High-Level Coupled-Cluster Energetics by Monte Carlo Sampling and Moment Expansions: Further Details and Comparisons

J. Emiliano Deustua,\textsuperscript{1} Jun Shen,\textsuperscript{1} and Piotr Piecuch\textsuperscript{1,2}\textsuperscript{a)}

\textsuperscript{1) Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, USA}
\textsuperscript{2) Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA}

(Dated: 23 February 2021)

We recently proposed a novel approach to converging electronic energies equivalent to high-level coupled-cluster (CC) computations by combining the deterministic CC($P$,$Q$) formalism with the stochastic configuration interaction (CI) and CC Quantum Monte Carlo (QMC) propagations. This article extends our initial study [J. E. Deustua, J. Shen, and P. Piecuch, \textit{Phys. Rev. Lett.} \textbf{119}, 223003 (2017)], which focused on recovering the energies obtained with the CC method with singles, doubles, and triples (CCSDT) using the information extracted from full CI QMC and CCSDT-MC, to the CIQMC approaches truncated at triples and quadruples. It also reports our first semi-stochastic CC($P$,$Q$) calculations aimed at converging the energies that correspond to the CC method with singles, doubles, triples, and quadruples (CCSDTQ). The ability of the semi-stochastic CC($P$,$Q$) formalism to recover the CCSDT and CCSDTQ energies, even when electronic quasi-degeneracies and triply and quadruply excited clusters become substantial, is illustrated by a few numerical examples, including the F–F bond breaking in F\textsubscript{2}, the automerization of cyclobutadiene, and the double dissociation of the water molecule.

\section{I. INTRODUCTION}

One of the main goals of quantum chemistry is to provide an accurate and systematically improvable description of many-electron correlation effects needed to determine molecular potential energy and property surfaces and understand chemical reactivity and various types of spectroscopy. In searching for the best solutions in this area, the size extensive methods based on the exponential wave function ansatz\textsuperscript{1,2} of coupled-cluster (CC) theory\textsuperscript{3–7}

\begin{equation}
|\Psi\rangle = e^T |\Phi\rangle,
\end{equation}

where

\begin{equation}
T = \sum_{n=1}^{N} T_n
\end{equation}

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 extensions to excited, open-shell, and multi-reference states\textsuperscript{8–12} are among the top contenders. In this study, we focus on the higher-rank members of the single-reference CC hierarchy beyond the basic CC singles and doubles (CCSD) level, where $T$ is truncated at $T_2$\textsuperscript{13–16} especially on the CC approach with singles, doubles, and triples (CCSDT), where $T$ is truncated at $T_3$\textsuperscript{17–19} and the CC approach with singles, doubles, triples, and quadruples (CCSDTQ), where $T$ is truncated at $T_4$\textsuperscript{20–22} This is motivated by the fact that in great many cases relevant to chemistry, including molecular properties at equilibrium geometries, multi-reference situations involving smaller numbers of strongly correlated electrons, as in the case of bond breaking and formation in the course of chemical reactions, noncovalent interactions, and photochemistry, the single-reference CCSD, CCSDT, CCSDTQ, etc. methods and their equation-of-motion (EOM)\textsuperscript{23–27} and linear responses\textsuperscript{28–33} extensions rapidly converge to the exact, full configuration interaction (FCI) limit, allowing one to incorporate the relevant many-electron correlation effects in a conceptually straightforward manner through particle-hole excitations from a single Slater determinant defining the Fermi vacuum without loss of accuracy as the system becomes larger characterizing truncated CI methods\textsuperscript{34}.

The convergence of the single-reference CCSD, CCSDT, CCSDTQ, etc. hierarchy toward FCI in situations other than larger numbers of strongly entangled electrons is fast, but costs of the post-CCSD computations needed to achieve a quantitative description, which are determined by the iterative $n_o^3 n_u^5$ steps in the CCSDT case and the iterative $n_o^4 n_u^6$ steps in the case of CCSDTQ, where $n_o$ ($n_u$) is the number of occupied (unoccupied) correlated orbitals, are usually prohibitively expensive. This is why part of the CC method development effort has been devoted to finding approximate ways of incorporating higher--than--two-body components of the cluster operator $T$, i.e., $T_n$ components with $n > 2$, and the analogous higher-order components of the EOMCC excitation, electron-attachment, and electron-detachment operators, which could reduce enormous computational costs of the CCSDT, CCSDTQ, and similar schemes, while eliminating failures of the CCSD($T$)\textsuperscript{38} CCSD(T)\textsuperscript{39} CCSDT\textsuperscript{1} CCSDT\textsuperscript{1,2} CCSDT\textsuperscript{1} CCSDT\textsuperscript{1,2} and other perturbative CC approaches (cf. Ref. \textsuperscript{40} for a review) that fail when bond

\textsuperscript{a)}Corresponding author; e-mail: piecuch@chemistry.msu.edu.
breaking, biradicals, and other typical multi-reference situations in chemistry are examined. In fact, the analogous effort has been taking place in other areas of many-body theory, such as studies of nuclear matter, where a systematic, computationally efficient, and robust incorporation of higher-order many-particle correlation effects is every bit as important as in the case of electronic structure theory and where the quantum-chemistry-inspired CC and EOMCC methods, thanks, in part, to our group’s involvement, have become quite popular (see, e.g., Ref. [50] and references therein). While substantial progress in the above area, reviewed, for example, in Refs. [11,48, and 54], has already been made, the search for the optimal solution that would allow us to obtain the results of the full CCSDT, full CCSDTQ, or similar quality at the fraction of the cost and without having to rely on perturbative concepts or user- and system-dependent ideas, such as the idea of active orbitals to select higher-than-two-body components of the cluster and EOMCC excitation operators, continues.

In order to address this situation, we have started exploring a radically new way of converging accurate electronic energetics equivalent to those obtained with the high-level CC approaches of the full CCSDT, full CCSDTQ, and similar types, at the small fraction of the computational cost and preserving the black-box character of conventional single-reference methods, even when higher-than-two-body components of the cluster and excitation operators characterizing potential energy surfaces along bond stretching coordinates become large. The key idea of the approach suggested in Ref. [58], which we have recently extended to excited states, is a merger of the deterministic formalism, abbreviated as CC($P,Q$) [57,61,63], which enables one to correct energies obtained with conventional as well as unconventional truncations in the cluster and EOMCC excitation operators for any category of many-electron correlation effects of interest, with the stochastic FCI Quantum Monte Carlo (FCIQMC) [54,67] and CC Monte Carlo (CCMC) [68,71] methods (cf. Refs. [72,73] for alternative ways of combining FCIQMC with the deterministic CC framework). As shown in Refs. [58] and [68], where we reported preliminary calculations aimed at recovering full CCSDT and EOMCCSDT [26,28] energetics, the resulting semi-stochastic CC($P,Q$) methodology, using the FCIQMC and CCMC ideas that were originally laid down in Refs. [64,67, and 68]. Thus, we divide this section into two subsections. In Section II A, we summarize the key elements of the semi-stochastic CC($P,Q$) formalism, focusing on the ground-state problem relevant to the calculations reported in this study. Section II B provides information about the semi-stochastic CC($P,Q$) methods developed and tested in this work, which aim at converging the CCSDT and CCSDTQ energies with the help of the FCIQMC, CISDT-MC, and CISDTQ-MC propagations.

II. THEORY AND ALGORITHMIC DETAILS

As pointed out in the Introduction, the semi-stochastic CC($P,Q$) approach proposed in Ref. [58] is based on combining the deterministic CC($P,Q$) framework, developed mainly in Refs. [57,61, and 63], with the CIQMC and CCMC ideas that were originally laid down in Refs. [64,67, and 68]. Thus, we divide this section into two subsections. In Section II A, we summarize the key elements of the semi-stochastic CC($P,Q$) formalism, focusing on the ground-state problem relevant to the calculations reported in this study. Section II B provides information about the semi-stochastic CC($P,Q$) methods developed and tested in this work, which aim at converging the CCSDT and CCSDTQ energies with the help of the FCIQMC, CISDT-MC, and CISDTQ-MC propagations.

A. Basic Elements of the Ground-State CC($P,Q$) Formalism

The CC($P,Q$) formalism has emerged out of our interest in generalizing the biorthogonal moment energy expansions, which in the past resulted in the completely renormalized (CR) CC and EOMCC approaches, including CR-CC(2,3), CR-EOMCC(2,3), and CR-EOMCC(2,3) methods, and their higher-order extensions, such that one can correct the CC/EOMCC energies obtained with unconventional truncations in the cluster and EOMCC excitation operators, in addition to the conventional ones at a given many-body rank, for essentially any category of many-electron correlation effects of interest. The CC($P,Q$) framework is general, i.e., it applies to ground as well as
excited states, but since this work deals with the calculations that aim at recovering the ground-state CCSDT and CCSDTQ energetics, in the description below we focus on the ground-state CC(P,Q) theory.

