Accelerating convergence of equation-of-motion coupled-cluster computations using the semi-stochastic CC(*P*;*Q*) formalism

Stephen H. Yuwono,[1] Arnab Chakraborty,[1] J. Emiliano Deustua,[1] Jun Shen,[1] and Piotr Piecuch[1,2,a]

[1] *Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA*

[2] *Department of Physics & Astronomy, Michigan State University, East Lansing, MI 48824, USA*

ABSTRACT

The recently proposed approach to excited electronic states, in which the deterministic equation-of-motion coupled-cluster (EOMCC) framework is merged with stochastic configuration interaction Quantum Monte Carlo (CIQMC) computations [J.E. Deustua *et al.*, J. Chem. Phys. **150**, 111101 (2019)], is combined with the noniterative energy corrections derived from the CC(*P*;*Q*) formalism. By examining vertical excitations in CH^+ at the equilibrium and stretched geometries and adiabatic excitations in CH and CNC, we demonstrate that the resulting semi-stochastic CC(*P*;*Q*) methodology converges target high-level energetics, represented in this study by the EOMCC method with singles, doubles, and triples, in the early stages of CIQMC propagations.

[a] Corresponding author. Electronic mail: piecuch@chemistry.msu.edu.

1. Introduction

One of the greatest challenges in quantum chemistry is the development of practical and yet robust and systematically improvable treatments of many-electron correlation effects, which are needed to accurately determine ground- and excited-state molecular potential energy surfaces and property functions that emerge in studies of chemical reactivity, spectroscopy, and photochemistry. It is nowadays well established that size-extensive methods based on the exponential wave function ansatz of coupled-cluster (CC) theory [1–5], $|\Psi_0\rangle = \exp(T)|\Phi\rangle$, where $T = \sum_{n=1}^{N} T_n$ is the cluster operator, T_n is the n-body component of T, N is the number of correlated electrons, and $|\Phi\rangle$ is the reference determinant, and their various extensions to excited, open-shell, and multireference states (see Refs. [6–8] for selected reviews) are excellent candidates for addressing this challenge. This is particularly true for the equation-of-motion (EOM) CC [9–11] approach to excited states, pursued in this study, and its linear-response (LR) [12–18] and sym-metry-adapted-cluster (SAC) configuration interaction (CI) [19] counterparts, which adopt the following representation of excited-state wave functions: $|\Psi_\mu\rangle = R_\mu|\Psi_0\rangle = R_\mu \exp(T)|\Phi\rangle$, where $R_\mu = r_{\mu,0}\mathbf{1} + \sum_{n=1}^{N} R_{\mu,n}$ is the linear excitation operator generating $|\Psi_\mu\rangle$ from the CC ground state $|\Psi_0\rangle$, $R_{\mu,n}$ is the n-body component of R_μ, and $\mathbf{1}$ is the unit operator.

One of the key challenges in EOMCC, LRCC, and SAC-CI, which has propelled much of the development work dealing with these methodologies for about three decades, has been how to incorporate higher–than–two-body components of the cluster and EOM excitation operators, i.e., the T_n and $R_{\mu,n}$ components with $n > 2$, needed to achieve a quantitative description, without running into enormous, often prohibitive, computational costs of the higher-order schemes, such as the EOMCC method with singles, doubles, and triples (EOMCCSDT) [20–22], where T and R_μ are truncated at T_3 and $R_{\mu,3}$, respectively, which is the excited-state counterpart of CCSDT [23,24], or the EOMCC approach with singles, doubles, triples, and quadruples (EOMCCSDTQ) [25,26], where T and R_μ are truncated at T_4 and $R_{\mu,4}$, respectively, which is an excited-state extension of CCSDTQ [27–29]. Focusing on EOMCC, one can make the computations a lot more affordable and reduce the iterative $n_o^3 n_u^5$ steps of EOMCCSDT or $n_o^4 n_u^6$ steps of EOMCCSDTQ to

$n_o^2 n_u^4$ by ignoring the T_n and $R_{\mu,n}$ components with $n > 2$, as in the EOMCC method with singles and doubles (EOMCCSD) [11], which extends the ground-state CCSD approach [30,31] to excited states, but EOMCCSD fails to describe excited-state potentials along bond stretching coordinates and excited states dominated by two- and other many-electron transitions, producing errors that often exceed 1 eV (cf. Refs. [20,21,32–37] for examples), while not being fully quantitative for singly excited states, where errors on the order of 0.3–0.5 eV are not uncommon [38] (n_o and n_u are the numbers of correlated occupied and unoccupied orbitals, respectively). Many ways of incorporating higher–than–two-body components of the T and R_μ operators at the fraction of the computational costs of high-level EOMCC methods, such as EOMCCSDT, have been proposed, resulting in noniterative perturbative corrections to EOMCCSD or LRCCSD excitation energies [39–44] and their completely renormalized (CR) [32,35–38,45–47] and iterative [39–42,48,49] counterparts, but methods of these types face several new challenges. For example, the EOMCC approaches utilizing perturbation theory to identify the leading post-EOMCCSD contributions have difficulties with describing multireference excited states characterized by substantial electronic quasi-degeneracies, while the more robust CR-EOMCC corrections to EOMCCSD, which are capable of handling doubly excited states and excited-state potentials along bond stretching coordinates, have difficulties with balancing ground- and excited-state energies. The latter issue is, in part, related to decoupling the low-order T_n and $R_{\mu,n}$ components with $n \leq 2$ from their higher–than–two-body counterparts in all noniterative corrections to EOMCCSD. One can eliminate many of the above problems, while keeping computational costs at the relatively low levels, by turning to the active-space EOMCC approaches [50], such as EOMCCSDt [20,21,33,51] and EOMCCSDtq [51,52], where one uses active orbitals to select the dominant T_n and $R_{\mu,n}$ amplitudes with $n > 2$ within the parent EOMCCSDT, EOMCCSDTQ, and similar schemes, but the resulting methods, in analogy to multireference theories, are no longer in a black-box category. Furthermore, with inadequate choices of active orbitals, they can miss information about certain classes of higher-order dynamical correlation effects that may be needed to achieve desired accuracy levels. While, in analogy to the ground-state case, these higher-order effects can be captured by the CC(P;Q) corrections [47] to the EOMCCSDt or EOMCCSDtq energies, and we have made progress toward extending our CC(P;Q)-inspired CC(t;3), CC(t,q;3), and CC(t,q;3,4)

codes [47,53,54] to excited states, which will be reported elsewhere, the resulting methods continue to rely on the user- and system-dependent active orbitals, which is not ideal.

