EOMCC($4p$-$2h$). At 4000 MC iterations and with just

1.8% of 4p-2h determinants in the P space, the DEA-EOMCC(P) errors are already within 5 millihartree and as soon as the i-FCIQMC propagations select 7% of 4p-2h determinants, the error in the DEA-EOMCC(P) energy becomes only ∼1 millihartree. This happens at just 10000 MC iterations, and if we allow i-FCIQMC to complete 50000 MC steps, the P spaces contain 15.3% of 4p-2h determinants and the error in energy of the X $^3B_1$ state reaches the sub-millihartree region. This fast convergence to DEA-EOMCC(4p-2h) remains true even when we go to the first excited singlet state A $^1A_1$, where the $>$ 13 millihartree error in the DEA-EOMCC(3p-1h) energies are quickly reduced to $<$ 2 millihartree at only 10000 MC iterations and with just 4.3% of 4p-2h determinants in the P space. The situation improves even more when the i-FCIQMC propagations reach 20000 MC time steps, where with only 7.6% of 4p-2h determinants the total electronic energy is already in the sub-millihartree regime. If we look at the A $^1A_1$ − X $^3B_1$ singlet triplet gap, it can be seen that the 0.487 kcal/mol error at $\tau = 0$ is quickly reduced to 0.2 kcal/mol at already 10000 MC iterations and by the time we reach 50000 MC iterations, the error is less than 0.1 kcal/mol relative to DEA-EOMCC(4p-2h). For the next state B $^1B_1$, which is mostly single reference in nature, the error at $\tau = 0$ is 11.129 millihartree. However, with only ∼5% and ∼8% of 4p-2h determinants in the P spaces and at just 10000 and 20000 MC iterations, respectively, the ∼2 millihartree and ∼1 millihartree marks are reached by the errors in the total electronic energy predicted by DEA-EOMCC(P). The situation in the corresponding B $^1B_1$ − X $^3B_1$ singlet–triplet gap is more favorable. For example, the 1.875 kcal/mol errors in the DEA-EOMCC($P|_{\tau=0}$) method is quickly brought down to 0.1 kcal/mol relative to DEA-EOMCC(4p-2h) at only 10000 MC time steps, with additional iterations further improving the results as seen in Table 4.5 and in Fig. 4.6. In case of the most challenging C $^1A_1$ state, the errors in the DEA-EOMCC(3p-1h) energies relative to DEA-EOMCC(4p-2h) is 7.987 millihartree. As the underlying i-FCIQMC wave function propagation gradually select more and more 4p-2h determinants, this error quickly gets down to the sub-millihartree regime. At 20000 MC iteration with 5% of 4p-2h determinants in the P space the DEA-

EOMCC(P) energies are about 3 millihartree away from the DEA-EOMCC(4p-2h) method and after 50000 MC iterations and with 13% of 4p-2h determinants in the P space, these errors are less than 1 millihartree. This is also reflected in the C $^1A_1$ − X $^3B_1$ singlet–triplet gap. The 3.847 kcal/mol error at $\tau = 0$ rapidly becomes $\sim$1 kcal/mol in absolute value at 20000 MC iterations and by the time the i-FCIQMC propagations complete 50000 time steps, the absolute value of the error is only about 0.2 kcal/mol from the high-level DEA-EOMCC(4p-2h) result. It is interesting to note that, due to the different rates of convergence of the low-lying states, the singlet–triplet gaps fluctuate before converging to the DEA-EOMCC(4p-2h) results.

Table 4.5 Convergence of the DEA-EOMCC(P) energies [abbreviated as DEA(P)] of the X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states of methylene, as described by the TZ2P basis set of Ref. [180], and of the corresponding adiabatic singlet–triplet gaps toward their parent DEA-EOMCC(4p-2h) values. The geometries of the X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states, optimized in the FCI calculations using the TZ2P basis set, were taken from Ref. [181]. The P spaces used in the DEA-EOMCC(P) calculations were defined as all 2p and 3p-1h determinants and subsets of 4p-2h determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The i-FCIQMC calculations were initiated by placing 500 walkers on the corresponding ROHF reference determinant for the X ^{3}B$_1$ state and the corresponding RHF reference determinants for the remaining states and the n_a parameter of the initiator algorithm was set at 3. In all the post-Hartree–Fock calculations, the lowest core orbital of the carbon atom was kept frozen.

MC	X ^{3}B$_1$		A ^{1}A$_1$		B ^{1}B$_1$		C ^{1}A$_1$		Singlet–Triplet Gap		
Iters.	DEA(P)[a]	%(4p-2h)[b]	DEA(P)[a]	%(4p-2h)[b]	DEA(P)[a]	%(4p-2h)[b]	DEA(P)[a]	%(4p-2h)[b]	ΔE_1[c]	ΔE_2[c]	ΔE_3[c]
0	14.118[d]	0.0	13.342[d]	0.0	11.129[d]	0.0	7.987[d]	0.0	0.487[d]	1.875[d]	3.847[d]
4000	4.863	1.8	4.611	1.5	4.477	2.0	6.344	1.3	0.158	0.242	−0.929
10000	2.252	4.4	1.953	4.3	2.085	4.8	4.794	3.3	0.188	0.105	−1.595
20000	1.323	7.0	0.996	7.6	1.163	7.9	3.161	5.1	0.205	0.101	−1.154
50000	0.448	15.3	0.316	20.1	0.394	18.3	0.701	13.2	0.083	0.034	−0.158
100000	0.079	38.2	0.028	51.9	0.050	48.5	0.078	40.7	0.032	0.018	0.001
200000	0.001	85.2	0.000	88.9	0.000	89.3	0.000	88.0	0.000	0.000	0.000
∞	−39.066449[e]		−39.048089[e]		−39.009764[e]		−38.967833[e]		11.521[f]	35.570[f]	61.881[f]

[a] Unless otherwise stated, all energies are reported as errors relative to DEA-EOMCC(4p-2h) in millihartree.

[b] The %(4p-2h) values are the percentages of 4p-2h determinants captured during the i-FCIQMC propagations [the $S_z = 1$ 4p-2h determinants of the B$_1$ symmetry in the case of the X ^{3}B$_1$ state, the $S_z = 0$ 4p-2h determinants of the B$_1$ symmetry in case of the B ^{1}B$_1$ state , and the $S_z = 0$ 4p-2h determinants of the A$_1$ symmetry in the case of the A ^{1}A$_1$, and the A$_1$ symmetry in case of the C ^{1}A$_1$ state].

[c] Unless otherwise specified, the singlet–triplet gaps are reported as errors relative to DEA-EOMCC(4p-2h) in kcal/mol; ΔE_1 = A ^{1}A$_1$ − X ^{3}B$_1$, ΔE_2 = B ^{1}B$_1$ − X ^{3}B$_1$, and ΔE_3 = C ^{1}A$_1$ − X ^{3}B$_1$.

[d] Equivalent to DEA-EOMCC(3p-1h).

[e] Total DEA-EOMCC(4p-2h) energy in hartree.

[f] The DEA-EOMCC(4p-2h) singlet–triplet gap in kcal/mol.

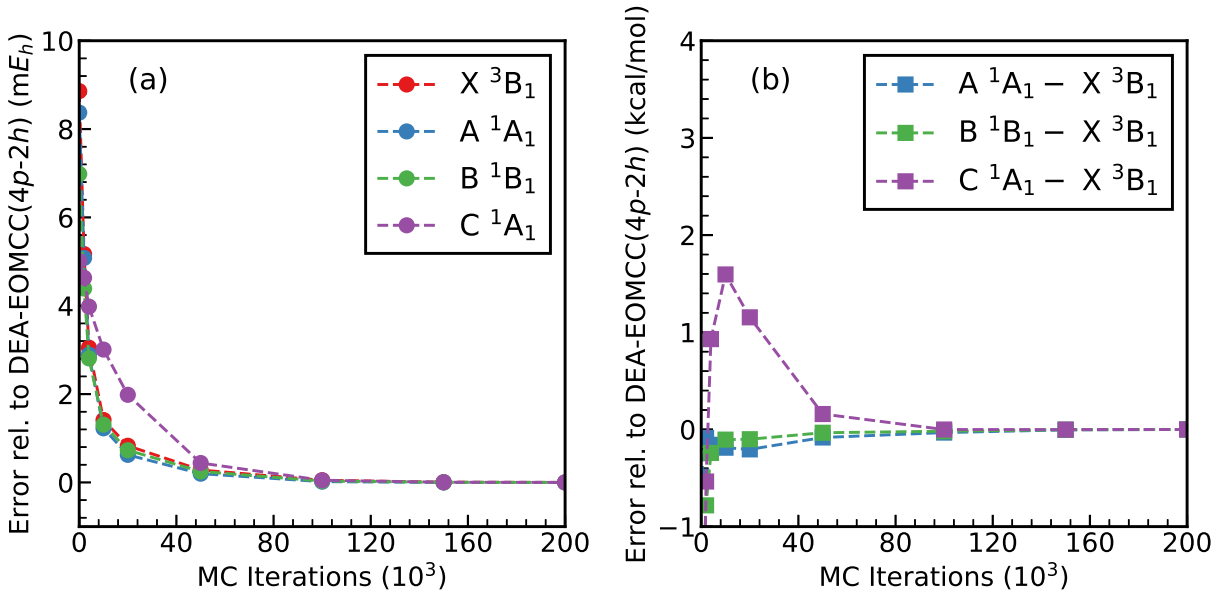

Figure 4.6 Convergence of (a) DEA-EOMCC(P) energies of X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states of methylene, as described by the TZ2P basis set, and (b) of the corresponding adiabatic singlet–triplet gaps towards their parent DEA-EOMCC(4p-2h) values.