According to the formal CC(P,Q) prescription, the ground-state energy of a N-electron system is determined in two steps. In the initial, iterative, CC(P) step, we solve the CC equations in the subspace \( \mathcal{H}^{(P)} \) of the \( N \)-electron Hilbert space \( \mathcal{H} \). We assume that subspace \( \mathcal{H}^{(P)} \), which we also call the \( P \) space, is spanned by the excited determinants \( |\Phi_K\rangle = E_K|\Phi\rangle \) that together with the reference determinant \( |\Phi\rangle \) provide the leading contributions to the target ground state \( |\Psi\rangle \) \( (E_K \) designates the usual elementary particle-hole excitation operator generating \( |\Phi_K\rangle \) from \( |\Phi\rangle \). In other words, we approximate the cluster operator \( T \) in Eq. (1) by

\[
T^{(P)} = \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} t_K E_K
\]

and solve the usual system of CC equations,

\[
\mathfrak{M}_K(P) = 0, \quad |\Phi_K\rangle \in \mathcal{H}^{(P)},
\]

where

\[
\mathfrak{M}_K(P) = \langle \Phi_K | \bar{H}^{(P)} | \Phi \rangle
\]

are the generalized moments of the \( P \)-space CC equations\(^{55-57}\) and

\[
\bar{H}^{(P)} = e^{-T^{(P)}} H e^{T^{(P)}} = (H e^{T^{(P)}})_C
\]

is the relevant similarity-transformed Hamiltonian, for the cluster amplitudes \( t_K \) (subscript \( C \) in Eq. (6) designates the connected operator product). Once the cluster operator \( T^{(P)} \) and the ground-state energy

\[
E^{(P)} = \langle \Phi | \bar{H}^{(P)} | \Phi \rangle
\]

that corresponds to it are determined, we proceed to the second step of CC(P;Q) considerations, which is the calculation of the noniterative correction \( \delta(P,Q) \) to the CC(P) energy \( E^{(P)} \) that accounts for the many-electron correlation effects captured by another subspace of the \( N \)-electron Hilbert space \( \mathcal{H} \), designated as \( \mathcal{H}^{(Q)} \) and called the \( Q \) space, which satisfies the condition \( \mathcal{H}^{(Q)} \subseteq (\mathcal{H}^{(0)} \oplus \mathcal{H}^{(P)})^\perp \), where \( \mathcal{H}^{(0)} \) is a one-dimensional subspace of \( \mathcal{H} \) spanned by the reference determinant \( |\Phi\rangle \). The formula for the \( \delta(P;Q) \) correction is\(^{57,58,60-63}\)

\[
\delta(P;Q) = \sum_{\langle \Phi_K\rangle \in \mathcal{H}^{(Q)}; \text{rank}(|\Phi_K\rangle) \leq \min(N^{(P)}, \Xi^{(Q)})} \ell_K(P) \mathfrak{M}_K(P),
\]

where integer \( N^{(P)} \) defines the highest many-body rank of the excited determinant \( |\Phi_K\rangle \) relative to \( |\Phi\rangle \) \( (\text{rank}(|\Phi_K\rangle)) \) for which moments \( \mathfrak{M}_K(P), \) Eq. (4), are still non-zero and \( \Xi^{(Q)} \) is the highest many-body rank of the excited determinant(s) \( |\Phi_K\rangle \) included in \( \mathcal{H}^{(Q)} \). In practical CC(P;Q) calculations, including those discussed in Section III, the \( \ell_K(P) \) coefficients entering Eq. (8) are calculated as

\[
\ell_K(P) = \langle \Phi | (1 + \lambda^{(P)}) \bar{H}^{(P)} | \Phi_K \rangle / D_K(P),
\]

where \( 1 \) is the unit operator,

\[
\Lambda^{(P)} = \sum_{|\Phi_K\rangle \in \mathcal{H}^{(P)}} \lambda_K(E_K)^\dagger
\]

is the hole-particle deexcitation operator defining the bra state \( |\psi^{(P)}\rangle = |\Phi | (1 + \lambda^{(P)}) e^{-T^{(P)}} \rangle \) corresponding to the CC(P) ket state \( |\psi^{(P)}\rangle = e^{-T^{(P)}} |\Phi\rangle \), and

\[
D_K(P) = E^{(P)} - \langle \Phi_K | \bar{H}^{(P)} | \Phi_K \rangle.
\]

One determines \( \Lambda^{(P)} \), or the amplitudes \( \lambda_K \) that define it, by solving the linear system of equations representing the left eigenstate CC problem\(^{10}\) in the \( P \) space, i.e.,

\[
\langle \Phi | (1 + \lambda^{(P)}) \bar{H}^{(P)} | \Phi_K \rangle = E^{(P)} \lambda_K, \quad |\Phi_K\rangle \in \mathcal{H}^{(P)},
\]

where \( E^{(P)} \) is the previously determined CC(P) energy. Once the noniterative correction \( \delta(P;Q) \) is determined, the CC(P;Q) energy is obtained as

\[
E^{(P+Q)} = E^{(P)} + \delta(P;Q).
\]

In practice, we often distinguish between the complete version of the CC(P;Q) theory, designated, following Refs. \(^{58,62}\), as CC(P;Q)\(_{\text{EN}}\), which uses the Epstein–Nesbet-like denominator \( D_K(P) \), Eq. (11), in calculating the \( \ell_K(P) \) amplitudes, and the approximate version of CC(P;Q), abbreviated as CC(P;Q)\(_{\text{MP}}\), which relies on the Möller–Plesset form of \( D_K(P) \) obtained by replacing \( \bar{H}^{(P)} \) in Eq. (11) by the bare Fock operator (cf., e.g., Refs. \(^{58,62}\) and \(^{63}\)). Both of these variants of the CC(P;Q) formalism are considered in this study.

We must now come up with the appropriate choices of the \( P \) and \( Q \) spaces entering the CC(P;Q) considerations that would allow us to match the quality of the high-level CC computations of the CCSDT, CCSDTQ, and similar type at the small fraction of the cost. As is often the case in the CC work, one could start from the conventional choices, where the \( P \) space \( \mathcal{H}^{(P)} \) is spanned by all excited \( |\Phi_{i_1 \ldots i_n}\rangle \) determinants with the excitation rank \( n \leq m_A \), where \( i_1, i_2, \ldots , a_1, a_2, \ldots \) designate the spin-orbitals occupied (unoccupied) in \( |\Phi\rangle \), and the \( Q \) space \( \mathcal{H}^{(Q)} \) by those with \( m_A < n < m_B \), where \( m_B \leq N \). In that case, one ends up with the well-established CR-CC\(_{\text{CCSDT}}(m_A, m_B)\) hierarchy\(^{52,57,63,75,80,83}\) including the aforementioned CR-CC\(_{\text{CCSDT}}(2,3)\) approximation, where \( m_A = 2 \) and \( m_B = 3 \), and the related CCSDT(2)\(_{\text{EN}}\) (cf., also, Refs. \(^{62,63}\), CCSD(T)\(_{\text{EN}}\)\(^{92,94}\)–\(^{95,96}\)) and similar schemes that allow one to correct the CCSD energies for triples. The CR-CC\(_{\text{CCSDT}}(2,3)\) method is useful, improving, for example,
poor performance of CCSD(T) in covalent bond breaking situations,
and in certain classes of noncovalent interactions, without a substantial increase of the computational effort, but neither CR-CC(2,3) nor its counterparts (which are all approximations to CC-CC(2,3)) are free from drawbacks. One of the main problems with CR-CC(2,3), CCSD(2)_T, Λ-CCSD(T), and other noniterative corrections to CCSD is the fact that, in analogy to CCSD(T), they decouple the higher-order T_n components with \( n > m_A \), such as \( T_3 \) or \( T_4 \), from their lower-order \( n \leq m_A \) (e.g., \( T_1 \) and \( T_2 \)) counterparts. This can result in substantial errors, for example when the activation energies and chemical reaction profiles involving rearrangements of \( \pi \) bonds and singlet–triplet gaps in certain classes of biradical species are examined. The automerization of cyclobutadiene, which is one of the benchmark examples in Section 111, provides an illustration of the challenges the noniterative corrections to CCSD, including CCSD(2)_T and CR-CC(2,3), face when the coupling of the lower-order \( T_1 \) and \( T_2 \) and higher-order \( T_3 \) clusters becomes significant (see Ref. 61 for further analysis and additional remarks). One can address problems of this type by using active orbitals to incorporate the dominant higher–than–doubly excited determinants, in addition to all singles and doubles, in the space, as in the successful CC(t;3), CC(t,q;3), and CC(t,q;3,4) hierarchy, which uses the CC(P;Q) framework to correct the results of the active-space CCSD, or CSDT, calculations for the remaining \( T_3 \) or \( T_3 \) and \( T_4 \) correlations that were not captured via active orbitals, but the resulting methods are no longer computational black boxes. The semi-stochastic CC(P;Q) methodology, introduced in Ref. 58 and further developed in this work, which takes advantage of the FCIQMC or truncated CIQMC/CCMC propagations that can identify the leading higher–than–doubly-excited determinants for the inclusion in the space, while using the noniterative \( \delta(P;Q) \) corrections to capture the remaining correlations of interest, offers an automated way of performing CC(P;Q) computations without any reference to the user- and system-dependent active orbitals. The semi-stochastic CC(P;Q) methods developed and tested in this study are discussed next.