Encouraged by our earlier ground-state work [55], in this study we choose an alternative way of utilizing the CC(P;Q) framework in the EOMCC calculations for excited electronic states, which allows us to achieve desired high accuracy levels without resorting to the active-space concepts by taking advantage of the stochastic CI Quantum Monte Carlo (QMC) methodology introduced in Refs. [56,57]. In analogy to our previous study [58], in the semi-stochastic CC(P;Q) approach to excited states proposed in this article, the dominant higher–than–two-body components of the T and R_μ operators within the high-level deterministic EOMCC computations, such as EOMCCSDT, are identified stochastically using CIQMC propagations, but, unlike in Ref. [58], we go one significant step further and accelerate convergence toward target EOMCC (e.g., EOMCCSDT) energetics by using the *a posteriori* CC(P;Q) corrections to account for those correlations of interest that the EOMCC computations carried out in the stochastically determined excitation subspaces have not been able to capture. By examining vertical excitations in the CH$^+$ ion at the equilibrium and stretched geometries and adiabatic excitations in the CH and CNC open-shell systems, we demonstrate that the extension of the semi-stochastic CC(P;Q) methodology of Ref. [55] to excited states is capable of rapidly converging target high-level EOMCC energetics, represented in this study by EOMCCSDT, out of the early stages of CIQMC propagations, even when the excited states of interest have a significant double excitation or multireference character.

2. Theory and algorithmic details

2.1. Brief synopsis of the CC(P;Q) formalism

We begin our description of the semi-stochastic approach proposed in this work by summarizing the key ingredients of the underlying CC(P;Q) theory, as applied to ground as well as excited states [47,53]. The CC(P;Q) formalism is a generalization of the biorthogonal moment expansions of Refs. [35,59,60], which in the past resulted in the CR-CC(2,3) [35,59,60], CR-EOMCC(2,3) [35–37], and δ-CR-EOMCC(2,3) [38,46] corrections to the CCSD and EOMCCSD energies, to arbitrary, i.e., conventional as well as unconventional, truncations used in the underlying CC/EOMCC computations. The CC(P;Q) energies are obtained in two steps. In the first step, abbreviated as CC(P) for the ground ($\mu = 0$) state and EOMCC(P) for excited ($\mu > 0$)

states, we solve the CC/EOMCC equations in the subspace of the N-electron Hilbert space called the P space, designated as $\mathcal{H}^{(P)}$, which is spanned by excited determinants $|\Phi_K\rangle = E_K |\Phi\rangle$ that 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 clarity of this brief 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., we start by determining the cluster operator $T^{(P)} = \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} t_K E_K$, with t_K representing the corresponding cluster amplitudes, and the ground-state CC(P) energy $E_0^{(P)} = \langle \Phi | \bar{H}^{(P)} | \Phi \rangle$, where $\bar{H}^{(P)} = \exp(-T^{(P)}) H \exp(T^{(P)})$. We then diagonalize the similarity-transformed Hamiltonian $\bar{H}^{(P)}$ 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,

$$R_\mu^{(P)} = r_{\mu,0} \mathbf{1} + \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} r_{\mu,K} E_K \quad \text{and} \quad L_\mu^{(P)} = \delta_{\mu,0} \mathbf{1} + \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} l_{\mu,K} (E_K)^\dagger,$$

respectively, where $r_{\mu,K}$ and $l_{\mu,K}$ designate the relevant amplitudes, which define the EOMCC(P) ket states $|\Psi_\mu^{(P)}\rangle = R_\mu^{(P)} e^{T^{(P)}} |\Phi\rangle$ and the CC(P)/EOMCC(P) bra states $\langle \tilde{\Psi}_\mu^{(P)}| = \langle \Phi | L_\mu^{(P)} e^{-T^{(P)}}$ satisfying $\langle \tilde{\Psi}_\mu^{(P)} | \Psi_\nu^{(P)} \rangle = \delta_{\mu\nu}$. Once this is done, we proceed 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 many-electron correlation effects captured by the second 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 formula for these corrections is

$$\delta_\mu(P;Q) = \sum_{|\Phi_K\rangle \in \mathcal{H}^{(Q)}} \ell_{\mu,K}(P) \, \mathfrak{M}_{\mu,K}(P), \tag{1}$$

where $\mathfrak{M}_{0,K}(P) = \langle \Phi_K | \bar{H}^{(P)} | \Phi \rangle$ and $\mathfrak{M}_{\mu,K}(P) = \langle \Phi_K | \bar{H}^{(P)} R_\mu^{(P)} | \Phi \rangle$ are the generalized moments of the CC(P) and EOMCC(P) equations, which correspond to projections of these equations on the Q-space determinants $|\Phi_K\rangle \in \mathcal{H}^{(Q)}$, and $\ell_{\mu,K}(P) = \langle \Phi | L_\mu^{(P)} \bar{H}^{(P)} | \Phi_K \rangle / D_{\mu,K}(P)$, with