Table 4.6 contains the results of our semi-stochastic, i-FCIQMC-driven, DIP-EOMCC(P) computations. At the start of the i-FCIQMC propagations, the total electronic energy of the X ^{3}B$_1$ state calculated by the DIP-EOMCC(P) approach has an error of 26.147 millihartree relative to DIP-EOMCC(4h-2p). This error is readily decreased to about 3 millihartree at only 4000 MC iterations, where the P spaces contain about 35% of 4h-2p determinants in the P space. Allowing the i-FCIQMC propagations to complete a total of 10000 and 20000 MC iterations results in errors of 1.357 and 0.766 millihartree, respectively. The percentage of 4h-2p determinants in the P spaces at this stage are 43.4% and 50.3%, respectively. For the first excited singlet state A ^{1}A$_1$, the 19.624 millihartree error reported at $\tau = 0$ is sharply decreased to 3 millihartree at just 4000 MC iteration and with $\sim$19% of 4h-2p determinants in the P space. By the time the i-FCIQMC propagations reach 10000 MC time steps, this error is already below 1 millihartree and this happens with 30% of 4h-2p determinants in the P space. Similarly, the -4.093 kcal/mol error in the A ^{1}A$_1$ $-$ X ^{3}B$_1$ singlet–triplet gap narrows down to just -0.152 kcal/mol at 4000 MC iterations and by the time 50000 MC time steps are completed, this error is only -0.028 millihartree away from the DIP-EOMCC(4h-2p) value. The next state, which is an open-shell singlet state, is referred to as B ^{1}B$_1$, and in this case the $\sim$22 millihartree errors in the DIP-EOMCC(3h-1p) energetics improves to only 3.681 millihartree at just 4000 MC iterations and with $\sim$33% of 4h-2p determinants in the P space. Then with i-FCIQMC propagation it steadily converges to DIP-EOMCC(4h-2p). By the time i-FCIQMC completes 20000 MC iterations, the errors in the total electronic energies in this state falls in the sub-millihartree regime. At this stage, the B ^{1}B$_1$ $-$ X ^{3}B$_1$ singlet–triplet gap is already converged to 0.032 kcal/mol compared to the target DIP-EOMCC(4h-2p) values, and this is a significant improvement from the DIP-EOMCC($P|_{\tau=0}$) result of -2.714. The final state considered here is the C ^{1}A$_1$ state, which has a strong multireference character as described above. In this case, the DIP-EOMCC(P) calculations start with an error of 19.355 millihartree at $\tau = 0$. After 4000 MC time steps and with 17.3% of 4h-2p determinants in the P space, this error is down to 8.869 millihartree. It gradually

decreases to 3.480 millihartree at 20000 MC iterations before reaching the sub-millihartree mark at 50000 MC iterations. At these stages, the percentages of the $4h$-$2p$ determinants are 36.6% and 51.7%, respectively. The convergence of the C $^1A_1$ − X $^3B_1$ singlet–triplet gap mirrors this convergence pattern. The −4.149 kcal/mol error in this gap shrinks to 1.703 kcal/mol at 20000 MC iterations and by the time one reaches 50000 MC iterations the error is only 0.149 kcal/mol (*cf.*, Fig. 4.7 for a graphical representation).

Table 4.6 Convergence of the DIP-EOMCC(P) energies [abbreviated as DIP(P)] of the X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states of methylene, as described by the TZ2P basis set of Ref. [180], and of the corresponding adiabatic singlet–triplet gaps toward their parent DIP-EOMCC(4h-2p) values. The geometries of the X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states, optimized in the FCI calculations using the TZ2P basis set, were taken from Ref. [181]. The P spaces used in the DIP-EOMCC(P) calculations were defined as all 2h and 3h-1p determinants and subsets of 4h-2p determinants extracted from the i-FCIQMC propagations with $\delta\tau = 0.0001$ a.u. The i-FCIQMC calculations were initiated by placing 500 walkers on the corresponding ROHF reference determinant for the X ^{3}B$_1$ state and the corresponding RHF reference determinants for the remaining states and the n_a parameter of the initiator algorithm was set at 3. In all the post-Hartree–Fock calculations, the lowest core orbital of the carbon atom was kept frozen.

MC	X ^{3}B$_1$		A ^{1}A$_1$		B ^{1}B$_1$		C ^{1}A$_1$		Singlet–Triplet Gap		
Iters.	DIP(P)[a]	%(4h-2p)[b]	DIP(P)[a]	%(4h-2p)[b]	DIP(P)[a]	%(4h-2p)[b]	DIP(P)[a]	%(4h-2p)[b]	ΔE_1[c]	ΔE_2[c]	ΔE_3[c]
0	26.147[d]	0.0	19.624[d]	0.0	21.821[d]	0.0	19.355[d]	0.0	−4.093[d]	−2.714[d]	−4.149[d]
4000	3.289	34.6	3.046	19.4	3.681	32.9	8.869	17.3	−0.152	0.246	3.502
10000	1.357	43.4	0.946	29.8	1.853	42.0	5.741	28.2	−0.257	0.311	2.751
20000	0.766	50.3	0.524	39.8	0.818	48.6	3.480	36.6	−0.152	0.032	1.703
50000	0.131	59.5	0.085	55.5	0.168	59.0	0.369	51.7	−0.028	0.023	0.149
100000	0.009	68.1	0.005	63.8	0.009	68.1	0.057	62.8	−0.003	0.000	0.030
200000	0.000	70.6	0.000	65.2	0.000	70.6	0.000	65.2	0.000	0.000	0.000
∞	−39.066449[e]		−39.048089[e]		−39.009764[e]		−38.967833[e]		10.701[f]	35.083[f]	61.191[f]

[a] Unless otherwise stated, all energies are reported as errors relative to DIP-EOMCC(4h-2p) in millihartree.

[b] The %(4h-2p) values are the percentages of 4h-2p determinants captured during the i-FCIQMC propagations [the $S_z = 1$ 4h-2p determinants of the B$_1$ symmetry in the case of the X ^{3}B$_1$ state, the $S_z = 0$ 4h-2p determinants of the B$_1$ symmetry in case of the B ^{1}B$_1$ state , and the $S_z = 0$ 4h-2p determinants of the A$_1$ symmetry in the case of the A ^{1}A$_1$, and the A$_1$ symmetry in case of the C ^{1}A$_1$ state].

[c] Unless otherwise specified, the singlet–triplet gaps are reported as errors relative to DIP-EOMCC(4h-2p) in kcal/mol; ΔE_1 = A ^{1}A$_1$ − X ^{3}B$_1$, ΔE_2 = B ^{1}B$_1$ − X ^{3}B$_1$, and ΔE_3 = C ^{1}A$_1$ − X ^{3}B$_1$.

[d] Equivalent to DIP-EOMCC(3h-1p).

[e] Total DIP-EOMCC(4h-2p) energy in hartree.

[f] The DIP-EOMCC(4h-2p) singlet–triplet gap in kcal/mol.

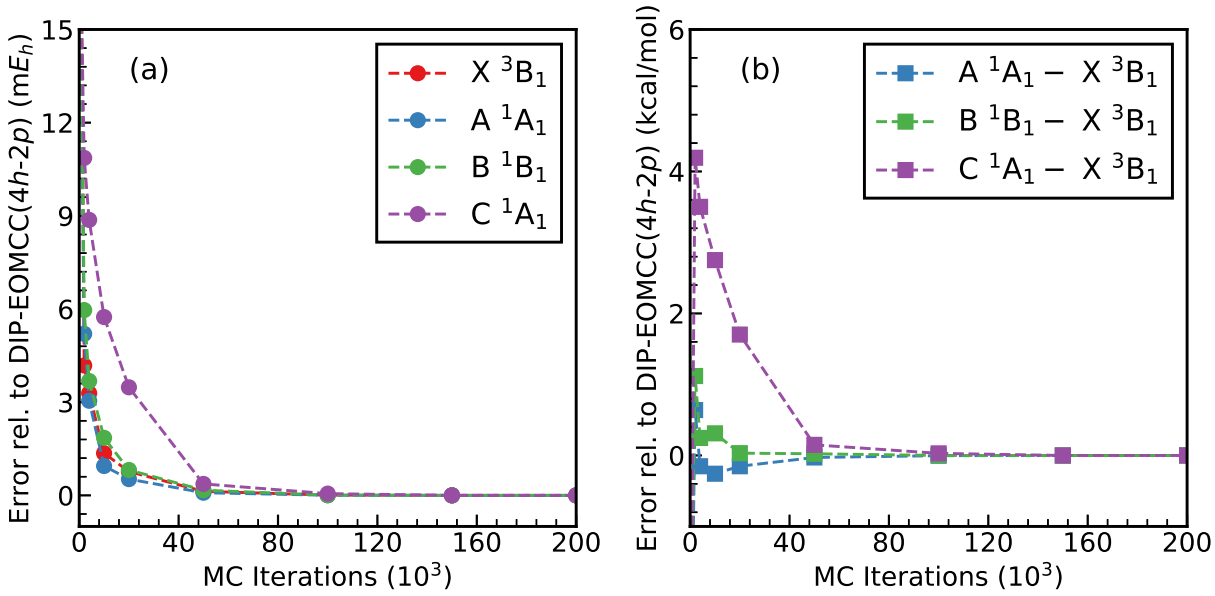

Figure 4.7 Convergence of (a) DIP-EOMCC(P) energies of X ^{3}B$_1$, A ^{1}A$_1$, B ^{1}B$_1$, and C ^{1}A$_1$ states of methylene, as described by the TZ2P basis set, and (b) of the corresponding adiabatic singlet–triplet gaps towards their parent DIP-EOMCC(4h-2p) values.