### B. Semi-stochastic CC(P;Q) Approaches Using FCIQMC and its Truncated CISDT-MC and CISDTQ-MC Counterparts

In our original examination of the semi-stochastic CC(P;Q) framework, and its recent extension to excited states, where we focused on converging the full CCSDT and EOMCCSDT energetics, we demonstrated that the FCIQMC and CCSDT-MC approaches are capable of generating meaningful \( P \) spaces for the subsequent CC(P)/EOMCC(P) iterations, which preceded the determination of the \( \delta(P;Q) \) moment corrections, already in the early stages of the respective QMC propagations. The main objective of this work is to explore if the same remains true when FCIQMC is replaced by its less expensive truncated CISDT-MC and CISDTQ-MC counterparts, in which spawning beyond the triply excited (CISDT-MC) or quadruply excited (CISDTQ-MC) determinants is disallowed, and if one can use the CIQMC-driven CC(P;Q) calculations to converge the higher-level CCSDT energetics with similar efficiency.

The key steps of the semi-stochastic CC(P;Q) algorithm exploited in this study, which allows us to converge the CCSDT and CCSDTQ energetics using the \( P \) spaces extracted from the FCIQMC and truncated CISDT-MC and CISDTQ-MC propagations, are as follows:

1. Initiate a CIQMC run appropriate for the CC method of interest by placing a certain number of walkers on the reference state \( |\Phi\rangle \), which in all of the calculations reported in this article is the restricted Hartree-Fock (RHF) determinant. Among the CIQMC schemes that can provide meaningful \( P \) spaces for the CC(P;Q) calculations targeting the CCSDT energetics are the FCIQMC approach used in our earlier work and the CISDT-MC and CISDTQ-MC methods examined in the present study. If the objective is to converge the CCSDTQ energetics, one can use FCIQMC or CISDTQ-MC, which are the two choices pursued in the present work, but not CISDT-MC, which ignores quadruply excited determinants. As in our earlier work, we adopt the initiator CIQMC (i-CIQMC) algorithm, originally proposed in Ref. 63, based on integer walker numbers.

2. After a certain number of CIQMC time steps, called MC iterations, i.e., after some QMC propagation time \( \tau \), extract a list of higher–than–doubly excited determinants relevant to the CC theory of interest to construct the \( P \) space for executing the CC(P) calculations. If one is interested in targeting the CCSDT-level energetics, the \( P \) space used in the CC(P) iterations consists of all singly and doubly excited determinants and a subset of triply excited determinants identified by the underlying FCIQMC, CISDT-MC, or CISDTQ-MC propagation, where each triply excited determinant in the subset is populated by at least \( n_P \) positive or negative walkers. In analogy to our previous studies, all of the CC(P) and CC(P;Q) calculations carried out in this work use \( n_P = 1 \). If the goal is to converge the CCSDTQ energetics, the \( P \) space for the CC(P) computations is defined as all singly and doubly excited determinants and a subset of triply and quadruply excited determinants identified by the underlying FCIQMC or CISDTQ-MC propaga-
3. Solve the CC($P$) and left-eigenstate CC($P$) equations, Eqs. (11) and (12), respectively, where $E^{(P)}$ is given by Eq. (7), for the cluster operator $T^{(P)}$ and the deexcitation operator $\Lambda^{(P)}$ in the $P$ space determined in step 2. If the objective is to converge the CC$^3$P and the deexcitation operator $\Lambda$ determined in step 2. If the objective is to increase the number of triples identified by the FCIQMC or CISDTQ-MC propagation at time $\tau$, respectively, defined using the lists of triples and quadruples, and the four-body components of $T^{(MC)}$ and $\Lambda^{(MC)}$, are the three-body and four-body components of $T^{(P)}$ and $\Lambda^{(P)}$, respectively, defined using the lists of triples and quadruples identified by the FCIQMC or CISDTQ-MC propagation at time $\tau$. 

4. Use the CC($P;Q$) correction $\delta(P;Q)$, Eq. (8), to correct the energy $E^{(P)}$ obtained in step 3 for the remaining correlation effects of interest, meaning those correlations that were not captured by the CC($P$) calculations performed at the time $\tau$ the list of higher–than–doubly excited determinants entering the relevant $P$ space was created. If the objective is to converge the CC$^3$P-level energetics, the $Q$ space entering the definition of $\delta(P;Q)$ consists of those triply excited determinants that in the FCIQMC, CISDT-MC, or CISDTQ-MC propagation at time $\tau$ are populated by less than $n_P$ positive or negative walkers (in this study, where $n_P = 1$, the triply excited determinants that were not captured by the FCIQMC, CISDT-MC, or CISDTQ-MC propagation at time $\tau$). If the goal is to recover the CCSDTQ-level energetics, the $Q$ space used to calculate $\delta(P;Q)$ consists of the triply and quadruply excited determinants that in the FCIQMC or CISDTQ-MC propagation at time $\tau$ are populated by less than $n_P$ positive or negative walkers.

5. Check the convergence of the CC($P;Q$) energy $E^{(P+Q)}$, Eq. (13), obtained in step 4, by repeating steps 2–4 at some later CIQMC propagation time $\tau' > \tau$. If the resulting energy $E^{(P+Q)}$ no longer changes within a given convergence threshold, the CC($P;Q$) calculation can be stopped. As pointed out in Refs. 58, 60, one can also stop it once the fraction (fractions) of higher–than–doubly excited determinants captured by the CIQMC propagation relevant to the target CC theory level, included in the $P$ space, is (are) sufficiently large to obtain the desired accuracy. This is further discussed in Section III, where the numerical results obtained in this study are presented.

The above semi-stochastic CC($P;Q$) algorithm, allowing us to recover the CCSDT and CCSDTQ energetics using the $P$ spaces identified with the help of FCIQMC or truncated CISDT-MC and CISDTQ-MC propagations, has been implemented by modifying our previously developed standalone deterministic CC($P;Q$) codes, which rely on the RHF, restricted open-shell Hartree-Fock, and integral routines in the GAMESS package, such that they could handle the stochastically determined lists of triples and quadruples, and by interfacing the resulting program with the $i$-CIQMC routines available in the HANDE software. As in our earlier semi-stochastic CC($P$)/EOMCC($P$) and CC($P;Q$) work, we rely on the original form of the initiator CIQMC ($i$-CIQMC) algorithm proposed in Ref. 63, where only those determinants that acquire walker population exceeding a preset value $n_a$ are allowed to attempt spawning new walkers onto empty determinants, but one could consider interfacing our CC($P;Q$) framework with the improved ways of converging CIQMC, such as the adaptive-shift method developed in Refs. 67 and 114. We will consider such an interface in the future.

In the case of the semi-stochastic CC($P;Q$) codes aimed at converging the CCSDT energetics, which we have extended in the present study by allowing them to work with the CISDT-MC and CISDTQ-MC approaches, in addition to the previously examined FCIQMC and CISDTQ-MC options, we follow the algorithm summarized in steps 1–5 without any alterations. In particular, all of the quantities entering Eq. (8) for the noniterative correction $\delta(P;Q)$ are treated in the present study fully. This is an improvement compared to our original semi-stochastic CC($P;Q$) computations utilizing FCIQMC and CCSDT-MC, reported in Ref. 58, where we adopted an approximation in which the three-body component $\Lambda_3^{(MC)}$ of the deexcitation operator $\Lambda^{(P)}$ was used to determine amplitudes $\ell_k(P)$ entering Eq. (8) was neglected. In analogy to this work, the similarity-transformed Hamiltonian $\hat{H}^{(P)}$, defining moments $\mathfrak{M}_k(P)$ and entering the linear system defined by Eq. (12), which is used to determine $\Lambda^{(P)}$, was treated in Ref. 58 fully, i.e., $\hat{H}^{(P)}$ employed in the CC($P;Q$) calculations aimed at recovering the CCSDT energetics was defined as $\left(H^{(P)} + T_1^{(MC)} + T_2^{(MC)}\right)_C$, so that the one- and two-body components of $\Lambda^{(P)}$ employed in Ref. 58 were properly relaxed in the presence of the three-body component $T_3^{(MC)}$ of the cluster operator $T^{(P)}$ obtained in the preceding CC($P$) calculations, but $\Lambda_3^{(MC)}$ was neglected. Although all of our numerical tests to date indicate that this approximation has a small effect on the results of the semi-stochastic CC($P;Q$) calculations utilizing full and truncated CIQMC and no effect on our main conclusions, we no longer use it in this work. In other words, all of the calculations reported in the present study rely
on the complete representations of \( \tilde{H}(P) \) and \( \Lambda(P) \) when constructing moments \( \mathcal{M}_k(P) \) and amplitudes \( \ell_k(P) \) entering Eq. (3). This means that \( \tilde{H}(P) \) and \( \Lambda(P) \) used to determine the CC(P,Q) correction \( \delta(P;Q) \) in the calculations aimed at the CCSDTQ energetics are defined as 
\( \langle He^{T_1+T_2+T_3^{(MC)}} \rangle_C \) and \( \Lambda_1 + \Lambda_2 + \Lambda_3^{(MC)} \), respectively.