$D_{\mu,K}(P) = E_\mu^{(P)} - \langle \Phi_K | \bar{H}^{(P)} | \Phi_K \rangle$ (one could replace the Epstein–Nesbet $D_{\mu,K}(P)$ denominator entering $\ell_{\mu,K}(P)$ by its Møller–Plesset analog, but, as shown in the past, for example in Refs. [36,54,59–61], the Epstein–Nesbet form is generally more effective). The final CC(P;Q) electronic energies are determined as

$$E_\mu^{(P+Q)} = E_\mu^{(P)} + \delta_\mu(P;Q). \tag{2}$$

The question arises how to define the P and Q spaces entering the CC(P;Q) considerations to obtain accurate ground- and excited-state energetics matching the quality of high-level CC/EOMCC calculations without incurring large computational costs of CCSDT/EOMCCSDT and similar approaches. One can try conventional choices where, for example, $\mathcal{H}^{(P)}$ is spanned by the singly and doubly excited determinants and $\mathcal{H}^{(Q)}$ by the triples, but, as already alluded to above, the resulting CR-CC(2,3) and CR-EOMCC(2,3) corrections to the CCSD and EOMCCSD energies, which decouple the low-order T_n and $R_{\mu,n}$ components with $n \le 2$ from their higher-order T_3 and $R_{\mu,3}$ counterparts, may have difficulties 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 aforementioned CC(t;3), CC(t,q;3), CC(t,q;3,4), etc. hierarchy, but the resulting methods are no longer computational black boxes. The semi-stochastic CC(P;Q) approach to excited-state calculations, which we describe next and which, following our earlier work [58], exploits the CIQMC propagations to identify the leading higher–than–doubly excited determinants pertinent to the CC/EOMCC calculations of interest, 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.

2.2. Semi-stochastic CC(P;Q) approach to ground and excited states

In our previous studies [55,58], we demonstrated that the CIQMC methodology of Refs. [56,57] is very good in identifying the leading determinants and generating meaningful P spaces for the deterministic CC(P)/EOMCC(P) calculations already in the early stages of QMC propagations without any *a priori* knowledge of the states being calculated. We show in this work that the excited-state CC(P;Q) corrections $\delta_\mu(P;Q)$, defined by Eq. (1), similarly to their $\mu = 0$

ground-state counterparts examined in Ref. [55], are highly effective in accounting for the many-electron correlation effects outside the stochastically determined P spaces. While the specific computations reported in this work, which aim at recovering the EOMCCSDT energetics, rely on the FCIQMC [56,57] propagations to identify the dominant triply excited determinants for defining the relevant P spaces, the algorithm summarized below is quite general, permitting the use of truncated CIQMC and CCMC [62,63] approaches and extensions to higher EOMCC levels than EOMCCSDT, such as EOMCCSDTQ (not implemented yet). In the stochastic part of the excited-state CC(P;Q) algorithm proposed in this work, we rely on the initiator CIQMC (i-CIQMC) approach developed in Ref. [57], which we also exploited in our earlier [55,58,64] studies, where only determinants with numbers of walkers equal to or exceeding a preset threshold n_a are allowed to attempt spawning, but we could certainly take advantage of improvements in the original i-CIQMC [57] and i-CCMC [65] algorithms, such as those recently reported in Refs. [66–68]. It is also worth pointing out that by combining the stochastic CIQMC and deterministic EOMCC ideas via the CC(P;Q) methodology, we can extract highly accurate excited-state information on the basis of relatively short CIQMC propagations for the ground state or the lowest-energy state of a given symmetry, without having to resort to the more complex excited-state CIQMC framework proposed in Refs. [69,70], although exploring the utility of the latter framework would be an interesting direction to pursue.

The key steps of the semi-stochastic CC(P;Q) algorithm proposed in this article, which builds upon the semi-stochastic CC(P)/EOMCC(P) framework suggested in Ref. [58] (steps 1–3 below) and which extends the previously developed [55] merger of the ground-state CC(P;Q) methodology with CIQMC or CCMC to excited states, are as follows:

1. Initiate a CIQMC (or CCMC) 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 (in CCMC, "excips" [63,65]) 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 (or CCMC) 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. 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 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/excips (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 (or CCMC) 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 (or CCMC) 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 $T^{(P)} = T_1 + T_2 + T_3^{(MC)}$, $R_\mu^{(P)} = r_{\mu,0}\mathbf{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3}^{(MC)}$, and $L_\mu^{(P)} = \delta_{\mu,0}\mathbf{1} + L_{\mu,1} + L_{\mu,2} + L_{\mu,3}^{(MC)}$, where the list of triples in $T_3^{(MC)}$, $R_{\mu,3}^{(MC)}$, and $L_{\mu,3}^{(MC)}$ is extracted from the ground-state CIQMC (or CCMC) 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 (or CCMC) propagations for the lowest states of these irreps to define the lists of triples in $R_{\mu,3}^{(MC)}$ and $L_{\mu,3}^{(MC)}$. We follow a similar procedure when targeting the CCSDTQ/EOMCCSDTQ-level energetics, in which case $T^{(P)} = T_1 + T_2 + T_3^{(MC)} + T_4^{(MC)}$, $R_\mu^{(P)} = r_{\mu,0}\mathbf{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3}^{(MC)} + R_{\mu,4}^{(MC)}$, and $L_\mu^{(P)} = \delta_{\mu,0}\mathbf{1} + L_{\mu,1} + L_{\mu,2} + L_{\mu,3}^{(MC)} + L_{\mu,4}^{(MC)}$.

4. Correct the CC(*P*) and EOMCC(*P*) energies for the missing correlations of interest that were not captured by the CIQMC (or CCMC) 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. (1).