4.3.2 Trimethylenemethane

Our final example is trimethylenemethane (TMM), a fascinating non-Kekulé hydrocarbon, in which four valence π electrons are de-localized over four closely spaced π-type orbitals. Assuming D_{3h} symmetry, which is the true point group symmetry of the minimum energy structure of the ground-state triplet surface of trimethylenemethane, the four MOs involved in the π-orbital network of trimethylenemethane consists of the nondegenerate $1a_2''$ orbital, the doubly degenerate $1e''$ orbitals, and the nondegenerate $2a_2''$ orbital. If, on the other hand, the symmetry relevant to the low-lying singlet states, C_{2v} is adopted, the $1a_2''$ and $2a_2''$ orbitals in the D_{3h} description becomes $1b_1$ and $3b_1$, respectively, and the doubly degenerate $1e''$ shell splits into the $1a_2$ and $2b_1$ components (see Fig. 3.14 for a pictorial representation of the orbitals). Although the electronic structure of trimethylenemethane has been well understood for decades (*cf.*, *e.g.*, Ref. [213] and references therein), an accurate characterization of its triplet ground state and low-lying singlet states and the energy separation between them continues to represent a significant challenge to quantum chemistry approaches [80, 88, 153, 154, 214–239]. The D_{3h}-symmetric triplet ground state, designated as $X\,^3A_2'$ (in a C_{2v} description adopted in this study, $X\,^3B_2$), which is dominated by the $|\{\text{core}\}(1a_2'')^2(1e_1'')^1(1e_2'')^1|$ configuration (in C_{2v}, $|\{\text{core}\}(1b_1)^2(1a_2)^1(2b_1)^1|$), is relatively easy to describe, but the next two states, which are the nearly degenerate singlets, are not. These states undergo Jahn–Teller distortion that lifts their exact degeneracy in a D_{3h} description, splitting the two states. The lower of the two singlets, characterized by a C_s-symmetric minimum that can be approximated by a twisted C_{2v} structure and is usually designated as the open-shell singlet $A\,^1B_1$ state and the second state is a C_{2v}-symmetric multi configurational state referred to as $B\ ^1A_1$. The first state emerges from the $|\{\text{core}\}(1b_1)^2(1a_2)^1(2b_1)^1|$ configuration, while the second one is dominated by the $|\{\text{core}\}(1b_1)^2(1a_2)^2|$ and $|\{\text{core}\}(1b_1)^2(2b_1)^2|$ closed-shell determinants. The open-shell singlet state $A\,^1B_1$, after the Jahn–Teller distortion is lower in energy compared to the $B\ ^1A_1$ state, but it has not been observed experimentally due to unfavorable Franck–Condon fac-

tors [229, 240], so we do not consider it in this work. The second singlet state B ^{1}A$_1$, on the other hand, has been detected in photoelectron spectroscopy experiments reported in Refs. [240, 241], which located it at 16.1 ± 0.1 kcal/mol above the X ^{3}A$_2'$ ground state. Thus, following our group's previous deterministic, active-orbital-based, CC(P;Q) work [88], the state-of-the-art DEA- and DIP-EOMCC computations with up to 4p-2h and 4h-2p excitations reported in Refs. [80, 153, 154], and the semi-stochastic CC(P;Q) calculations reported in Ref. [102], we focused on the D$_{3h}$-symmetric X ^{3}A$_2'$ ground state and the C$_{2v}$-symmetric B ^{1}A$_1$ singlet state and the adiabatic singlet–triplet gap between them (see, Fig. 4.8 for a schematic illustration of the Jahn–Teller splitting in trimethylenemethane and the singlet–triplet gap targeted). Similar to our previous work, we utilized the geometries optimized using the the spin-flip density functional theory (SF-DFT) and the 6-31G(d) basis reported in Ref. [231]. The purely electronic singlet–triplet gap derived from the experiments, obtained by subtracting the zero-point vibrational energy correction ΔZPVE resulting from the SF-DFT/6-31G(d) calculations reported in Ref. [231] from the experimental B ^{1}A$_1$ – X ^{3}A$_2'$ gap determined in Refs. [240, 241], is 18.1 kcal/mol. However, this estimate depends on the source of the information about the ΔZPVE correction. For example, if one replaces the ΔZPVE value obtained in the SF-DFT/6-31G(d) calculations reported in Ref. [231] by its CCSD(T)/6-311++G(2d,2p) estimate and accounts for the core polarization effects determined with the help of the CCSD(T)/cc-pCVQZ computations, combining the resulting information with the experimental B ^{1}A$_1$ – X ^{3}A$_2'$ separation determined in Refs. [240, 241], the purely electronic, experimentally derived, adiabatic $\Delta E_{\text{S-T}}$ gap increases to 19.4 kcal/mol [237]. The DEA-EOMCC(4p-2h)/6-31G(d) and DIP-EOMCC(4h-2p)/6-31G(d) approaches using the SF-DFT/6-31G(d) geometries predict this gap to be 19.856 kcal/mol and 19.907 kcal/mol, respectively. The remainder of this section describes how the DEA-EOMCC(P) and DIP-EOMCC(P) employing the 6-31G(d) basis set perform in this challenging situation. Following our previous semi-stochastic CC(P;Q) approach [102], we used the i-CISDTQ-MC wave function propagation for constructing the P spaces. Based on our previous experience

in semi-stochastic methods, replacing FCIQMC propagations with CISDTQ-MC will not affect the rate of convergence of our DEA-EOMCC(P) and DIP-EOMCC(P) approaches, while offering additional computational savings in the QMC part. In case of the DEA-EOMCC calculations, we used the TMM^{2+} dication as the closed-shell reference and for the DIP-EOMCC calculations, we used the TMM^{2-} dianion as the reference determinant. The leading $4p$-$2h$ ($4h$-$2p$) determinants captured during the i-CISDTQ-MC propagations at various time steps τ are the determinants of B_1 symmetry with $S_z = 1$ in the case of the X $^3A_2'$ state, and the $S_z = 0$ A_1 symmetric determinants in case of the B $^1A_1$ state. For the CIQMC propagations we used the ROHF reference for the X $^3A_2'$ state and the RHF reference for B $^1A_1$ state. The one- and two-body integrals used in the DEA-EOMCC (DIP-EOMCC) calculations are extracted from the TMM^{2+} (TMM^{2-}) system. All the results of the semi-stochastic calculations are reported in Table 4.7 and Figs. 4.9 and 4.10. We first discuss the results of our DEA-EOMCC(P) calculations in the next paragraph and then the following paragraph contains the results of DIP-EOMCC(P) computations.

The results reported in Table 4.7 and Figure 4.9 demonstrate that, for both the X $^3A_2'$ and B $^1A_1$ states, the large errors which range from 19 to 27 millihartree at $\tau = 0$ get reduced to half its value with just 1% of all $4p$-$2h$ determinants in the stochastically determined P spaces captured by the i-CISDTQ-MC runs which happens at 10000 MC iteration. After the additional 20000 MC time steps, which results in capturing 7–8% of all $4p$-$2h$ determinants, the errors go down to $\sim$3 millihartree for both states. This error in the energies of the two states reaches the sub-millihartree regime when the underlying P spaces contain 18–26% of the $4p$-$2h$ determinants and this happens at 50000 MC iteration. For the B $^1A_1$ – X $^3A_2'$ singlet–triplet gap, $\Delta E_{\text{S-T}}$, the $\sim$5 kcal/mol error relative to DEA-EOMCC($4p$-$2h$) at $\tau = 0$ quickly gets reduced to 1 kcal/mol at 10000 MC iterations. By the time we reach 20000 QMC iterations the singlet–triplet gap is only 0.349 kcal/mol away from the target 19.856 kcal/mol value obtained with DEA-EOMCC($4p$-$2h$).

From Table 4.7 and Figure 4.10 we can see that the DIP-EOMCC($P|_{\tau=0}$) calculations,

where the $R_{\mu,4h\text{-}2p}$ components are completely neglected, produce very large, 26–30 millihartree, errors for both the states relative to the parent DIP-EOMCC($4h$-$2p$) method. As soon as we propagate a little in imaginary time *i.e.*, after only 10000 QMC iterations the DIP-EOMCC(P) calculated energies for the $\mathrm{X}\,^3\mathrm{A}_2'$ and $\mathrm{B}\,^1\mathrm{A}_1$ states are 4.259 and 9.124 millihartree away from their target values. This is a significant improvement, especially considering the fact that the P spaces at this time step contain only 7.3% of $4h$-$2p$ determinants for the $\mathrm{X}\,^3\mathrm{A}_2'$ state and 7.7% of $4h$-$2p$ determinants in case of the $\mathrm{B}\,^1\mathrm{A}_1$ state. The $\mathrm{B}\,^1\mathrm{A}_1 - \mathrm{X}\,^3\mathrm{A}_2'$ gap at this point is $\sim$3 kcal/mol away from its DIP-EOMCC($4h$-$2p$) predicted target of 19.907 kcal/mol. After 30000 QMC iterations, where the P spaces for the singlet and the triplet states contain about 30% and 16% of $4h$-$2p$ determinants respectively. At this point the error in the singlet–triplet gap is less than 1 kcal/mol and the underlying electronic states are only 1–3 millihartree in error.

Table 4.7 Convergence of the DEA-EOMCC(P) and DIP-EOMCC(P) energies [abbreviated as DEA(P) and DIP(P), respectively] of the X ^{3}A$_2'$ and B ^{1}A$_1$ states of TMM, as described by the 6-31G(d) basis set of Ref. [178], and of the corresponding adiabatic singlet–triplet (S–T) gaps toward their parent DEA-EOMCC(4p-2h) and DIP-EOMCC(4h-2p) values. The geometries of the X ^{3}A$_2'$ and B ^{1}A$_1$ states, optimized using the SF-DFT/6-31G(d) calculations, were taken from Ref. [231]. The P spaces used in the DEA-EOMCC(P) calculations were defined as all 2p and 3p-1h determinants and subsets of 4p-2h determinants extracted from the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. The P spaces used in the DIP-EOMCC(P) calculations were defined as all 2h and 3h-1p determinants and subsets of 4h-2p determinants extracted from the i-CISDTQ-MC propagations with $\delta\tau = 0.0001$ a.u. In all the post-SCF calculations, the lowest core orbitals of the carbon atoms were kept frozen and the spherical components of the carbon d orbitals were employed throughout.