We have, however, introduced an approximation in the semi-stochastic C(C(P,Q)) routines that are used to converge the CCSDTQ-level energetics. Given the pilot nature of these routines, the noniterative correction \( \delta(P;Q) \) that they produce corrects the \( E(P) \) energy, which is obtained in this case by solving the CC(P) equations in the space of all singles and doubles and subsets of triples and quadruples captured by FCIQMC or CISDCT-MC, for the remaining triples not included in the P space, but the quadruples contributions to \( \delta(P;Q) \) are ignored. This approximation is acceptable, since in the \( \tau = \infty \) limit, where the \( P \) space contains all triples and quadruples, i.e., the corresponding Q space is empty, the uncorrected CC(P) and partially or fully corrected CC(P,Q) calculations recover the CCSDTQ energetics. All of our tests to date, including those discussed in Section III indicate that the convergence of the CC(P,Q) computations, in which the quadruples component of \( \delta(P;Q) \) is ignored, toward CCSDTQ is rapid, even when the \( T_4 \) effects become significant, so the above approximation does not seem to have a major effect on the convergence rate, but we will implement the full correction \( \delta(P;Q) \) due to the missing triples as well as quadruples in the future to examine if one can accelerate convergence toward CCSDTQ even further.

As explained in Refs. 58 and 60 (cf., also, Ref. 59), the semi-stochastic CC(P,Q) approaches of the type summarized above offer a number of advantages. Among them are substantial savings in the computational effort compared to the parent high-level CC theories they target and a systematic behavior of the resulting \( E(P+Q) \) energies as \( \tau \) approaches \( \infty \). The latter feature is a direct consequence of the fact that if we follow the definitions of the \( P \) and \( Q \) spaces introduced in steps 2 and 4 above, the initial, \( \tau = 0 \), CC(P,Q) energies are identical to those obtained with CR-CC(2,3) or CR-CC(2,4), which are approximations to CCSDT and CCSDTQ, respectively, that account for some \( T_3 \) (CR-CC(2,3)) or \( T_3 \) and \( T_4 \) (CR-CC(2,4)) correlations. In the \( \tau = \infty \) limit, the CC(P,Q) energies \( E(P+Q) \) become equivalent to their respective high-level CC parents, which account for the \( T_n \) components with \( n > 2 \), such as \( T_3 \) or \( T_3 \) and \( T_4 \) fully, so that the QMC propagation time \( \tau \) becomes a parameter connecting CR-CC(2,3) with CCSDT and CR-CC(2,4) with CCSDTQ. In the case of our current implementation of the semi-stochastic CC(P,Q) approach aimed at converging the CCSDTQ energetics, where the quadruples contributions to correction \( \delta(P;Q) \) are ignored, the initial, \( \tau = 0 \), CC(P,Q) energy is equivalent to that obtained with the CR-CC(2,3) approach, i.e., the QMC propagation time \( \tau \) connects CR-CC(2,3) with CCSDTQ. When \( \tau \) approaches \( \infty \), the uncorrected CC(P) energies \( E(P) \) converge to their CCSDT and CCSDTQ parents as well, but the convergence toward CCSDT and CCSDTQ is in this case slower, since the CC(P) energies at \( \tau = 0 \) are equivalent to those of CCSD, which has no information about the \( T_n \) components with \( n > 2 \), and, as shown in our earlier work, and as clearly demonstrated in the present study, the CC(P,Q) corrections \( \delta(P;Q) \) greatly accelerate the convergence toward the target CC energetics. The above relationships between the semi-stochastic CC(P) and CC(P,Q) approaches and the deterministic CCSD, CR-CC(2,3)/CR-CC(2,4), and CCSDT/CCSDTQ theories are also helpful when debugging the CC(P) and CC(P,Q) codes.

As far as the savings in the computational effort offered by the semi-stochastic CC(P,Q) methods, when compared to their high-level CC parents, such as CCSDT or CCSDTQ, are concerned, they were already discussed in Refs. 58 and 60, so here we focus on the information relevant to the calculations discussed in Section III. There are three main factors that contribute to these savings. First, the computational times associated with the early stages of the CIQMC walker propagations, which are sufficient to recover the parent CCSDT or CCSDTQ energetics to within small fractions of a millihartree when the semi-stochastic CC(P,Q) framework is employed, are very short compared to the converged CIQMC runs. They are already short when one uses FCIQMC, and they are even shorter when one replaces FCIQMC by the CISDT-MC and CISDCT-MC truncations.

Second, the CC(P) calculations using small fractions of higher-than-doubly excited determinants, which is how the \( P \) spaces used in these calculations look like when the early stages of the CIQMC walker propagations are considered, are much faster than the parent CC computations. For example, when the most expensive \( \langle \Phi_{ijk}^{abc}[H,T_3]\Phi \rangle \) or \( \langle \Phi_{ijk}^{abc}[\tilde{H}^{(2)},T_3]\Phi \rangle \) contributions to the CCSDT equations, where \( \tilde{H}^{(2)} = e^{-T_1-T_2}He^{T_1+T_2} \), are isolated and implemented using programming methods similar to those exploited in selected CI algorithms (rather than the usual diagrammatic techniques that assume continuous excitation manifolds labeled by all occupied and all unoccupied orbitals), one can accelerate 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 triplets included in the \( P \) space, captured with the help of CIQMC propagations. Other contributions to the CCSDT equations that involve \( T_3 \) or the projections on the triply excited determinants, such as \( \langle \Phi_{ijl}^{abc}[H,T_3]\Phi \rangle \) and \( \langle \Phi_{ijl}^{abc}[H,T_2]\Phi \rangle \), may offer additional speedups, on the order of \( (D/d) \). Our current CC(P) codes are still in the pilot stages, but the speedups on the order of \( (D/d) \) in the determination of the most expensive \( \langle \Phi_{ijl}^{abc} [H,T_3] \Phi \rangle \) (or \( \langle \Phi_{ijl}^{abc} [\tilde{H}^{(2)},T_3] \Phi \rangle \)) terms are attainable. Similar remarks apply to the CC(P)/CC(P,Q) calculations aimed at converging the CCSDTQ energetics, where one can considerably speed up the determination of the most expensive \( \langle \Phi_{ijkl}^{abcd} [H,T_4] \Phi \rangle \) or \( \langle \Phi_{ijkl}^{abcd} [\tilde{H}^{(2)},T_4] \Phi \rangle \) contri-
butions and other terms containing the $T_3$ and $T_4$ clusters and the projections on the triply and quadruply excited determinants. It should also be noted that the CC($P$) calculations do not require storing the entire $T_3$ and $T_4$ vectors. The $T_3^{\text{(MC)}}$ and $T_4^{\text{(MC)}}$ operators use much smaller numbers of amplitudes than their full $T_3$ and $T_4$ counterparts.

Third, the computation of the noniterative correction $\delta(P;Q)$ is much less expensive than a single iteration of the target CC calculation. In the case of the CC($P;Q$) calculations aimed at converging the CCSDT energetics, the computational time required to determine the corresponding correction $\delta(P;Q)$ scales no worse than $\sim 2n_3^4n_4^4$, which is much less than the $n_3^5n_4^5$ scaling of each iteration of CCSDT. In the case of the CC($P;Q$) approach aimed at CCSDTQ, the computational time required to determine correction $\delta(P;Q)$ scales as $\sim 2n_3^3n_4^4$ in the case of the contributions due to the remaining triples and is on the order of $n_3^5n_4^5$ in the case of the quadruples part of $\delta(P;Q)$, when the more complete CC($P;Q$)$_\text{EN}$ approach is used, or $n_3^2n_4^5$, when the CC($P;Q$)$_\text{MP}$ form of $\delta(P;Q)$ is employed. This is all much less than the $n_3^5n_4^5$ scaling of every CCSDTQ iteration. As mentioned above, in our current implementation of the semi-stochastic CC($P;Q$) approach aimed at converging the CCSDTQ energetics, the quadruples contribution to correction $\delta(P;Q)$ is neglected, so the computational time required to obtain $\delta(P;Q)$ scales as $\sim 2n_3^3n_4^4$, at worst, which points to the usefulness of such an approximation, especially that the convergence of the resulting CC($P;Q$) energies toward CCSDTQ, as shown in Section III, is very fast.