5. Check the convergence of the resulting $E_\mu^{(P+Q)}$ energies calculated using Eq. (2) by repeating steps 2–4 at some later CIQMC (or CCMC) 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. [55], 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. [55], 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 or i-CCMC walker/excip 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. [55,58], when the most expensive $\langle \Phi_{ijk}^{abc} | [H, T_3] | \Phi \rangle$ (or $\langle \Phi_{ijk}^{abc} | [\bar{H}^{(2)}, T_3] | \Phi \rangle$, where $\bar{H}^{(2)} = \exp(-T_1 - T_2) H \exp(T_1 + T_2)$) and $\langle \Phi_{ijk}^{abc} | [\bar{H}^{(P)}, R_{\mu,3}] | \Phi \rangle$ 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 labeled 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 $\langle \Phi_{ijk}^{abc} | [H, T_2] | \Phi \rangle$ and $\langle \Phi_{ijk}^{abc} | [\bar{H}^{(P)}, R_{\mu,2}] | \Phi \rangle$ or $\langle \Phi_{ij}^{ab} | [H, T_3] | \Phi \rangle$ and $\langle \Phi_{ij}^{ab} | [\bar{H}^{(P)}, R_{\mu,3}] | \Phi \rangle$, 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 $\langle \Phi_{ijk}^{abc} | [H, T_3] | \Phi \rangle$ and $\langle \Phi_{ijk}^{abc} | [\bar{H}^{(P)}, R_{\mu,3}] | \Phi \rangle$ contributions 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 $\sim 2 n_o^3 n_u^4$, as opposed to the $n_o^3 n_u^5$ scaling of every CCSDT and EOMCCSDT iteration).

3. Numerical examples and discussion

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) calculations toward the desired EOMCC energetics, represented in this study by EOMCCSDT, we carried out benchmark calculations for the frequently studied (cf., e.g., Refs. [18,20,21,32,35,39–41,43,44,49–51,71]) vertical excitations in the CH^+ ion at the equilibrium (Table 1 and Fig. 1 (a) and (b)) and stretched (Table 2 and Fig. 1 (c) and (d)) geometries, which we previously used to test the EOMCC(P) framework [58], and the adiabatic excitations in the challenging open-shell CH (Table 3) and CNC (Table 4) systems, which have low-lying excited states dominated by two-electron transitions that require the EOMCCSDT theory level to obtain a reliable description [26,36,45,50,72–75]. The CH^+ ion was described by the [5s3p1d/3s1p] basis set of Ref. [71] and we used the aug-cc-pVDZ [76,77] and DZP[4s2p1d] [78,79] bases for the CH and CNC species, respectively. Following Refs. [55,58] (cf., also, Ref. [64]), we used the HANDE software package

[80,81] 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 [82,83], 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). It should be noted that 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)/CR-EOMCC(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 ∞. Each i-FCIQMC run was initiated by placing 1500 walkers on the relevant reference function (see Tables 1–4 for the details) and we set the initiator parameter n_a at 3. All of the i-FCIQMC propagations used the time step $\Delta\tau$ 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.

3.1. 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 (labeled 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 $2\,^1\Delta$). Two C–H internuclear separations were considered, the

equilibrium distance $R = R_e = 2.13713$ bohr (Table 1 and Fig. 1 (a) and (b)) and the stretched $R = 2R_e$ geometry (Table 2 and Fig. 1 (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 $T^{(P)} = T_1 + T_2 + T_3^{(MC)}$, $R_\mu^{(P)} = r_{\mu,0}\mathbf{1} + R_{\mu,1} + R_{\mu,2} + R_{\mu,3}^{(MC)}$, and $L_\mu^{(P)} = \delta_{\mu,0}\mathbf{1} + L_{\mu,1} + L_{\mu,2} + L_{\mu,3}^{(MC)}$, respectively, where the list of triples defining the three-body components $T_3^{(MC)}$, $R_{\mu,3}^{(MC)}$, and $L_{\mu,3}^{(MC)}$ at a given time τ was obtained from the ground-state i-FCIQMC propagation at the same value of τ. The $T_3^{(MC)}$ 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 $\overline{H}^{(P)}$ 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 $R_{\mu,3}^{(MC)}$ component of $R_\mu^{(P)}$ and the $L_{\mu,3}^{(MC)}$ 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 $^1B_1(C_{2v})$ component of $1\,^1\Pi$ for the $^1\Pi$ states and the $^1A_2(C_{2v})$ component of $1\,^1\Delta$ for the $^1\Delta$ states (C_{2v} is the largest Abelian subgroup of the true point group of CH^+, $C_{\infty v}$; our codes cannot handle non-Abelian symmetries). As implied by Eq. (1), 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. [20,21,58], the $2\,^1\Sigma^+$, $2\,^1\Pi$, $1\,^1\Delta$, and $2\,^1\Delta$ states of CH^+ at $R = R_e$ and all of the excited states of the stretched $CH^+/R = 2R_e$ system, which we calculated in this work, are characterized by substantial multireference correlations that originate from large two-electron excitation contributions (the $2\,^1\Delta$ state at $R = 2R_e$ also has significant triple excitations [20,21,58]). 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 $2\,^1\Pi$, $2\,^1\Sigma^+$, and both $^1\Delta$ states, respectively, at $R = R_e$ and ~14–144 millihartree when the excited states at $R = 2R_e$ are considered (see Tables 1 and 2). The EOMCCSD energies for the $3\,^1\Sigma^+$, $4\,^1\Sigma^+$, and $1\,^1\Pi$ 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 1 and 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 article, but there are cases, especially the $4\,^1\Sigma^+$ and $2\,^1\Delta$ states at $R = 2R_e$, where the differences between the CR-EOMCC(2,3) and parent EOMCCSDT energies, which are about 13 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 older active-space EOMCCSDt calculations for CH^+ reported in Refs. [20,21] and the more recent semi-stochastic EOMCC(P) calculations for the same system described in Ref. [58] are telling us that the incorporation 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 1 and 2 and Fig. 1. In agreement with our previous work [58], 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. (1) accelerate the convergence toward EOMCCSDT even further. As shown in Tables 1 and 2 and Fig. 1, 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 the early stages of the i-FCIQMC propagations, when no more than 5–10 % of all triples are includes in the relevant P spaces. This is true even when the highly complex $4\,^1\Sigma^+$ and $2\,^1\Delta$ states at $R = 2R_e$, 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 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 times 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 1 and 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. [20,21], 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 the 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 space consisting of the 3σ, $1\pi_x$, $1\pi_y$, and 4σ valence orbitals amounts to about 26–29 % of all triples included in the respective EOMCC diagonalization spaces [20,21]. 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. [20,21], and those that define the $R_{\mu,3}^{(MC)}$ 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. 2, where we compare the distributions of the differences between the $R_{\mu,3}^{(MC)}$ amplitudes and their full