	DEA-EOMCC(P)					DIP-EOMCC(P)				
MC	X ^{3}A$_2'$		B ^{1}A$_1$		S–T Gap	X ^{3}A$_2'$		B ^{1}A$_1$		S–T Gap
Iters.	DEA(P)[a]	%(4p-2h)[b]	DEA(P)[a]	%(4p-2h)[b]	ΔE[c]	DIP(P)[d]	%(4h-2p)[e]	DIP(P)[d]	%(4h-2p)[e]	ΔE[f]
0	19.933[g]	0.0	27.755[g]	0.0	4.909[g]	26.513[h]	0.0	29.721[h]	0.0	2.013[h]
4000	16.852	0.4	19.554	0.3	1.695	8.053	3.8	16.608	2.1	5.368
10000	10.967	1.4	12.653	1.4	1.058	4.259	7.3	9.124	7.7	3.053
20000	6.092	3.6	6.648	4.1	0.349	1.747	13.9	3.720	18.5	1.238
30000	3.274	6.9	3.738	8.4	0.291	2.649	15.9	1.080	30.4	−0.985
40000	1.583	11.8	1.901	15.4	0.200	1.526	22.3	0.229	43.0	−0.814
50000	0.727	18.3	0.843	26.0	0.073	1.110	28.8	0.044	54.0	−0.669
80000	0.048	53.9	0.092	74.3	0.028	0.007	54.7	0.001	61.9	−0.004
∞	−155.399202[i]		−155.367559[i]		19.856[j]	−155.399528[k]		−155.367804[k]		19.907[l]

[a] Unless otherwise stated, all energies are reported as errors relative to DEA-EOMCC(4p-2h) in millihartree.
[b] The %(4p-2h) values are the percentages of 4p-2h determinants captured during the i-CISDTQ-MC propagations [the $S_z = 1$ 4p-2h determinants of the B$_1$ symmetry in case of the X ^{3}A$_2'$ state and the $S_z = 0$ 4p-2h determinants of the A$_1$ symmetry in case of the B ^{1}A$_1$ state].
[c] Unless otherwise specified, the singlet–triplet gaps ΔE = B ^{1}A$_1$ − X ^{3}A$_2'$ are reported as errors relative to DEA-EOMCC(4p-2h) in kcal/mol.
[d] Unless otherwise stated, all energies are reported as errors relative to DIP-EOMCC(4h-2p) in millihartree.
[e] The %(4h-2p) values are the percentages of 4h-2p determinants captured during the i-CISDTQ-MC propagations [the $S_z = 1$ 4h-2p determinants of the B$_1$ symmetry in case of the X ^{3}A$_2'$ state and the $S_z = 0$ 4h-2p determinants of the A$_1$ symmetry in case of the B ^{1}A$_1$ state].
[f] Unless otherwise specified, the singlet–triplet gaps ΔE = B ^{1}A$_1$ − X ^{3}A$_2'$ are reported as errors relative to DIP-EOMCC(4h-2p) in kcal/mol.
[g] Equivalent to DEA-EOMCC(3p-1h).
[h] Equivalent to DIP-EOMCC(3h-1p).
[i] Total DEA-EOMCC(3p-1h) energy in hartree.
[j] The DEA-EOMCC(4p-2h) singlet–triplet gap in kcal/mol.
[k] Total DIP-EOMCC(3h-1p) energy in hartree.
[l] The DIP-EOMCC(4h-2p) singlet–triplet gap in kcal/mol.

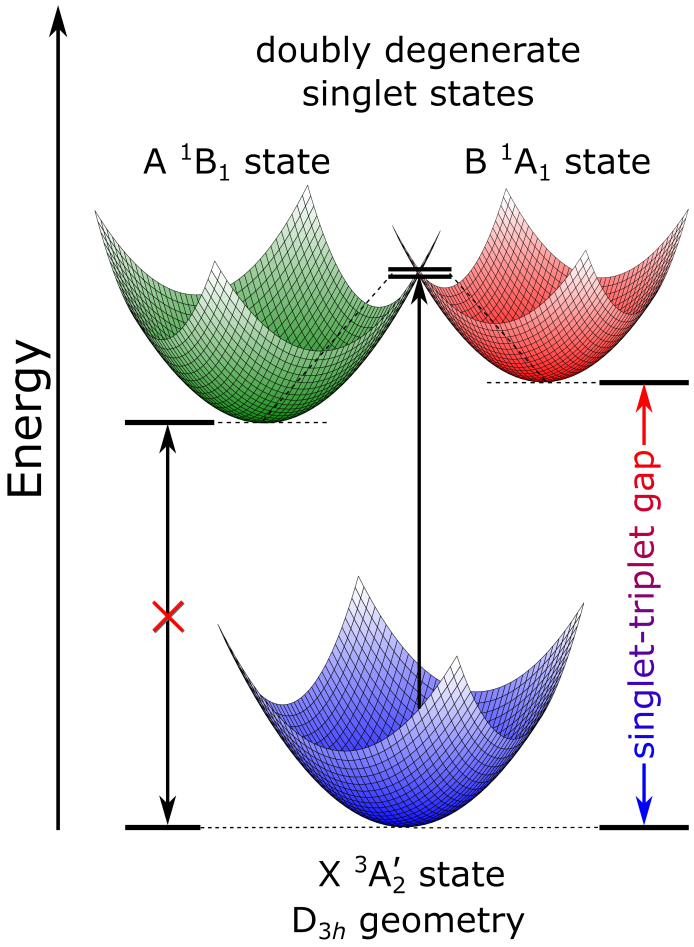

Figure 4.8 Jahn-Teller distortion in the trimethylenemethane molecule. At the geometry of the D_{3h}-symmetric triplet ground state (shown in blue), trimethylenemethane has a doubly degenerate singlet excited state. Due to Jahn–Teller distortion these states split into an open-shell singlet state A $^1B_1$ (shown in green) and a multi-configurational singlet state B $^1A_1$ (shown in red). Although the A $^1B_1$ state becomes the first excited state, it is not observed experimentally due to unfavorable Franck–Condon factors. Thus, we calculate the singlet–triplet gap between the ground triplet and the second excited singlet state B $^1A_1$.

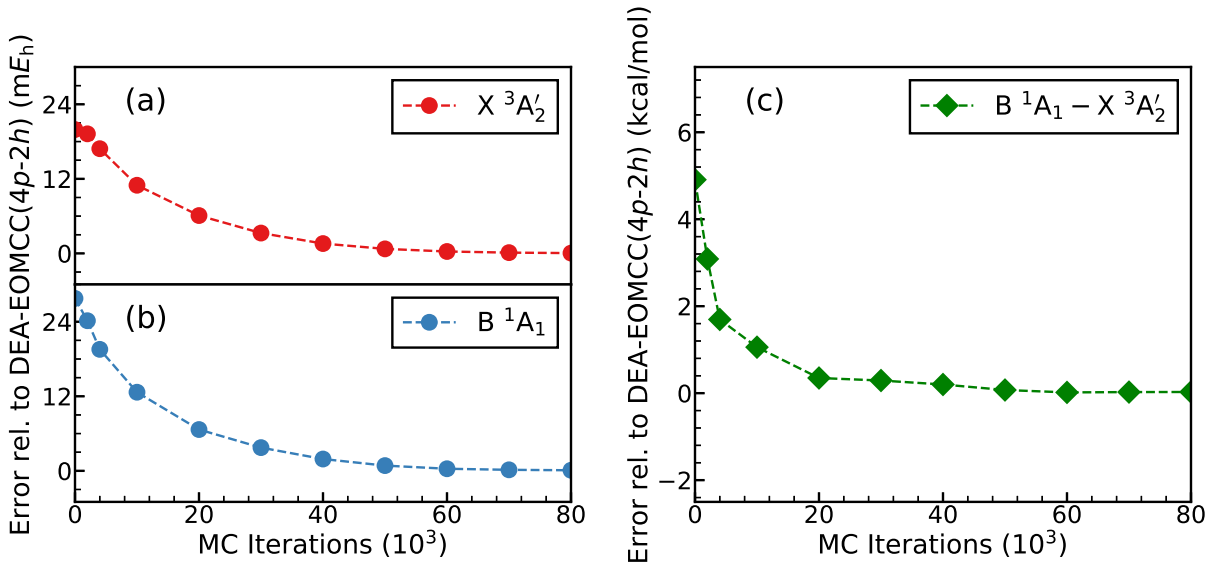

Figure 4.9 Convergence of DEA-EOMCC(P) energies of (a) X $^3A_2'$ and (b) B $^1A_1$ states of TMM, as described by the 6-31G(d) basis set, and (c) of the corresponding adiabatic singlet–triplet gap towards their parent DEA-EOMCC(4p-2h) values.

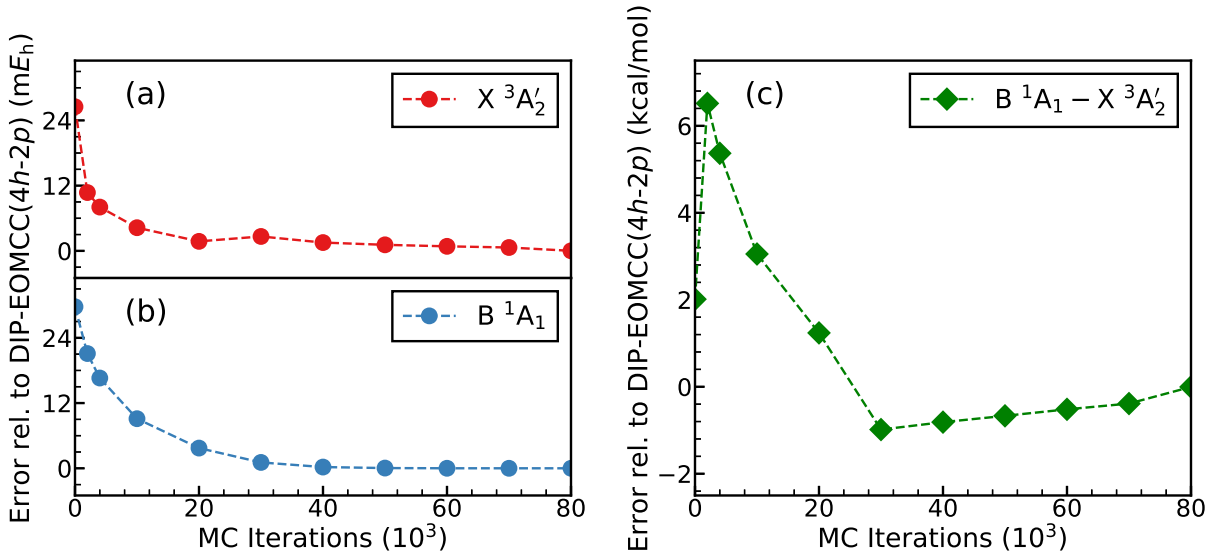

Figure 4.10 Convergence of DIP-EOMCC(P) energies of (a) X $^3A_2'$ and (b) B $^1A_1$ states of TMM, as described by the 6-31G(d) basis set, and (c) of the corresponding adiabatic singlet–triplet gap towards their parent DIP-EOMCC(4h-2p) values.