### III. NUMERICAL EXAMPLES

In order to demonstrate the benefits offered by the semi-stochastic CC($P;Q$) framework, especially the new CC($P;Q$) approaches implemented in this work that replace FCIQMC by the less expensive CISDT-MC and CISDTQ-MC propagations, we applied the FCIQMC-, CISDT-MC-, and CISDTQ-MC-driven CC($P;Q$) methods aimed at converging the CCSDT and CCSDTQ energetics to a few molecular problems, for which the parent full CCSDT and CCSDTQ results had previously been determined or were not too difficult to be recalculated. Thus, we carried out an extensive series of the CISDT-MC- and CISDTQ-MC-driven CC($P;Q$) calculations, along with the analogous computations using FCIQMC, which was utilized in our earlier study to examine the ability of the semi-stochastic CC($P;Q$) approaches using various types of CIQM to recover the CCSDT energetics for the F–F bond dissociation in the fluorine molecule (Section III A) and the automerization of cyclobutadiene (Section III B). In order to illustrate the performance of the FCIQMC- and CISDTQ-MC-driven CC($P;Q$) methods in calculations aimed at converging the CCSDTQ energetics, we considered the symmetric stretching of the O–H bonds in the water molecule (Section III C). We chose bond breaking in F$_2$, which is accurately described by full CCSDT, and CCSDT-MC-driven CC($P;Q$) work, as well as the preceding deterministic CC($P;Q$)-based CC(t;3) calculations reported in Ref. 53. Our choice of the automerization of cyclobutadiene, which is accurately described by CCSDT as well, was motivated by similar reasons. We studied this problem, where all noniterative triples corrections to CCSD, including CCSD(T), Δ-CCSD(T), CCSD(2)$_\text{T}$, and CR-CC(2,3) fail, using the deterministic CC(t;3) approach exploiting the CC($P;Q$) ideas in Ref. 61, and we studied it again using the semi-stochastic CC($P;Q$) framework utilizing FCIQMC and CCSDT-MC in Ref. 53. We would like to explore now what the effect of replacing FCIQMC propagations by their less expensive CISDT-MC and CISDTQ-MC counterparts on the convergence of the CC($P;Q$) energies toward CCSDT is. We would also like to learn if the incorporation of the previously neglected three-body component of the deexcitation operator $\Lambda^{(P)}$, which is used to construct amplitudes $\ell K(P)$ entering Eq. 5, helps the accuracy of the resulting semi-stochastic CC($P;Q$) energies. We studied the $C_2v$-symmetric double dissociation of H$_2$O, since by simultaneously stretching both O–H bonds by factors exceeding 2, one ends up with a catastrophic failure of CCSDT 61,116. One needs an accurate description of the $T_3$ and $T_4$ clusters to obtain a more reliable description of the water potential energy surface in that region.

Following our earlier semi-stochastic and deterministic CC($P;Q$) work, which also provides the parent CCSDT 57,58,61,63 and CCSDTQ 63 energetics, we used the cc-pVDZ, 119 cc-pVTZ, 119 and aug-cc-pVTZ 120 basis sets for F$_2$ and the cc-pVDZ bases for cyclobutadiene and water. For consistency with Refs. 57, 58, and 61, in all of the post-RHF computations for the F–F bond breaking in F$_2$ and the automerization of cyclobutadiene, the core electrons corresponding to the 1s shells of the fluorine and carbon atoms were kept frozen. As in Refs. 63 and 118, which provide the reference CCSDTQ data and, in the case of Ref. 118, the geometries of the equilibrium and stretched water molecule used in our semi-stochastic CC($P;Q$) calculations aimed at converging the CCSDTQ energetics, we correlated all electrons. Each of the relevant i-FCIQMC (all systems), i-CISDT-MC ($F_2$ and cyclobutadiene), and i-CISDTQ-MC (all systems) runs was initiated by placing 100 walkers on the RHF reference determinant and we set the initiator parameter $n_i$ at 3. All of the i-FCIQMC, i-CISDT-MC, and i-CISDTQ-MC propagations used the time step $\delta\tau$ of 0.0001 a.u.

#### A. Bond Breaking in F$_2$

We begin our discussion of the semi-stochastic CC($P;Q$) calculations carried out in this study with the
F–F bond dissociation in the fluorine molecule, as described by the cc-pVDZ basis set using the Cartesian components of $d$ orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Figs. 1–3). In analogy to Ref. 58, where our initial FCIQMC components of the cc-pVDZ basis set using the Cartesian orbitals (see Table I and Fig...
the convergence of the underlying CC\((P)\) energies, allowing one to recover the parent CCSDT energies to within \(\sim 0.1\) millihartree once about 40\% \((R = R_e\\text{ and } 1.5R_e)\) or 15–20\% \((R = 2R_e,\) and 5\(R_e\)) of the triples are captured by the FCIQMC, CISDT-MC, and CISDTQ-MC propagations.

The results shown in Table II and Figs. 1–3 demonstrate that it is practically irrelevant whether one uses FCIQMC or one of its less expensive truncated forms, such as CISDT-MC and CISDTQ-MC examined in this study, to identify the leading triply excited determinants for the inclusion in the \(P\) space used in the CC\((P;Q)\) and the underlying CC\((P)\) calculations. Clearly, as \(\tau\) approaches \(\infty\), the FCIQMC, CISDT-MC, and CISDTQ-MC propagations converge to completely different limits (FCI in the case of FCIQMC, CISDT-MC in the case of CISDT-MC, and CISDTQ in the case of CISDTQ-MC), but this has virtually no impact on the convergence patterns observed in our semi-stochastic CC\((P)\) and CC\((P;Q)\) calculations. This is a consequence of the fact that the uncorrected CC\((P)\) and corrected CC\((P;Q)\) computations are capable of recovering the parent high-level CC energetics, such as those corresponding to full CCSDT discussed in this subsection, based on the information extracted from the early stages of the corresponding CIQMC runs. In particular, if we are targeting CCSDT, all we need from the CIQMC calculations is a meaningful list of the leading triply excited determinants, which any CIQMC calculation that is allowed to sample the triples subspace of the Hilbert space, even the crude CISDT-MC approach, can provide. One can see, for example, in Table II that the fractions of triples captured by the FCIQMC, CISDT-MC, and CISDTQ-MC runs at the various numbers of MC iterations (various propagation times \(\tau\)) are very similar. Detailed inspection of the corresponding lists of triply excited determinants shows that while the numbers of walkers on the individual determinants may substantially differ, the lists of triples identified by the FCIQMC, CISDT-MC, and CISDTQ-MC propagations, especially the more important ones that result in larger \(T_{\text{MC}}^3\) amplitudes in the subsequent deterministic CC\((P)\) steps, are not much different. Once the lists of the leading triples are identified, we turn to the CC\((P)\) computations, correcting them for the remaining triples not captured by CIQMC, and this makes the semi-stochastic CC\((P)\) and CC\((P;Q)\) calculations rather insensitive to the type of the CIQMC approach used to construct these lists.

All of the above observations regarding the ability of the semi-stochastic CC\((P;Q)\) calculations using the FCIQMC, CISDT-MC, and CISDTQ-MC propagations to rapidly converge the full CCSDT energetics remain true when the cc-pVDZ basis set is replaced by its larger cc-pVTZ and aug-cc-pVTZ counterparts (both using the spherical components of \(d\) and \(f\) functions). This is illustrated in Table III where we examine the stretched \(F_2\) molecule, in which the \(F\–F\) distance \(R\) is set at \(2R_e\). We chose \(R = 2R_e\), since, in analogy to the previously discussed cc-pVDZ basis set, the \(T_3\) effects at this geometry, obtained by calculating differences of the respective CCSDT and CCSD energies, which are \(-62.819\) millihartree, when the cc-pVTZ basis set is employed, and \(-65.036\) millihartree, when the aug-cc-pVTZ basis is used, are not only very large, but also larger, in absolute value, than the corresponding CCSDT dissociation energies (differences between the CCSDT energies at \(R = 5R_e\), where the \(F\–F\) bond is broken, and \(R = R_e\) obtained with the cc-pVTZ and aug-cc-pVTZ basis sets are about 57 and 60 millihartree, respectively). We also chose it, since the \(R = 2R_e\) stretch of the \(F\–F\) bond length is large enough for the conventional CCSD\((T)\) approach to fail in a major way when the cc-pVTZ and aug-cc-pVTZ basis sets are employed, resulting in the \(-26.354\) and \(-27.209\) millihartree errors relative to CCSDT, respectively. The CCSD\((2)\) correction to CCSD or the equivalent CR-CC\((2,3)\)\(_A\) approximation, represented in Table II by the \(\tau = 0\) CC\((P;Q)_{\text{MP}}\) results, helps, but large differences between the CCSD\((2)\)\(_T\) and CCSDT energies, of 9.211 millihartree in the cc-pVTZ case and 9.808 millihartree when the aug-cc-pVTZ basis set is employed, remain. The CR-CC\((2,3)\)\(_D\) approach, represented in Table II by the \(\tau = 0\) CC\((P;Q)_{\text{EN}}\) data, is more effective than other triples corrections to CCSD, reducing the large errors relative to CCSDT observed in the CCSD\((T)\) and CCSD\((2)\)\(_T\) calculations to 4.254 (cc-pVTZ) and 5.595 (aug-cc-pVTZ) millihartree, but none of the above results are as good as the energies resulting from the semi-stochastic CC\((P;Q)\) calculations using FCIQMC, CISDT-MC, and CISDTQ-MC.