EOMCCSDT counterparts resulting from the EOMCC(P) computations at 4000 (Fig. 2 (a)), 10000 (Fig. 2 (b)), and 50000 (Fig. 2 (c)) MC iterations 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 using the 3σ, $1\pi_x$, $1\pi_y$, and 4σ active orbitals to define the corresponding triples space (Fig. 2 (d); all EOM vectors R_μ needed to construct Fig. 2, corresponding to the EOMCC(P), EOMCCSDt, and EOMCCSDT calculations, were normalized to unity). As shown in Fig. 2 (cf. Figs. 2 (c) and 2 (d)), the small differences between the $R_{\mu,3}^{(MC)}$ 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. [20,21]. This is in perfect agreement with the small errors relative to EOMCCSDT characterizing the two calculations, which are 0.302 millihartree in the former case (cf. Table 2) and 0.576 millihartree in the case of EOMCCSDt [20,21]. 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 $R_{\mu,3}^{(MC)}$ 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. 2, 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 (~30 %), and in Table 2, where the error in the 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 ~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 the CC(t;3) [47,53,54], whose performance in excited-state calculations will be reported by us in a separate 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. We have already observed this in our previous ground-state work [55], 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. [20,21,47,50,51,53], 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. [54]), 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 Tables 1 and 2 for CH[+] and in the remaining Tables 3 and 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 1, one needs to capture about 50 % of triples in the *P* space to reach a 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 2), one reaches a ~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 Section 2.2, 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 ~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. [84] for the analogous comments made in the context of comparing costs of the CCSDt computations with those of CC(t;3)).

3.2. CH

Similar convergence patterns in the semi-stochastic EOMCC(P) and CC(P;Q) calculations are observed for the CH radical (see Table 3). In this case, following our earlier deterministic EOMCC work, including the CR-EOMCC [36,45] and electron-attachment (EA) EOMCC [36,85,86] approaches, and a wide range of EOMCC computations, including the high EOMCCSDT and EOMCCSDTQ levels, published by Hirata [26], 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 [26,36,45,85,86], 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 [87], 1.1031 Å for the A $^2\Delta$ state [87], 1.1640 Å for the B $^2\Sigma^-$ state [88], and 1.1143 Å for the C $^2\Sigma^+$ state [89] (cf. Table 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 and 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 [26] 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 deexcitation 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 our earlier papers [36,45,85,86] and as shown in Ref. [26], 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. [36] or Table II of Ref. [45]), 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, 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 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-stochastic CC(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, 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-stochastic CC(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 spaces, the CC(P;Q) energies of the A $^2\Delta$ and B $^2\Sigma^-$ states are within fractions of a millihartree from EOMCCSDT. These are considerable 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-stochastic EOMCC(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 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.

3.3. 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 [36,74,75], we considered the X $^2\Pi_g$ ground state and 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. [74] 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 $T_3^{(\mathrm{MC})}$, $R_{\mu,3}^{(\mathrm{MC})}$, and $L_{\mu,3}^{(\mathrm{MC})}$ 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 4 and in agreement with one of our previous studies [36], 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 4). The excessively large, > 111 millihartree, error in the EOMCCSD energy of the $B\,^2\Sigma_u^+$ 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. [36]). 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 4). By incorporating the dominant triply excited determinants captured by the i-FCIQMC propagations in the respective P spaces, the semi-stochastic CC(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.

4. Conclusions

We have demonstrated that the semi-stochastic, CIQMC-driven, CC(P;Q) algorithm developed in this work, which is based on correcting the results of the EOMCC(P) diagonalizations, using the ideas presented in Refs. [55,58], for the effects of higher–than–doubly excited determinants that were not captured by the corresponding CIQMC propagations, is capable of faithfully reproducing the parent high-level EOMCC energetics, such as EOMCCSDT, out of the early stages of CIQMC runs. We have illustrated the performance of this algorithm by reporting the results of the CC(P;Q) and underlying CC(P) and EOMCC(P) calculations aimed at converging the EOMCCSDT energetics for a few benchmark systems, including CH^+ at the equilibrium and stretched geometries and the open-shell CH and CNC species. While the details of the rate of convergence of the semi-stochastic CC(P;Q) calculations toward EOMCCSDT, especially when one wants to tighten it, may depend on the specific excited state being calculated, the combination of the CIQMC wave function sampling with the CC(P;Q) corrections to the EOMCC(P) energies proposed in this work is as efficient as its ground-state counterpart developed in Ref. [55]. We have shown that the excited-state CC(P;Q) corrections accelerate the convergence of the uncorrected EOMCC(P) energies in the same way as their ground-state analogs, allowing

21

us to reach millihartree or sub-millihartree accuracy levels relative to the parent EOMCC (in this study, EOMCCSDT) methodology with the relatively small fractions of higher–than–doubly excited determinants captured in the early stages of the CIQMC runs. By relaxing the one and two-body components of the cluster and EOM excitation and deexcitation operators in the presence of their higher-order counterparts, which are determined using the excitation lists provided by CIQMC, they are capable of considerably improving the CR-EOMCC(2,3) corrections to EOMCCSD without making the calculations a lot more expensive. This is true for both the excited states dominated by one-electron transitions and for the more strongly correlated multireference states having significant double excitation character.