CHAPTER 5

CONCLUDING REMARKS AND FUTURE OUTLOOK

In this dissertation, we have discussed some of the recent advances in the CC and EOMCC theories to which the author of this thesis has had the opportunity to contribute during his doctoral work in Professor Piotr Piecuch's group. In particular, we have explored the semi-stochastic, CIQMC-driven, CC(P;Q) framework and and its extension to ground and excited states of open-shell systems. We have also discussed the semi-stochastic formulation of the particle nonconserving EA, IP, DEA, and DIP EOMCC frameworks, again taking advantage of CIQMC.

In the first part of this dissertation, we have discussed the CC theory as one of the best ways of approaching the many-electron correlation problem in molecular systems in a computationally tractable manner. After highlighting the advantages of the CC theory and some of its key challenges, and discussing its extensions to excited and open-shell states via the EOMCC formalism, we have focused on the CC(P;Q) methodology, where the flexibility in defining the P and Q excitation spaces allows one to obtain highly accurate energetics equivalent or very close to those obtained with the high-level CCSDT, CCSDTQ, *etc.* methods and their EOMCC extensions at small fractions of the computational costs, even when higher–than–two-body components of the cluster and EOM excitation operators become large, nonperturbative, and strongly coupled to their lower-rank components. We have also discussed the particle nonconserving EOMCC formalisms of the EA, IP, DEA, and DIP EOMCC types that offer an elegant and orthogonally spin-adapted approach to open-shell species within the SR framework. This includes problems involving electronic excitation spectra of radicals and singlet–triplet gaps in biradicals. We have also briefly reviewed the stochastic FCIQMC wave function propagation and sampling approach and its truncated CIQMC counterparts.

In the second part of this dissertation, we have focused on the semi-stochastic CC(P;Q) methodology that combines the flexible deterministic CC(P;Q) framework with the stochas-

tic CIQMC wave function propagations to automatically identify the P and Q spaces needed in the CC(P;Q) calculations without any reference to the previously exploited user- and system-dependent active-orbital concepts. By examining the excitation spectra of CH^+, CH, and CNC, we have demonstrated the ability of the semi-stochastic CC(P;Q) methodology to converge ground and excited states, including challenging non-singlet excited states of open-shell systems, out of the early stages of CIQMC propagations. Many of the excited states examined in this dissertation are dominated by two-electron transitions, making them difficult to describe by the majority of the existing quantum chemistry approaches. We have also applied the semi-stochastic CC(P;Q) methodology to converge the CCSDT energetics of the lowest singlet and triplet states of several biradical species, including methylene, $(HFH)^-$, cyclobutadiene, cyclopentadienyl cation, and trimethylenemethane, and the corresponding singlet–triplet gaps, which is another challenging problem for many quantum chemistry methods because one has to balance the predominantly weakly correlated triplet states with the multiconfigurational, often strongly correlated, singlet states.

In the third part of this dissertation, we have extended the semi-stochastic ideas to the particle nonconserving EOMCC formalisms of the EA, IP, DEA, and DIP types. We have demonstrated that by combining CIQMC wave function propagations with the deterministic EA- and IP-EOMCC frameworks, one can converge the high-level EA-EOMCC(3p-2h) and IP-EOMCC(3h-2p) energetics at small fractions of the computational costs out of the early stages of CIQMC propagations, even in the presence of strong 3p-2h and 3h-2p correlations. We have done this by studying the C_2N, CNC, N_3, and NCO radicals, where we have calculated the low-lying doublet states and the associated adiabatic excitation energies. We have then extended the semi-stochastic CIQMC-driven ideas to the DEA-EOMCC and DIP-EOMCC formalisms, which has allowed us to converge the high-level DEA-EOMCC(4p-2h) and DIP-EOMCC(4h-2p) energetics in an automated and efficient manner using small fractions of 4p-2h and 4h-2p amplitudes, again identified in the early stages of the underlying CIQMC runs. To illustrate the efficiency of our semi-stochastic DEA-EOMCC(4p-2h)

and DIP-EOMCC($4h$-$2p$) approaches, we have investigated the ground and three low-lying singlet excited states of methylene, along with the corresponding singlet–triplet gaps, and the trimethylenemethane biradical, where we have determined the ground triplet and the low-lying and multiconfigurational singlet states and the corresponding singlet–triplet gap.

While this dissertation has explored the semi-stochastic CC(P;Q) approach aimed at CCSDT/EOMCCSDT and the EA/IP/DEA/DIP EOMCC methodologies aimed at a highly accurate description of $3p$-$2h$/$3h$-$2p$/$4p$-$2h$/$4h$-$2p$ correlations, showing a lot of promise, there are still several areas that need to be examined. For example, recent work on the active space [84] and QMC-driven [101] CC(P;Q) frameworks targeting the ground-state CCSDTQ energetics have shown encouraging results, and hence it would be useful to extend the semi-stochastic CC(P;Q) methodology investigated in this dissertation to target the EOMCCSDTQ energetics of excited states. The selected-CI-driven [120] and adaptive [121] CC(P;Q) approaches have shown promising results in converging CCSDT energies as well, so it would be interesting to extend them to target EOMCCSDT or even CCSDTQ/EOMCCSDTQ. In the case of the particle nonconserving EOMCC approaches, the CC(P;Q)-type moment corrections have not been implemented yet. It would, thus, be very beneficial to extend the noniterative CC(P;Q) corrections to the particle nonconserving EOMCC schemes. Based on the significant acceleration toward the desired high-level CC/EOMCC energetics these corrections offer in the particle conserving cases, one may anticipate that they will be very effective in the EA-EOMCC, IP-EOMCC, DEA-EOMCC, and DIP-EOMCC approaches as well. Combining the particle nonconserving EOMCC formalisms with the selected-CI-based or adaptive selections of higher-order correlations, examined in this work using CIQMC-driven ideas, is another interesting aspect worth investigation.

In analogy to other post-SCF *ab initio* quantum chemistry approaches, in the longer-term, one can also think of extending the applicability of the methods developed in this dissertation to larger and more complex molecular systems containing hundreds of electrons and dozens of non-hydrogen atoms using techniques such as the fragment molecular orbital

(FMO) approach [259–261], the effective fragment potential (EFP) embedding scheme [262] and its merger with FMO abbreviated as EFMO [263–265], the cluster-in-molecule framework [266–269], and, ultimately, the quantum mechanics/molecular mechanics [270–274] and polarizable continuum [275, 276] models, to name a few examples. By doing so, especially when combined with code parallelization across multiple multi-core nodes, at least some of our semi-stochastic CC(P;Q) and EOMCC algorithms developed in this dissertation can be made applicable to studies of molecular electronic excitation spectra and singlet–triplet gaps in solution and condensed phases. Clearly, additional molecular applications of the semi-stochastic CC(P;Q) and EA/IP/DEA/DIP EOMCC approaches investigated in this work would be very useful too.

REFERENCES

[1] M. S. Gordon and D. G. Truhlar, Theor. Chim. Acta **71**, 1 (1987).

[2] F. Coester, Nucl. Phys. **7**, 421 (1958).

[3] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).

[4] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).

[5] J. Čížek, Adv. Chem. Phys. **14**, 35 (1969).

[6] J. Čížek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (1971).

[7] J. Paldus, J. Čížek, and I. Shavitt, Phys. Rev. A **5**, 50 (1972).

[8] J. Paldus and X. Li, Adv. Chem. Phys. **110**, 1 (1999).

[9] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[10] K. A. Brueckner, Phys. Rev. **100**, 36 (1955).

[11] J. Goldstone, Proc. R. Soc. London A Math. Phys. Eng. Sci. **239**, 267 (1957).

[12] J. Hubbard, Proc. R. Soc. London A Math. Phys. Eng. Sci. **240**, 539 (1957).

[13] N. M. Hugenholtz, Physica **23**, 481 (1957).

[14] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, Phys. Rev. A **47**, 2738 (1993).

[15] K. Kowalski and P. Piecuch, Phys. Rev. A **61**, 052506 (2000).

[16] P. Piecuch and K. Kowalski, Int. J. Mol. Sci. **3**, 676 (2002).

[17] D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. **112**, 182 (2012).

[18] F. A. Evangelista, J. Chem. Phys. **149**, 030901 (2018).

[19] B. O. Roos, Adv. Chem. Phys. **69**, 399 (1987).

[20] M. W. Schmidt and M. S. Gordon, Annu. Rev. Phys. Chem. **49**, 233 (1998).

[21] P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, Chem. Rev. **112**, 108 (2012).

[22] D. Roca-Sanjuán, F. Aquilante, and R. Lindh, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 585 (2012).

[23] S. Chattopadhyay, R. K. Chaudhuri, U. S. Mahapatra, A. Ghosh, and S. S. Ray, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **6**, 266 (2016).

[24] K. Kowalski and P. Piecuch, Int. J. Quantum Chem. **80**, 757 (2000).

[25] P. Piecuch, R. Toboła, and J. Paldus, Chem. Phys. Lett. **210**, 243 (1993).

[26] P. Piecuch and J. Paldus, Phys. Rev. A **49**, 3479 (1994).