Indeed, as shown in Table III we observe a rapid error reduction relative to the parent CCSDT data once we start migrating the triply excited determinants identified during the FCIQMC, CISDT-MC, and CISDTQ-MC propagations into the underlying \(P\) space. With about 20–30\% (cc-pVTZ) or 30–40\% (aug-cc-pVTZ) of the triples in the \(P\) space, the 62.819 and 65.036 millihartree errors resulting from the initial CCSD (\(\tau = 0\)) CC\((P)\) computations decrease to a 1–2 millihartree level when the CC\((P)\) method is employed. The CC\((P;Q)\) corrections due to the remaining triples not captured by FCIQMC, CISDT-MC, and CISDTQ-MC accelerate the convergence toward CCSDT even further, with the semi-stochastic CC\((P;Q)_{\text{EN}}\) approach being particularly efficient in this regard. With only 2–4\% of the triples in the stochastically determined \(P\) spaces, captured after 20,000–30,000 \(\delta\tau = 0.0001\) a.u. MC iterations, which are the very early stages of the FCIQMC, CISDT-MC, and CISDTQ-MC propagations, the CC\((P;Q)_{\text{EN}}\) calculations recover the full CCSDT energetics corresponding to the cc-pVTZ and aug-cc-pVTZ basis sets to within 0.1–0.2 millihartree. After 50,000 (cc-pVTZ) or 60,000 (aug-cc-pVTZ) MC iterations, where the FCIQMC, CISDT-MC, and CISDTQ-MC runs are still far from convergence, capturing only about 20–30\% (cc-pVTZ) or 30–40\% (aug-cc-pVTZ) of the triples, the errors in the CC\((P;Q)_{\text{EN}}\) energies relative to CCSDT reduce to a 10
microhartree level. In analogy to the cc-pVDZ basis set, the CC\((P;Q)\)\textsubscript{MP} correction is less accurate than its CC\((P;Q)\)\textsubscript{EN} counterpart when the cc-pVTZ and aug-cc-pVTZ basis sets are employed, recovering the CCSDT energies to within 0.1–0.2 millihartree after 50,000 rather than 20,000–30,000 MC iterations, i.e., after about 20–30 % rather than 2–4 % of the triples are captured by the CIQMC propagations, but the overall error reduction compared to the underlying CC\((P)\) calculations or the various noniterative triples corrections to CCSD is still impressive.

Similarly to the cc-pVDZ basis set, the semi-stochastic CC\((P;Q)\) calculations using larger cc-pVTZ and aug-cc-pVTZ bases are rather insensitive to the type of the CIQMC approach used to identify the leading triples for the inclusion in the \(P\) space. Based on the results in Table \(\text{II}\) one might try to argue that the energies obtained with the uncorrected CC\((P)\) approach using the CISDT-MC propagations are characterized by slower convergence compared to their CISDTQ-MC- and FCIQMC-driven counterparts, but this would be misleading, since CISDT-MC captures the leading triples at a somewhat slower rate, while being less expensive than CISDTQ-MC and FCIQMC at the same time. For example, the CISDT-MC-driven CC\((P)\) computations for \(F_2\) at \(R = 2R_e\) using the cc-pVTZ basis set need 60,000 \(\delta\tau = 0.0001\) a.u. MC iterations to reach a \(\sim\)1 millihartree accuracy relative to the corresponding CCSDT energy. The CC\((P)\) approach using CISDTQ-MC and FCIQMC reaches the same accuracy level sooner, after 50,000 MC iterations. One should keep in mind, however, that it takes 60,000 MC time steps for the CISDT-MC propagation to capture about 30 % of the triples, needed to reach a \(\sim\)1 millihartree accuracy level in the subsequent CC\((P)\) calculations, and the analogous CISDTQ-MC and FCIQMC runs capture a similar fraction of the triples after 50,000 time steps. Ultimately, one needs to remember that all CIQMC-driven CC\((P)\) computations considered in this subsection converge to CCSDT as \(\tau \to \infty\), independent of the type of the CIQMC approach used to define the underlying \(P\) spaces, as long as the CIQMC propagation is allowed to spawn walkers on the triply excited determinants. Perhaps more importantly, the CC\((P;Q)\) corrections to the CC\((P)\) energies make the convergence toward CCSDT not only much faster, but also less dependent on the type of the CIQMC approach used in the calculations, since they take care of the triples that were not captured by the respective QMC propagations.

Before discussing our next molecular example, it is worth pointing out that the FCIQMC-driven CC\((P;Q)\) calculations reported in Tables \(\text{II}\) and \(\text{III}\) and Fig. \(\text{I}\) in which, as explained in Section \(\text{IIIB}\) we used complete representations of \(\bar{H}(P)\) and \(\Lambda(P)\) in determining corrections \(\delta(P;Q)\), approach the parent CCSDT energetic of the stretched \(F_2\) system in the early stages of the underlying FCIQMC propagations faster than the analogous calculations reported in Ref. \(\text{58}\), where the three-body component of \(\Lambda(P)\) was neglected. For example, the CC\((P;Q)\) energies of \(F_2\) at \(R = 2R_e\) using the aug-cc-pVTZ basis set obtained in this work after 10,000, 20,000, and 30,000 \(\delta\tau = 0.0001\) a.u. MC iterations of the underlying FCIQMC propagation differ from the corresponding CCSDT energy by 1.594, 0.382, and 0.138 millihartree, respectively (see Table \(\text{III}\)). The analogous energy differences reported in Ref. \(\text{58}\) of 3.770, 1.661, and 0.454 millihartree, respectively, are noticeably larger (see Table \(\text{II}\) in the Supplemental Material to Ref. \(\text{58}\)). In fact, by comparing the FCIQMC-, CISDT-MC-, and CISDTQ-MC-based CC\((P;Q)\) energies shown in Tables \(\text{III}\) and \(\text{IV}\) and Figs. \(\text{I}\) and \(\text{IV}\) determined by treating the deexcitation operator \(\Lambda(P)\) in Eq. \(\text{9}\) fully, i.e., by defining \(\Lambda(P)\) as \(\Lambda_1 + \Lambda_2 + \Lambda_3^{\text{(MC)}}\), with their FCIQMC- and CCSDT-MC-based counterparts obtained in Ref. \(\text{58}\), where \(\Lambda(P)\) was approximated by \(\Lambda_1 + \Lambda_2\), we can conclude that as long as \(\Lambda_3^{\text{(MC)}}\) is not neglected one can replace FCIQMC by CISDTQ-MC or, even, CISDT-MC and still improve the rate of convergence of the CC\((P;Q)\) energies toward CCSDT in the early stages of the QMC propagations compared to that reported in Ref. \(\text{58}\).

The above observations, combined with the superior performance of the CC\((P;Q)\)\textsubscript{EN} approach compared to its CC\((P;Q)\)\textsubscript{MP} counterpart, suggest that a complete treatment of correction \(\delta(P;Q)\), as dictated by Eqs. \(\text{8}\), \(\text{9}\), and \(\text{10}\), is more important, especially when one is interested in accelerating convergence of the semi-stochastic CC\((P;Q)\) calculations for stretched or more multireference molecules in the early stages of the QMC propagations, than the actual type of the underlying CIQMC approach. It is interesting to examine if the same remains true when other molecular examples, including those discussed in the next two subsections, are considered.