The results reported in this study are promising, but, in addition to code optimization, we need more testing, including larger molecules and larger basis sets, for which parent EOMCCSDT data can be generated (cf., e.g., Refs. [90–92]), and examining extensions of our semi-stochastic CC(P;Q) methodology beyond the CCSDT and EOMCCSDT levels. It would also be interesting to consider various ways of modifying the semi-stochastic excited-state CC(P;Q) algorithm proposed in this work, such as the state-specific generalization suggested at the EOMCC(P) level in Ref. [58], where one would use as many stochastically determined P spaces as the number of the calculated states by taking advantage of the excited-state CIQMC framework proposed in Refs. [69,70]. As shown, for example, in Table 2 and Fig. 1 (d), the CC(P;Q) energies of the $4\,^1\Sigma^+$ state of CH^+ at $R = 2R_e$, while accelerating convergence of the uncorrected EOMCC(P) energies toward EOMCCSDT, converge to EOMCCSDT slower than the CC(P;Q) calculations for the $n\,^1\Sigma^+$ states with $n = 1 - 3$. Similarly, the CC(P;Q) energies of the $2\,^1\Delta$ state of CH^+ converge toward the parent EOMCCSDT energetics somewhat slower than those obtained for the $1\,^1\Delta$ state (see Tables 1 and 2). Thus, using the excited-state CIQMC framework of Refs. [69,70] to adjust each P space employed in the CC(P;Q) calculations to the corresponding excited electronic state of interest is a direction worth pursuing. This may become especially important for the high-lying excitations, for which the P spaces determined for the lowest-energy states of the relevant symmetries may no longer be adequate. Last, but not least, we plan to explore other ways of capturing the leading higher–than–doubly excited determinants to define the P spaces needed in the CC(P;Q) considerations, such as the adaptive CI [93,94], adaptive sampling CI [95,96], heat-bath CI [97–99], and related methodologies, to mention a few representative examples.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

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

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

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

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

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

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

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

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

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

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

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

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

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

[14] D. Mukherjee and P.K. Mukherjee, Chem. Phys. **39**, 325 (1979).

[15] H. Sekino and R.J. Bartlett, Int. J. Quantum Chem., Symp. **18**, 255 (1984).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[33] K. Kowalski, S. Hirata, M. Włoch, P. Piecuch, and T.L. Windus, J. Chem. Phys. **123**, 074319 (2005).

[34] C.D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).

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

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

[37] J.J. Lutz and P. Piecuch, Comput. Theor. Chem. **1040**–**1041**, 20 (2014).

[38] P. Piecuch, J.A. Hansen, and A.O. Ajala, Mol. Phys. **113**, 3085 (2015).

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

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

[41] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **105**, 1451 (1996).

[42] O. Christiansen, H. Koch, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **256**, 185 (1996).

[43] S. Hirata, M. Nooijen, I. Grabowski, and R.J. Bartlett, J. Chem. Phys. **114**, 3919 (2001); **115**, 3967 (2001) [Erratum].

[44] T. Shiozaki, K. Hirao, and S. Hirata, J. Chem. Phys. **126**, 244106 (2007).

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

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

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

[48] H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **244**, 75 (1995).

[49] O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **103**, 7429 (1995).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[68] V.A. Neufeld and A.J.W. Thom, J. Chem. Theory Comput. **16**, 1503 (2020).

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

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

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

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

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

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

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

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

[77] R.A. Kendall, T.H. Dunning Jr., and R.J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

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

[79] T.H. Dunning Jr. and P.J. Hay, in *Methods of Electronic Structure Theory*, edited by H.F. Schaefer III (Plenum, New York, 1977), pp. 1–27.

[80] 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. Software **3**, e9 (2015).

[81] 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.-Q. Xu, and A.J.W. Thom, J. Chem. Theory Comput. **15**, 1728 (2019).

[82] 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 Jr., J. Comput. Chem. **14**, 1347 (1993).

[83] 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).

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

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

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

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

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

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

[90] D. Kánnár, A. Tajti, and P.G. Szalay, J. Chem. Theory Comput. **13**, 202 (2017).

[91] P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, and D. Jacquemin, J. Chem. Theory Comput. **15**, 1939 (2019).

[92] P.-F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama, and D. Jacquemin, J. Chem. Theory Comput. **16**, 1711 (2020).

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

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

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

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

[97] A.A. Holmes, H.J. Changlani, and C.J. Umrigar, J. Chem. Theory Comput. **12**, 1561 (2016).

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

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

Table 1. Convergence of the CC(P)/EOMCC(P) and CC($P;Q$) energies toward CCSDT/EOMCCSDT for the CH$^+$ ion, calculated using the [5s3p1d/3s1p] basis set of Ref. [71], 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 $^1\Sigma^+$ states, the $3\sigma \to 1\pi$ state of the $^1B_1(C_{2v})$ symmetry for the $^1\Pi$ states, and the $3\sigma^2 \to 1\pi^2$ 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 the i-FCIQMC.