[27] K. Jankowski, J. Paldus, I. Grabowski, and K. Kowalski, J. Chem. Phys. **97**, 7600 (1992), **101**, 1759 (1994) [Erratum].

[28] K. Jankowski, J. Paldus, I. Grabowski, and K. Kowalski, J. Chem. Phys. **101**, 3085 (1994).

[29] G. D. Purvis, III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[30] J. M. Cullen and M. C. Zerner, J. Chem. Phys. **77**, 4088 (1982).

[31] G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer, III, J. Chem. Phys. **86**, 2881 (1987).

[32] P. Piecuch and J. Paldus, Int. J. Quantum Chem. **36**, 429 (1989).

[33] J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 7041 (1987), **89**, 3401 (1988) [Erratum].

[34] G. E. Scuseria and H. F. Schaefer, III, Chem. Phys. Lett. **152**, 382 (1988).

[35] M. R. Hoffman and H. F. Schaefer, III, Adv. Quant. Chem. **18**, 207 (1986).

[36] N. Oliphant and L. Adamowicz, J. Chem. Phys. **95**, 6645 (1991).

[37] S. A. Kucharski and R. J. Bartlett, Theor. Chim. Acta **80**, 387 (1991).

[38] P. Piecuch and L. Adamowicz, J. Chem. Phys. **100**, 5792 (1994).

[39] S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **97**, 4282 (1992).

[40] K. Emrich, Nucl. Phys. A **351**, 379 (1981).

[41] J. Geertsen, M. Rittby, and R. J. Bartlett, Chem. Phys. Lett. **164**, 57 (1989).

[42] D. C. Comeau and R. J. Bartlett, Chem. Phys. Lett. **207**, 414 (1993).

[43] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).

[44] P. Piecuch and R. J. Bartlett, Adv. Quantum Chem. **34**, 295 (1999).

[45] H. Monkhorst, Int. J. Quantum Chem. Symp. **11**, 421 (1977).

[46] E. Dalgaard and H. Monkhorst, Phys. Rev. A **28**, 1217 (1983).

[47] M. Takahashi and J. Paldus, J. Chem. Phys. **85**, 1486 (1986).

[48] H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).

[49] H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **93**, 3345 (1990).

[50] H. Nakatsuji, Chem. Phys. Lett. **59**, 362 (1978).

[51] K. Kowalski and P. Piecuch, J. Chem. Phys. **115**, 643 (2001).

[52] K. Kowalski and P. Piecuch, Chem. Phys. Lett. **347**, 237 (2001).

[53] S. A. Kucharski, M. Włoch, M. Musiał, and R. J. Bartlett, J. Chem. Phys. **115**, 8263 (2001).

[54] S. Hirata, J. Chem. Phys. **121**, 51 (2004).

[55] M. Kállay and J. Gauss, J. Chem. Phys. **121**, 9257 (2004).

[56] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).

[57] P. Piecuch and K. Kowalski, in *Computational Chemistry: Reviews of Current Trends*, Vol. 5, edited by J. Leszczyński (World Scientific, Singapore, 2000) pp. 1–104.

[58] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 18 (2000).

[59] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, Int. Rev. Phys. Chem. **21**, 527 (2002).

[60] P. Piecuch, K. Kowalski, I. S. O. Pimienta, P.-D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kuś, and M. Musiał, Theor. Chem. Acc. **112**, 349 (2004).

[61] K. Kowalski and P. Piecuch, J. Chem. Phys. **120**, 1715 (2004).

[62] M. Włoch, J. R. Gour, K. Kowalski, and P. Piecuch, J. Chem. Phys. **122**, 214107 (2005).

[63] P. Piecuch and M. Włoch, J. Chem. Phys. **123**, 224105 (2005).

[64] P. Piecuch, M. Włoch, J. R. Gour, and A. Kinal, Chem. Phys. Lett. **418**, 467 (2006).

[65] M. Włoch, M. D. Lodriguito, P. Piecuch, and J. R. Gour, Mol. Phys. **104**, 2149 (2006), **104**, 2991 (2006) [Erratum].

[66] M. Włoch, J. R. Gour, and P. Piecuch, J. Phys. Chem. A **111**, 11359 (2007).

[67] P. Piecuch, M. Włoch, M. Lodriguito, and J. R. Gour, in *Recent Advances in the Theory of Chemical and Physical Systems*, Progress in Theoretical Chemistry and Physics, Vol. 15, edited by S. Wilson, J.-P. Julien, J. Maruani, E. Brändas, and G. Delgado-Barrio (Springer, Dordrecht, 2006) pp. 45–106.

[68] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. **109**, 3268 (2009).

[69] G. Fradelos, J. J. Lutz, T. A. Wesołowski, P. Piecuch, and M. Włoch, J. Chem. Theory Comput. **7**, 1647 (2011).

[70] J. Shen and P. Piecuch, Chem. Phys. **401**, 180 (2012).

[71] N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).

[72] N. Oliphant and L. Adamowicz, J. Chem. Phys. **96**, 3739 (1992).

[73] P. Piecuch, N. Oliphant, and L. Adamowicz, J. Chem. Phys. **99**, 1875 (1993).

[74] P. Piecuch and L. Adamowicz, J. Chem. Phys. **102**, 898 (1995).

[75] L. Adamowicz, P. Piecuch, and K. B. Ghose, Mol. Phys. **94**, 225 (1998).

[76] P. Piecuch, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **110**, 6103 (1999).

[77] K. Kowalski and P. Piecuch, J. Chem. Phys. **113**, 8490 (2000).

[78] J. R. Gour, P. Piecuch, and M. Włoch, J. Chem. Phys. **123**, 134113 (2005).

[79] P. Piecuch, Mol. Phys. **108**, 2987 (2010).

[80] J. Shen and P. Piecuch, J. Chem. Phys. **138**, 194102 (2013).

[81] K. Kowalski and P. Piecuch, J. Chem. Phys. **122**, 074107 (2005).

[82] P.-D. Fan, K. Kowalski, and P. Piecuch, Mol. Phys. **103**, 2191 (2005).

[83] M. Horoi, J. R. Gour, M. Włoch, M. D. Lodriguito, B. A. Brown, and P. Piecuch, Phys. Rev. Lett. **98**, 112501 (2007).

[84] N. P. Bauman, J. Shen, and P. Piecuch, Mol. Phys. **115**, 2860 (2017).

[85] J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. **233**, 81 (1995).

[86] J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. **258**, 581 (1996).

[87] J. Shen and P. Piecuch, J. Chem. Phys. **136**, 144104 (2012).

[88] J. Shen and P. Piecuch, J. Chem. Theory Comput. **8**, 4968 (2012).

[89] S. H. Yuwono, I. Magoulas, J. Shen, and P. Piecuch, Mol. Phys. **117**, 1486 (2019).

[90] I. Magoulas, N. P. Bauman, J. Shen, and P. Piecuch, J. Phys. Chem. A **122**, 1350 (2018).

[91] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[92] D. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[93] K. Ghanem, A. Y. Lozovoi, and A. Alavi, J. Chem. Phys. **151**, 224108 (2019).

[94] K. Ghanem, K. Guther, and A. Alavi, J. Chem. Phys. **153**, 224115 (2020).

[95] A. J. W. Thom, Phys. Rev. Lett. **105**, 263004 (2010).

[96] R. S. T. Franklin, J. S. Spencer, A. Zoccante, and A. J. W. Thom, J. Chem. Phys. **144**, 044111 (2016).

[97] J. S. Spencer and A. J. W. Thom, J. Chem. Phys. **144**, 084108 (2016).

[98] C. J. C. Scott and A. J. W. Thom, J. Chem. Phys. **147**, 124105 (2017).

[99] J. E. Deustua, J. Shen, and P. Piecuch, Phys. Rev. Lett. **119**, 223003 (2017).

[100] S. H. Yuwono, A. Chakraborty, J. E. Deustua, J. Shen, and P. Piecuch, Mol. Phys. **118**, e1817592 (2020).

[101] J. E. Deustua, J. Shen, and P. Piecuch, J. Chem. Phys., **154**, 124103 (2021).

[102] A. Chakraborty, S. H. Yuwono, J. E. Deustua, J. Shen, and P. Piecuch, J. Chem. Phys.

157, 134101 (2022).

[103] J. Whitten and M. Hackmeyer, J. Chem. Phys. **51**, 5584 (1969).

[104] C. Bender and E. Davidson, Phys. Rev. **183**, 23 (1969).

[105] B. Huron, J. P. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).

[106] R. Buenker and S. Peyerimhoff, Theor. Chim. Acta. **35**, 33 (1974).

[107] J. Schriber and F. Evangelista, J. Chem. Phys. **144**, 161106 (2016).

[108] J. Schriber and F. Evangelista, J. Chem. Theory Comput. **13**, 5354 (2017).

[109] N. M. Tubman, J. Lee, T. Takeshita, M. Head-Gordon, and K. Whaley, J. Chem. Phys. **145**, 044112 (2016).

[110] N. M. Tubman, C. Freeman, D. Levine, D. Hait, M. Head-Gordon, and K. Whaley, J. Chem. Theory Comput. **16**, 2139 (2020).

[111] W. Liu and M. Hoffmann, J. Chem. Theory Comput. **12**, 1169 (2016), **12**, 3000 (2016) [Erratum].

[112] N. Zhang, W. Liu, and M. Hoffmann, J. Chem. Theory Comput. **16**, 2296 (2020).

[113] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, J. Chem. Theory Comput. **12**, 3674 (2016).

[114] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, J. Chem. Theory Comput. **13**, 1595 (2017).

[115] J. Li, M. Otten, A. A. Holmes, S. Sharma, and C. J. Umrigar, J. Chem. Phys. **149**, 214110 (2018).

[116] Y. Garniron, A. Scemama, P.-F. Loos, and M. Caffarel, J. Chem. Phys. **147**, 034101 (2017).