### B. Automerization of Cyclobutadiene

Our next example is the challenging and frequently studied\(61,116,117,124–141\) automerization of cyclobutadiene (see Fig. \(\text{I}\)). In this case, in order to obtain reliable energetics using computational means, especially the activation energy, one has to provide an accurate and well-balanced description of the nondegenerate closed-shell reactant (or the equivalent product) species, in which the many-electron correlation effects have a predominantly dynamical character, and the quasi-degenerate, biradicaloid transition state characterized by substantial non-dynamical correlations. Experiment suggests that the activation energy for the automerization of cyclobutadiene is somewhere between 1.6 and 10 kcal/mol\(124,126\). The most accurate single- and multi-reference calculations performed to date, reviewed, for example, in Refs. \(\text{61}\), \(\text{117}\), and \(\text{140}\), imply that the purely electronic value of the energy barrier falls into the 6–10 kcal/mol range. In particular, as pointed out in Ref. \(\text{61}\) (cf., also Ref. \(\text{116}\)), one can obtain a reliable description of the activation energy using the full CCSDT approach. Given this information and the methodolog-
ical nature of the present study, in which we had to perform a large number of semi-stochastic CC(\(P\)) and CC(\(P;Q\)) calculations, exploring three different types of the CIQMC method, including FCIQMC, CISDT-MC, and CISD\(TQ\)-MC, and probing many values of the QMC propagation time \(\tau\), in a discussion below we focus on converging the CCSDT energetics obtained using the spherical cc-pVDZ basis set. As shown in Ref. \[61\] and Table \[11\] the CCSDT/cc-pVDZ activation energy characterizing the autoinertization of cyclobutadiene, assuming the reactant/product and transition-state geometries obtained with the multireference average-quadratic CC (MR-AQCC) approach\[\text{12}\text{13}\] in Ref. \[13\], which we adopt in the CC(\(P\)) and CC(\(P;Q\)) calculations reported in this work as well, is 7.627 kcal/mol, in reasonable agreement with the most accurate ab initio results reported to date. The results of our semi-stochastic CC(\(P\)) and CC(\(P;Q\)) calculations, aimed at recovering the CCSDT/cc-pVDZ energetics of the reactant and transition-state species and the corresponding activation energy using the FCIQMC, CISDT-MC, and CISD\(TQ\)-MC propagations to identify the leading triply excited determinants for constructing the underlying \(P\) spaces, are summarized in Table \[11\] and Fig. \[\text{e}\].

As already mentioned, all of the noniterative triples corrections to CCSD, including CCSD(\(T\)), \(\Lambda\)-CCSD(\(T\)), CCSD(\(2\))\(_{\tau}\), and CR-CC(\(2;3\)), perform very poorly in this case, producing activation barriers in a 16–17 kcal/mol range when the cc-pVDZ basis set is considered\[61\text{17}\] instead of \(\sim\)8 kcal/mol obtained with CCSDT (it should be noted that the 16–17 kcal/mol values are also outside the experimentally derived and most accurate theoretically determined ranges of 1.6–10 kcal/mol and 6–10 kcal/mol, respectively). They improve the CCSD activation energy, which is even worse (about 21 kcal/mol; see the \(\tau = 0\) CC(\(P\)) barrier in Table \[11\], but the improvements offered by the noniterative triples corrections to CCSD are far from sufficient. This, in particular, applies to the CCSD(\(2\))\(_{\tau}\) = CR-CC(\(2;3\))\(_{\Lambda}\) and CR-CC(\(2;3\))\(_{\text{D}}\) approaches, represented in Table \[11\] by the \(\tau = 0\) CC(\(P;Q\))\(_{\text{MP}}\) and CC(\(P;Q\))\(_{\text{EN}}\) data, respectively, where errors in the resulting activation energies relative to CCSDT are 9.611 kcal/mol (126 \%) in the former case and 8.653 kcal/mol (113 \%) in the case of the latter method. As explained in Ref. \[61\], the poor performance of the noniterative triples corrections to CCSD in describing the autoinertization of cyclobutadiene is a consequence of neglecting the coupling between the \(T_3\) clusters and their lower-order \(T_1\) and \(T_2\) counterparts, which is accounted for in CCSDT, but ignored in methods such as CCSD(T), \(\Lambda\)-CCSD(T), CCSD(\(2\))\(_{\tau}\), and CR-CC(\(2;3\)). This coupling is particularly large at the transition-state geometry, where the magnitude of \(T_3\) contributions, defined as the absolute value of the difference between the CCSDT and CCSD energies, is nearly 48 millihartree, when the cc-pVDZ basis set is employed, and where errors in the CCSD(T), \(\Lambda\)-CCSD(T), CCSD(\(2\))\(_{\tau}\), and CR-CC(\(2;3\)) energies relative to CCSDT range from about 14 to 20 millihartree, as opposed to \(\sim\)1–5 millihartree obtained for the reactant (see Refs. \[61\] and \[11\] and Table \[11\]). In analogy to bond breaking in \(F_2\), if we want to capture the coupling of the \(T_1\), \(T_2\), and \(T_3\) clusters without having to solve full CCSDT equations, while preserving the idea of noniterative triples corrections to energies obtained in lower-order CC calculations, we must solve for the \(T_1\) and \(T_2\) amplitudes, which in the CCSD(T), \(\Lambda\)-CCSD(T), CCSD(\(2\))\(_{\tau}\), and CR-CC(\(2;3\)) approaches are obtained with CCSD, in the presence of the dominant \(T_3\) components by incorporating some triples in the iterative CC steps, and then correct the resulting energies for the remaining \(T_3\) effects neglected in the CC iterations. Again, this can be done deterministically by solving the active-space CCSDt equations, in which the dominant \(T_3\) amplitudes are selected using active orbitals, and correcting the CCSDt energies for the remaining \(T_3\) correlations using the CC(\(P;Q\)) corrections \(\delta(\(P;Q\)\)) as in the CC(\(t;3\)) calculations reported in Ref. \[61\], or by turning to the semi-stochastic form of the CC(\(P;Q\)) formalism pursued in this study, which eliminates the need for defining active orbitals, when identifying the leading triples, by resorting to CIQMC propagations. Interestingly, using the CCSD(T)-type correction to CCSDt, as in the afore-mentioned CCSD(T)-h approach, in the calculations for the automerization of cyclobutadiene worsens the activation energies obtained with CCSDt, moving them away from their parent CCSDT values.\[61\] This underlines the significance of treating corrections \(\delta(\(P;Q\))\) due to the correlation effects outside the underlying \(P\) spaces as completely as possible, following Eqs. \[\text{8}\text{, 9}\text{, 10}\text{, and 11}\], avoiding drastic approximations in these equations that lead to the triples corrections of CCSD(T).

As shown in Table \[11\] and Fig. \[\text{e}\] the semi-stochastic CC(\(P;Q\)) calculations using FCIQMC, CISDT-MC, and CISD\(TQ\)-MC are remarkably efficient in capturing the desired \(T_3\) correlation effects. Independent of the type of the CIQMC approach, they allow us to converge the CCSDT values of the transition-state and activation energies, which are poorly described by the noniterative triples corrections to CCSD, to within 1–2 millihartree or 1–2 kcal/mol out of the early stages of the CIQMC propagations, while further improving an accurate description of the reactant by methods such as CR-CC(\(2;3\))\(_\text{D}\). Similarly to \(F_2\), the performance of the CC(\(P;Q\))\(_\text{EN}\) approach, which uses the Epstein–Nesbet form of the \(D_K(\(P\))\) denominator in calculating the \(\delta K(\(P\))\) amplitudes entering Eq. \[\text{8}\], is particularly impressive. With just 5–6 \% of the triples in the stochastically determined \(P\) spaces, captured by the FCIQMC, CISDT-MC, and CISD\(TQ\)-MC propagations after 30,000 \(\delta\tau = 0.0001\) a.u. MC iterations, i.e., almost instantaneously, the CC(\(P;Q\))\(_\text{EN}\) approach reduces the initial 0.848 millihartree, 14.636 millihartree, and 8.653 kcal/mol errors in the reactant, transition-state, and activation energies relative to CCSDT obtained in the \(\tau = 0\) CC(\(P;Q\))\(_\text{EN}\) or CR-CC(\(2;3\))\(_\text{D}\) calculations by factors of 2–4, to 0.228–0.279 millihartree, 3.648–6.651 millihartree, and 2.146–
such as 30,000–40,000 FCIQMC, CISDT-MC, and CISDTQ-MC calculations, especially if we realize that the early stages of the underlying CC(P;Q) energies, which use the Møller–Plesset (MP) correction, are 92 microhartree, 1.556 millihartree, and 0.919–2.011 kcal/mol, respectively. These are remarkable improvements compared to the initial CR-CC(2,3)D values, especially if we consider the early stages of the FCIQMC, CISDT-MC, and CISDTQ-MC calculations, such as 30,000–40,000 δτ = 0.0001 a.u. MC time steps, are all very fast and 5–6% or 12–16% are small fractions of the triples that result in large speedups in the underlying CC(P) calculations and significant reductions in the T3 amplitude storage requirements. After 50,000 MC iterations, where, as shown in Fig. 3, the FCIQMC, CISDT-MC, and CISDTQ-MC runs are still far from convergence, capturing about 20–30% of the triples, i.e., still relatively small fractions of all triply excited determinants, the CC(P;Q)EN calculations recover the CCSDT values of the reactant, transition-state, and activation energies to within 22–57 microhartree, 0.243–0.602 millihartree, and 0.138–0.343 kcal/mol, respectively, which is a massive error reduction compared to CR-CC(2,3)D and other noniterative triples corrections to CCSD. As in the case of bond breaking in F2, the CC(P;Q)MP correction, which uses the Møller–Plesset D3(P) denominator in Eq. 6 instead of its more elaborate Epstein–Nesbet form given by Eq. 11, is less accurate than its CC(P;Q)EN counterpart, but its ability to accelerate convergence of the underlying CC(P) energies and improving the results obtained with CR-CC(2,3) and other triples corrections to CCSD is still quite impressive. For example, with about 20–30% of the triples captured by the FCIQMC, CISDT-MC, and CISDTQ-MC propagations after 50,000 MC iterations, the differences between the CC(P;Q)MP reactant, transition-state, and activation energies and their CCSDT counterparts, of 0.877–1.235 millihartree, 1.488–2.238 millihartree, and 0.361–0.629 kcal/mol, are much smaller than the analogous errors relative to CCSDT resulting from the corresponding CC(P) calculations, which are 6.895–9.202 millihartree, 9.727–12.495 millihartree, and 1.601–2.067 kcal/mol, respectively, although they are not as small as the aforementioned 22–57 microhartree, 0.243–0.602 millihartree, and 0.138–0.343 kcal/mol errors obtained using the CC(P;Q)EN correction.