MC Iters. ($\times 10^3$)	1 $^1\Sigma^+$			2 $^1\Sigma^+$		3 $^1\Sigma^+$		4 $^1\Sigma^+$		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 $^1\Sigma^+$ ground state) and EOMCC(P) (excited states) energies relative to the corresponding CCSDT and EOMCCSDT data, in millihartree [58].

[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 $^1\Sigma^+$ = 1 $^1A_1(C_{2v})$ ground state for the $^1\Sigma^+$ 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 τ = 0.0 a.u. are identical to the energies obtained in the CCSD and EOMCCSD calculations. The τ = 0.0 a.u. CC($P;Q$) energies are equivalent to the CR-CC(2,3) (the ground state) and 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 CCSDT and EOMCCSDT calculations (see Refs. [20,21]).

Table 2. Same as Table 1 for the stretched C–H internuclear distance $R = 2R_e = 4.27426$ bohr.

MC Iters. ($\times 10^3$)	1 $^1\Sigma^+$			2 $^1\Sigma^+$		3 $^1\Sigma^+$		4 $^1\Sigma^+$		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	0.006	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	

Table 3. Convergence of the CC(P)/EOMCC(P) and CC(P;Q) energies toward CCSDT/EOMCCSDT for the CH radical, calculated using the aug-cc-pVDZ [76,77] 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 $^2\Pi$ ground 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 the i-FCIQMC.

MC Iters. ($\times 10^3$)	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^d	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. [26,36,45], which are 1.1197868 Å for the X $^2\Pi$ state [87], 1.1031 Å for the A $^2\Delta$ state [87], 1.1640 Å for the B $^2\Sigma^-$ state [88], and 1.1143 Å for the C $^2\Sigma^+$ state [89]. 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. [26,36,45] (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 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 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.

Table 4. Convergence of the CC(P)/EOMCC(P) and CC($P;Q$) energies toward CCSDT/EOMCCSDT for the CNC molecule, calculated using the DZP[4s2p1d] basis set of Refs. [78,79]. 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 $^2B_{2g}(D_{2h})$ determinant for the X $^2\Pi_g$ ground state and the $3\sigma_u \rightarrow 1\pi_g$ state of the $^2B_{1u}(D_{2h})$ 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 the i-FCIQMC.

MC Iters. ($\times10^3$)	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[d]	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) or EOMCC(P) (the remaining states) energies relative to the corresponding CCSDT or EOMCCSDT data, in millihartree, calculated at the equilibrium C–N distances optimized in Ref. [74], 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 orbitals were frozen in all correlated calculations.