[117] Y. Garniron, T. Applencourt, K. Gasperich, A. Benali, A. Ferte, J. Paquier, B. Pradines, R. Assaraf, P. Reinhardt, J. Toulouse, P. Barbaresco, N. Renon, G. David, J.-P. Malrieu, M. Veril, M. Caffarel, P.-F. Loos, E. Giner, and A. Scemama, J. Chem. Theory Comput. **15**, 3591 (2019).

[118] P.-F. Loos, Y. Damour, and A. Scemama, J. Chem. Phys. **153**, 176101 (2020).

[119] J. J. Eriksen, T. A. Anderson, J. E. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann,

S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar, and J. Gauss, J. Phys. Chem. Lett. **11**, 8922 (2020).

[120] K. Gururangan, J. E. Deustua, J. Shen, and P. Piecuch, J. Chem. Phys. **155**, 174114 (2021).

[121] K. Gururangan and P. Piecuch, J. Chem. Phys. **159**, 084108 (2023).

[122] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **44**, 335 (1949).

[123] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[124] W. K. Hastings, Biometrika **57**, 97 (1970).

[125] W. L. McMillan, Phys. Rev. **138**, A442 (1965).

[126] D. Schiff and L. Verlet, Phys. Rev. **160**, 208 (1967).

[127] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[128] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr., J. Chem. Phys. **77**, 5593 (1982).

[129] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[130] D. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **16**, 3081 (1977).

[131] D. M. Ceperley, Rev. Mineral. Geochem. **71**, 129 (2010).

[132] J. Toulouse, R. Assaraf, and C. J. Umrigar, Adv. Quantum Chem. **73**, 285 (2016).

[133] J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975).

[134] D. J. Klein and H. M. Pickett, J. Chem. Phys. **64**, 4811 (1976).

[135] J. B. Anderson, J. Chem. Phys. **65**, 4121 (1976).

[136] J. B. Anderson, Int. J. Quantum Chem. **15**, 109 (1979).

[137] J. E. Deustua, S. H. Yuwono, J. Shen, and P. Piecuch, J. Chem. Phys. **150**, 111101 (2019).

[138] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **102**, 3629 (1995).

[139] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **102**, 6735 (1995).

[140] J. R. Gour, P. Piecuch, and M. Włoch, Int. J. Quantum Chem. **106**, 2854 (2006).

[141] J. R. Gour and P. Piecuch, J. Chem. Phys. **125**, 234107 (2006).

[142] Y. Ohtsuka, P. Piecuch, J. R. Gour, M. Ehara, and H. Nakatsuji, J. Chem. Phys. **126**, 164111 (2007).

[143] M. Nooijen and J. G. Snijders, Int. J. Quantum Chem. Symp. **26**, 55 (1992).

[144] M. Nooijen and J. G. Snijders, Int. J. Quantum Chem. **48**, 15 (1993).

[145] J. F. Stanton and J. Gauss, J. Chem. Phys. **101**, 8938 (1994).

[146] R. J. Bartlett and J. F. Stanton, in *Reviews in Computational Chemistry*, Vol. 5, edited by K. B. Lipkowitz and D. B. Boyd (VCH Publishers, New York, 1994) pp. 65–169.

[147] M. Nooijen and R. J. Bartlett, J. Chem. Phys. **106**, 6441 (1997).

[148] M. Nooijen, Int. J. Mol. Sci. **3**, 656 (2002).

[149] M. Musiał, A. Perera, and R. J. Bartlett, J. Chem. Phys. **134**, 114108 (2011).

[150] M. Musiał, S. A. Kucharski, and R. J. Bartlett, J. Chem. Theory Comput. **7**, 3088 (2011).

[151] T. Kuś and A. I. Krylov, J. Chem. Phys. **135**, 084109 (2011).

[152] T. Kuś and A. I. Krylov, J. Chem. Phys. **136**, 244109 (2012).

[153] J. Shen and P. Piecuch, Mol. Phys. **112**, 868 (2014).

[154] A. O. Ajala, J. Shen, and P. Piecuch, J. Phys. Chem. A **121**, 3469 (2017).

[155] S. J. Stoneburner, J. Shen, A. O. Ajala, P. Piecuch, D. G. Truhlar, and L. Gagliardi, J. Chem. Phys. **147**, 164120 (2017).

[156] J. Shen and P. Piecuch, Mol. Phys. **119**, e1966534 (2021).

[157] S. Gulania, E. F. Kjønstad, J. F. Stanton, H. Koch, and A. I. Krylov, J. Chem. Phys. **154**, 114115 (2021).

[158] K. E. Schmidt and M. H. Kalos, in *Applications of the Monte Carlo method in Statistical Physics*, Vol. 36 (Springer-Verlag, Berlin, 1984).

[159] N. S. Blunt, S. D. Smart, G. H. Booth, and A. Alavi, J. Chem. Phys. **143**, 134117 (2015).

[160] N. S. Blunt, G. H. Booth, and A. Alavi, J. Chem. Phys. **146**, 244105 (2017).

[161] S. Hirata, J. Chem. Phys. **121**, 51 (2004).

[162] M. Ehara, J. R. Gour, and P. Piecuch, Mol. Phys. **107**, 871 (2009).

[163] J. A. Hansen, P. Piecuch, J. J. Lutz, and J. R. Gour, Phys. Scr. **84**, 028110 (2011).

[164] J. Olsen, A. M. S. de Merás, H. Jensen, and P. Jørgensen, Chem. Phys. Lett. **154**, 380 (1989).

[165] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[166] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6769 (1992).

[167] T. H. Dunning, Jr., J. Chem. Phys. **53**, 2823 (1970).

[168] T. H. Dunning, Jr. and P. J. Hay, in *Methods of Electronic Structure Theory*, Vol. 2, edited by H. F. Schaefer, III (Plenum, New York, 1977) pp. 1–28.

[169] J. E. Deustua, I. Magoulas, J. Shen, and P. Piecuch, J. Chem. Phys. **149**, 151101 (2018).

[170] J. S. Spencer, N. S. Blunt, W. A. Vigor, F. D. Malone, W. M. C. Foulkes, J. J. Shepherd, and A. J. W. Thom, J. Open Res. Softw. **3**, e9 (2015).

[171] J. S. Spencer, N. S. Blunt, S. Choi, J. Etrych, M.-A. Filip, W. M. C. Foulkes, R. S. T. Franklin, W. J. Handley, F. D. Malone, V. A. Neufeld, R. Di Remigio, T. W. Rogers, C. J. C. Scott, J. J. Shepherd, W. A. Vigor, J. Weston, R. Xu, and A. J. W. Thom, J. Chem. Theory Comput. **15**, 1728 (2019).

[172] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).

[173] G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B.

Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. G. Vallejo, B. Westheimer, M. Włoch, P. Xu, F. Zahariev, and M. S. Gordon, J. Chem. Phys. **152**, 154102 (2020).

[174] F. Zahariev, P. Xu, B. M. Westheimer, S. Webb, J. Galvez Vallejo, A. Tiwari, V. Sundriyal, M. Sosonkina, J. Shen, G. Schoendorff, M. Schlinsog, T. Sattasathuchana, K. Ruedenberg, L. B. Roskop, A. P. Rendell, D. Poole, P. Piecuch, B. Q. Pham, V. Mironov, J. Mato, S. Leonard, S. S. Leang, J. Ivanic, J. Hayes, T. Harville, K. Gururangan, E. Guidez, I. S. Gerasimov, C. Friedl, K. N. Ferreras, G. Elliott, D. Datta, D. D. A. Cruz, L. Carrington, C. Bertoni, G. M. J. Barca, M. Alkan, and M. S. Gordon, J. Chem. Theory Comput. **19**, 7031 (2023).

[175] M. Zachwieja, J. Mol. Spectrosc. **170**, 285 (1995).

[176] R. Kępa, A. Para, M. Rytel, and M. Zachwieja, J. Mol. Spectrosc. **178**, 189 (1996).

[177] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure: Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).

[178] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[179] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).

[180] T. H. Dunning, Jr., J. Chem. Phys. **55**, 716 (1971).

[181] C. D. Sherrill, M. L. Leininger, T. J. V. Huis, and H. F. Schaefer, III, J. Chem. Phys. **108**, 1040 (1998).

[182] C. W. Bauschlicher and P. R. Taylor, J. Chem. Phys. **85**, 6510 (1986).

[183] C. W. Bauschlicher, S. R. Langhoff, and P. R. Taylor, J. Chem. Phys. **87**, 387 (1987).

[184] A. D. McLean, P. R. Buenker, R. M. Escribano, and P. Jensen, J. Chem. Phys. **87**, 2166 (1987).

[185] D. C. Comeau, I. Shavitt, P. Jensen, and P. R. Bunker, J. Chem. Phys. **90**, 6491 (1989).

[186] D. B. Knowles, J. Ramon, A. Collado, G. Hirsch, and R. J. Buekner, J. Chem. Phys. **92**, 585 (1990).

[187] X. Li, P. Piecuch, and J. Paldus, Chem. Phys. Lett. **224**, 267 (1994).

[188] P. Piecuch, X. Li, and J. Paldus, Chem. Phys. Lett. **230**, 377 (1994).

[189] A. Balková and R. J. Bartlett, J. Chem. Phys. **102**, 7116 (1995).

[190] X. Li and J. Paldus, J. Chem. Phys. **124**, 174101 (2006).

[191] J. D. Watts and R. J. Bartlett, J. Chem. Phys. **93**, 6104 (1990).

[192] J. R. Gour, P. Piecuch, and M. Włoch, Mol. Phys. **108**, 2633 (2010).

[193] P. Jensen and P. R. Bunker, J. Chem. Phys. **89**, 1327 (1988).

[194] E. R. Davidson, D. Feller, and P. Phillips, Chem. Phys. Lett. **76**, 416 (1980).

[195] N. C. Handy, Y. Yamaguchi, and H. F. Schaefer, III, J. Chem. Phys. **84**, 4481 (1986).

[196] J. R. Hart, A. K. Rappe, S. M. Gorun, and T. H. Upton, J. Phys. Chem. **96**, 6264 (1992).