[b]Errors in the CC($P;Q$) energies relative to the corresponding CCSDT or EOMCCSDT data, in millihartree, calculated at the equilibrium C–N distances optimized in Ref. [74] (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 $^2B_{2g}(D_{2h})$ component of the X $^2\Pi_g$ ground state and the lowest $^2B_{1u}(D_{2h})$ state for the A $^2\Delta_u$ 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 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 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.
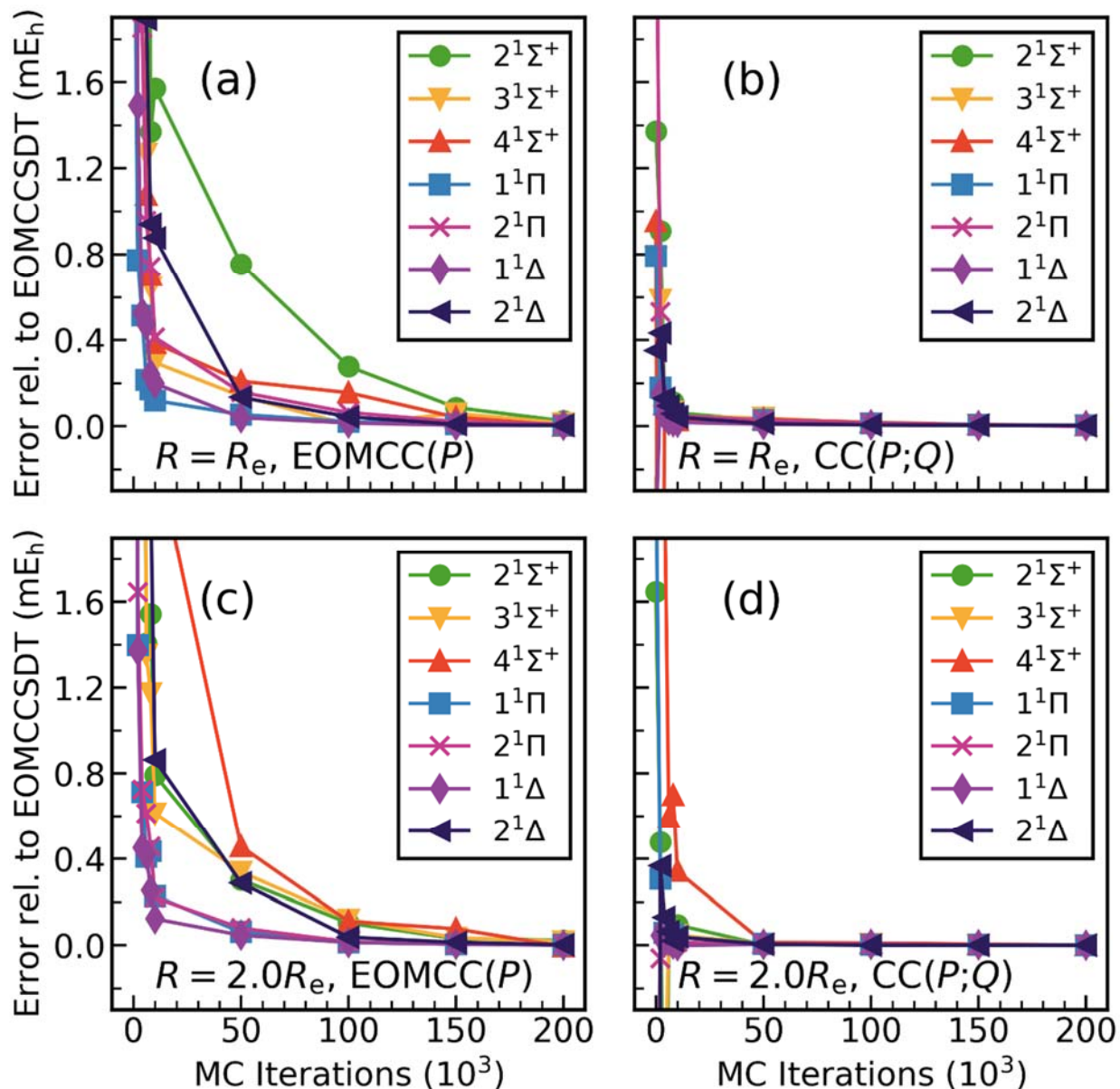
Figure 1. Convergence of the EOMCC(P) (panels (a) and (c)) and CC(P;Q) (panels (b) and (d)) energies toward EOMCCSDT for the three lowest-energy excited states of the $^1\Sigma^+$ symmetry, two lowest states of the $^1\Pi$ symmetry, and two lowest $^1\Delta$ states of the CH$^+$ ion, as described by the [5s3p1d/3s1p] basis set of Ref. [71], 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)).
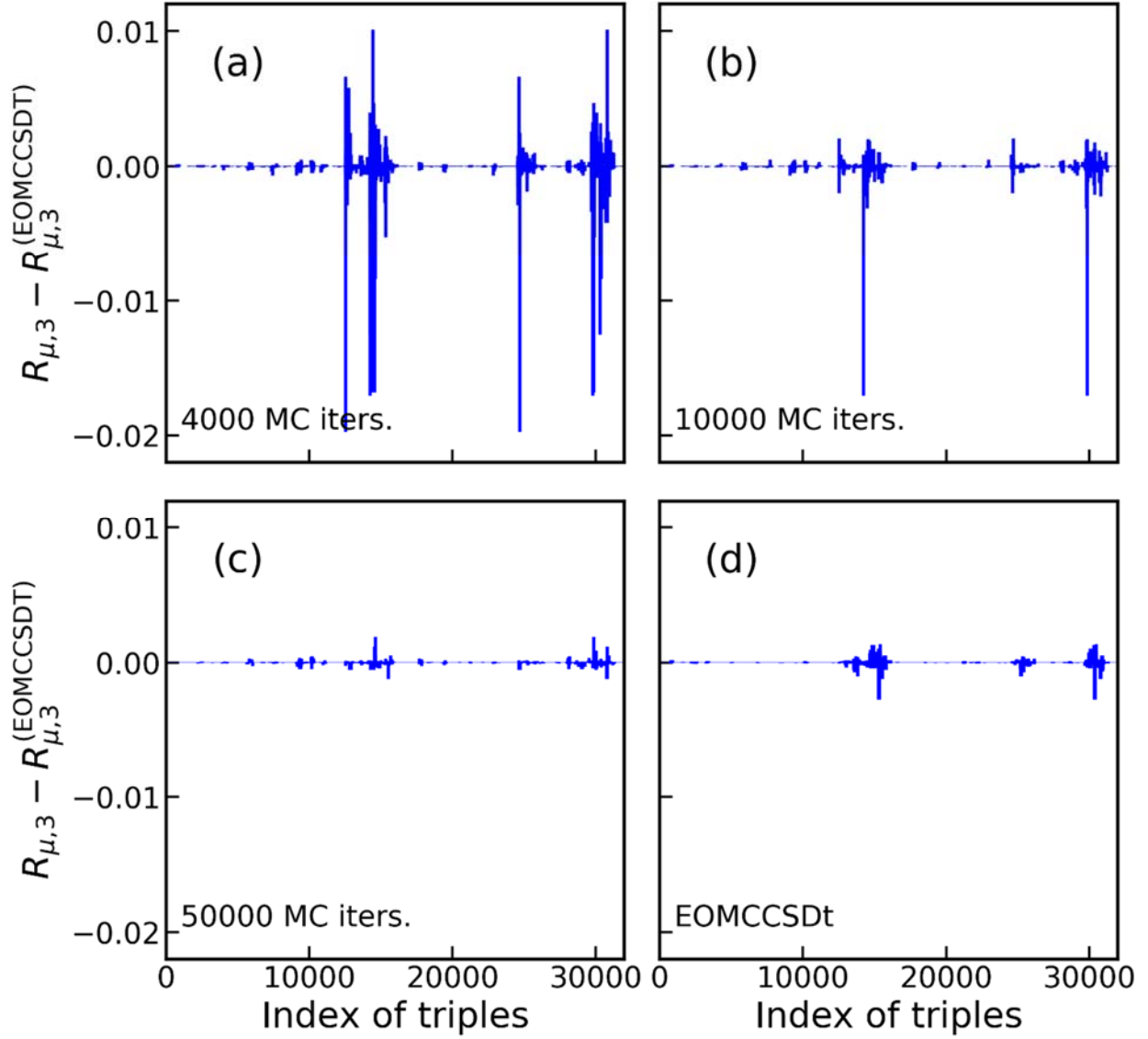
Figure 2. A comparison of the distributions of the differences between the $R_{\mu,3}^{(MC)}$ amplitudes and their EOMCCSDT counterparts resulting from the EOMCC(P) computations at (a) 4000, (b) 10000, and (c) 50000 MC iterations using $\Delta\tau = 0.0001$ a.u. for the $2\,^1\Sigma^+$ state of CH$^+$ 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.