[197] P. Piecuch, J. R. Gour, and M. Włoch, Int. J. Quantum Chem. **108**, 2128 (2008).

[198] J. Shen, Z. Kou, E. Xu, and S. Li, J. Chem. Phys. **134**, 044134 (2011).

[199] H. Schurkus, D.-T. Chen, H.-P. Cheng, G. Chan, and J. Stanton, J. Chem. Phys. **152**, 234115 (2020).

[200] P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993).

[201] P. G. Szalay and R. J. Bartlett, J. Chem. Phys. **103**, 3600 (1995).

[202] M. Eckert-Maksić, M. Vazdar, M. Barbatti, H. Lischka, and Z. B. Maksić, J. Chem. Phys. **125**, 064310 (2006).

[203] T. Saito, S. Nishihara, S. Yamanaka, Y. Kitagawa, T. Kawakami, S. Yamada, H. Isobe, M. Okumura, and K. Yamaguchi, Theor. Chem. Acc. **130**, 749 (2011).

[204] A. Balková and R. J. Bartlett, J. Chem. Phys. **101**, 8972 (1994).

[205] S. V. Levchenko and A. I. Krylov, J. Chem. Phys. **120**, 175 (2004).

[206] P. M. Zimmerman, J. Phys. Chem. A **121**, 4712 (2017).

[207] J.-N. Boyn and D. A. Mazziotti, J. Chem. Phys. **154**, 134103 (2021).

[208] E. Monino, M. Boggio-Pasqua, A. Scemama, D. Jacquemin, and P.-F. Loos, J. Phys. Chem. A **126**, 4664 (2022).

[209] C. A. Coulson, J. Chim. Phys. Phys.-Chim. Biol. **45**, 243 (1948).

[210] H. C. Longuet-Higgins, J. Chem. Phys. **18**, 265 (1950).

[211] P. Dowd, J. Am. Chem. Soc. **88**, 2587 (1966).

[212] R. J. Baseman, D. W. Pratt, M. Chow, and P. Dowd, J. Am. Chem. Soc. **98**, 5726 (1976).

[213] W. T. Borden and E. R. Davidson, Acc. Chem. Res. **14**, 69 (1981).

[214] D. R. Yarkony and H. F. Schaefer, III, J. Am. Chem. Soc. **96**, 3754 (1974).

[215] W. J. Hehre, L. Salem, and M. R. Willcott, J. Am. Chem. Soc. **96**, 4328 (1974).

[216] J. H. Davis and W. A. Goddard, III, J. Am. Chem. Soc. **99**, 4242 (1977).

[217] D. M. Hood, R. M. Pitzer, and H. F. Schaefer, III, J. Am. Chem. Soc. **100**, 2227 (1978).

[218] D. M. Hood, H. F. Schaefer, III, and R. M. Pitzer, J. Am. Chem. Soc. **100**, 8009 (1978).

[219] D. A. Dixon, R. Foster, T. A. Halgren, and W. N. Lipscomb, J. Am. Chem. Soc. **100**, 1359 (1978).

[220] D. Feller, W. T. Borden, and E. R. Davidson, J. Chem. Phys. **74**, 2256 (1981).

[221] S. B. Auster, R. M. Pitzer, and M. S. Platz, J. Am. Chem. Soc. **104**, 3812 (1982).

[222] W. T. Borden, E. R. Davidson, and D. Feller, Tetrahedron **38**, 737 (1982).

[223] P. M. Lahti, A. R. Rossi, and J. A. Berson, J. Am. Chem. Soc. **107**, 2273 (1985).

[224] A. Skanche, L. J. Schaad, and B. A. Hess, Jr., J. Am. Chem. Soc. **110**, 5315 (1988).

[225] S. Olivella and J. Salvador, Int. J. Quantum Chem. **37**, 713 (1990).

[226] T. P. Radhakrishnan, Tetrahedron Lett. **32**, 4601 (1991).

[227] A. S. Ichimura, N. Koga, and H. Iwamura, J. Phys. Org. Chem. **7**, 207 (1994).

[228] W. T. Borden, Mol. Cryst. Liq. Cryst. **232**, 195 (1993).

[229] C. J. Cramer and B. A. Smith, J. Phys. Chem. **100**, 9664 (1996).

[230] B. Ma and H. F. Schaefer, III, Chem. Phys. **207**, 31 (1996).

[231] L. V. Slipchenko and A. I. Krylov, J. Chem. Phys. **117**, 4694 (2002).

[232] L. V. Slipchenko and A. I. Krylov, J. Chem. Phys. **118**, 6874 (2003).

[233] L. V. Slipchenko and A. I. Krylov, J. Chem. Phys. **123**, 084107 (2005).

[234] J. Brabec and J. Pittner, J. Phys. A **110**, 11765 (2006).

[235] J. Shen, T. Fang, S. Li, and Y. Jiang, J. Phys. Chem. A **112**, 12518 (2008).

[236] X. Li and J. Paldus, J. Chem. Phys. **129**, 174101 (2008).

[237] A. Perera, R. W. Molt, Jr., V. F. Lotrich, and R. J. Bartlett, Theor. Chem. Acc. **133**, 1514 (2014).

[238] S. Sinha Ray, S. Manna, A. Ghosh, R. K. Chaudhuri, and S. Chattopadhyay, Int. J. Quantum Chem. **119**, e25776 (2019).

[239] S. Chattopadhyay, ACS Omega **6**, 1668 (2021).

[240] P. G. Wenthold, J. Hu, R. R. Squires, and W. C. Lineberger, J. Am. Chem. Soc. **118**, 475 (1996).

[241] P. G. Wenthold, J. Hu, R. R. Squires, and W. C. Lineberger, J. Am. Soc. Mass Spectrom. **10**, 800 (1999).

[242] A. Chakraborty, S. Basumallick, J. Shen, and P. Piecuch, "Extension of the Semi-Stochastic Equation-of-Motion Coupled-Cluster Formalism to the Single and Double Electron Attachment and Ionization Potential Schemes", in preparation for J. Phys. Chem. A.

[243] M. Musiał and R. J. Bartlett, J. Chem. Phys. **119**, 1901 (2003).

[244] M. Kamiya and S. Hirata, J. Chem. Phys. **125**, 074111 (2006).

[245] P.-D. Fan, M. Kamiya, and S. Hirata, J. Chem. Theor. Comput. **3**, 1036 (2007).

[246] M. Musiał, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **118**, 1128 (2003).

[247] M. Musiał and R. J. Bartlett, Chem. Phys. Lett. **384**, 210 (2004).

[248] Y. J. Bomble, J. C. Saeh, J. F. Stanton, P. G. Szalay, and M. Kállay, J. Chem. Phys. **122**, 154107 (2005).

[249] M. Wladyslawski and M. Nooijen, in *Low-Lying Potential Energy Surfaces*, ACS Symposium Series, Vol. 828, edited by M. R. Hoffmann and K. G. Dyall (American Chemical Society, Washington, D.C., 2002) pp. 65–92.

[250] K. W. Sattelmeyer, H. F. Schaefer, III, and J. F. Stanton, Chem. Phys. Lett. **378**, 42 (2003).

[251] M. Ehara, P. Piecuch, J. J. Lutz, and J. R. Gour, Chem. Phys. **399**, 94 (2012).

[252] N. P. Bauman, J. A. Hansen, M. Ehara, and P. Piecuch, J. Chem. Phys. **141**, 101102 (2014).

[253] N. P. Bauman, J. A. Hansen, and P. Piecuch, J. Chem. Phys. **145**, 084306 (2016).

[254] J. R. Reimets, J. Shen, M. Kianinia, C. Bradac, I. Aharonovich, M. J. Ford, and P. Piecuch, Phys. Rev. B **102**, 144105 (2020).

[255] W. Park, J. Shen, S. Lee, P. Piecuch, M. Filatov, and C. H. Choi, J. Phys. Chem. Lett. **12**, 9720 (2021).

[256] W. Park, J. Shen, S. Lee, P. Piecuch, T. Joo, M. Filatov(Gulak), and C. H. Choi, J. Phys. Chem. C **126**, 14976 (2022).

[257] W. Dobrautz, S. D. Smart, and A. Alavi, J. Chem. Phys. **151**, 094104 (2019).

[258] G. Herzberg, *Molecular Spectra and Molecular Structure 1: Spectra of Diatomic Molecules* (Van Nostrand, London, 1967).

[259] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).

[260] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **120**, 6832 (2004).

[261] D. G. Fedorov, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **7**, e1322 (2017).

[262] M. S. Gordon, L. Slipchenko, H. Li, and J. H. Jensen, in *Annual Reports in Computational Chemistry*, edited by D. C. Spellmeyer and R. Wheeler (Elsevier, 2007) pp. 177–193.

[263] C. Steinmann, D. G. Fedorov, and J. H. Jensen, J. Phys. Chem. A **114**, 8705 (2010).

[264] C. Steinmann, D. G. Fedorov, and J. H. Jensen, PLoS One **7**, e41117 (2012).

[265] C. Steinmann, D. G. Fedorov, and J. H. Jensen, PLoS One **8**, e60602 (2013).

[266] S. Li, J. Shen, W. Li, and Y. Jiang, J. Chem. Phys. **125**, 074109 (2006).

[267] W. Li, J. R. Gour, P. Piecuch, and S. Li, J. Chem. Phys. **131**, 114109 (2009).

[268] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).

[269] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 6721 (2010).

[270] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).

[271] Y. Zhang, T.-S. Lee, and W. Yang, J. Chem. Phys. **110**, 46 (1999).

[272] H. Lin and D. G. Truhlar, J. Phys. Chem. A **109**, 3991 (2005).

[273] G. Giudetti, I. Polyakov, B. L. Grigorenko, S. Faraji, A. V. Nemukhin, and A. I. Krylov, J. Chem. Theoy Comput. **18**, 5056 (2022).

[274] A. Zlobin, J. Belyaeva, and A. Golovin, J. Chem. Inf. Model. **63**, 546 (2023).

[275] J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).

[276] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. B **113**, 6378 (2009).