In analogy to the fluorine molecule, the semi-stochastic CC(P;Q) calculations aimed at converging the CCSDT results for the automerization of cyclobutadiene are generally insensitive to the type of the CIQMC approach used to identify the leading triples for the inclusion in the underlying P spaces. It is sufficient to resort to the least expensive forms of the CIQMC propagations capable of capturing the triples, such as CISDT-MC or CISDTQ-MC, to obtain the fast convergence of the CC(P;Q) reactant, transition-state, and activation energies toward their CCSDT parents observed in Table III and Fig. 3. Treating the CC(P;Q) correction δ(P;Q) fully, following Eqs. (5), (9), and (11), is, however, important. We have already discussed the benefits of using the Epstein–Nesbet form of the D3(P) denominator, Eq. (11), in determining the ℓK(P) amplitudes entering Eq. (8). A complete treatment of the deexcitation operator Λ(P) in Eq. (9), which in the case of the triples corrections to the CC(P) energies considered here means representing it as Λ1 + Λ2 + Λ3(MC), is important too. One can consider an approximation in which the three-body component Λ3(MC) is neglected, which is what we did in Ref. 58, but it is generally better, especially in the earlier stages of the CIQMC propagations, to keep all of the relevant many-body components of Λ(P) in calculating the ℓK(P) amplitudes that enter the CC(P;Q) correction δ(P;Q). This can be illustrated by comparing the results of the FCIQMC-driven CC(P;Q) computations shown in Table III where we used a complete representation of Λ(P), in which the three-body component Λ3(MC) was included, with the analogous results reported in Ref. 58, where Λ3(MC) was neglected. For example, the differences between the CC(P;Q)EN reactant, transition-state, and activation energies and their CCSDT counterparts obtained in this work after 40,000 δτ = 0.0001 a.u. time steps of the FCIQMC propagation are 92 microhartree, 1.556 millihartree, and 0.919 kcal/mol, respectively. The analogous energy differences reported in Ref. 58 are 0.489 millihartree, 3.235 millihartree, and 1.7 kcal/mol, respectively, i.e., they are substantially larger. Ultimately, when the propagation time τ becomes longer, different ways of handling the Λ(P) operator or different ways of defining the D3(P) denominator in Eq. (9) become less important, but if we are interested in accurately approximating the parent CC energetics in the early stages of the underlying CIQMC propagations, treating these quantities fully is essential.

As shown in this subsection and Section III A, using complete representations of the Λ(P) and H(P) operators and the Epstein–Nesbet-type denominators D3(P) in determining corrections δ(P;Q) benefits the semi-stochastic CC(P;Q) calculations aimed at converging the CCSDT energetics. In Section III C, which is the final part of our discussion of the numerical results obtained in this work, we investigate if similar applies to the CIQMC-driven CC(P;Q) computations targeting CCSDTQ.

C. Double Dissociation of H2O

Our last example, which illustrates the ability of the semi-stochastic CC(P) and CC(P;Q) approaches to converge the CCSDTQ energetics, is the C2v-symmetric cut of the ground-state potential energy surface of the water molecule, in which both O–H bonds are simultaneously stretched without changing the ∠(H–O–H) angle, resulting in large T3 and T4 contributions. Following Ref. 118 and consistent with our earlier deterministic CC(P;Q)
study, where we also obtained the reference CCSDTQ energies, we used the spherical cc-pVQZ basis set, correlated all electrons, and considered four stretches of the O–H bonds, including $R_{O-H} = 1.5R_e, 2R_e, 2.5R_e$, and $3R_e$, in addition to the equilibrium geometry, $R_{O-H} = R_e$. We used the same geometries, which the reader can find in Ref. 118, in the semi-stochastic CC($P$) and CC($P;Q$) calculations for H$_2$O carried out in this work, summarized in Table IV and Fig. 6. The authors of Ref. 118 obtained the CCSDTQ energies too, but we rely on our own CCSDTQ data, published in Ref. 63 and recalculated in this study, since Ref. 118 does not provide the CCSDTQ results for $R_{O-H} = 2.5R_e$ and $3R_e$ and the CCSDTQ energies for $R_{O-H} = 1.5R_e$ and $2R_e$ reported in Ref. 118 are in slight disagreement with the correctly converged values.

Up to twice the equilibrium O–H bond lengths, the CCSDT approach provides an accurate description of the electronic energies of water, resulting in the 0.493, 1.423, and $1.405$ millihartree signed errors relative to FCI at $R_{O-H} = R_e$, $1.5R_e$, and $2R_e$, respectively, when the cc-pVQZ basis set is employed, but when $R_{O-H} > 2R_e$, CCSDT completely fails.$^{63,118}$ and the CCSD(T), CCSD(2)$_T$ or CR-CC(2,3)$_A$ (in Table IV) $\tau = 0$ CC($P;Q$)$_{MP}$, CR-CC(2,3)$_D$ (in Table IV) $\tau = 0$ CC($P;Q$)$_{EN}$, CCSDt, and CC(t;3) approximations to CCSDT, which were examined in Refs. 63, 77, 88, and 118 fail with it (CCSD(T) fails already at $R_{O-H} = 2R_e$).

In particular, the difference between the CCSDT and FCI energies obtained with the cc-pVQZ basis set at $R_{O-H} = 2.5R_e$ is $-24.752$ millihartree. At $R_{O-H} = 3R_e$, the situation becomes even more dramatic, with the CCSDT/cc-pVQZ energy falling $40.126$ millihartree below its FCI counterpart.$^{63,118}$ One needs to incorporate $T_4$ clusters to reduce these massive errors in the $R_{O-H} > 2R_e$ region, and in order to do it in a reliable manner one has to use full CCSDTQ or one of the robust approximations to it, such as the CCSDTq, CC(t,q;3), and CC(t;q;3,4) methods tested in Ref. 63. The conventional $T_3$ plus $T_4$ corrections to CCSD, such as CCSD(TQ)$^{144}$ or their CCSD(2)$_T^{88,90}$ and CR-CC(2,4)$^{75,76,83}$ counterparts examined in Refs. 63 and 88 do not suffice. The CCSDT(2)$_Q$ quadruples correction to CCSDT$^{90}$ is not robust enough either.$^{88}$

When the cc-pVQZ basis set is employed, the differences between the CCSDQ and FCI energies at $R_{O-H} = R_e, 1.5R_e, 2R_e, 2.5R_e$, and $3R_e$ are $0.019, 0.121, 0.030, -2.361$, and $-4.733$ millihartree, respectively.$^{88}$ which is a huge improvement over CCSDT. One might argue the need for the inclusion of $T_3$ clusters with $n > 4$ at $R_{O-H} = 2.5R_e$ and $3R_e$, or try to obtain further improvements in describing the $R_{O-H} > 2R_e$ region by replacing the RHF reference determinants used throughout this work by their unrestricted counterparts, but studies of this kind are outside the scope of this article. The goal of the calculations for the water molecule discussed in this subsection is to explore the potential offered by the semi-stochastic CC($P$) and CC($P;Q$) approaches, especially the CC($P;Q$) corrections to the CC($P$) energies calculated with the help of the FCIQMC and CISD-TQ-MC propagations, in converging the CCSDTQ energetics obtained with the spin- and symmetry-adapted RHF references.

As shown in Table IV and Fig. 6, the semi-stochastic CC($P;Q$) calculations using FCIQMC and CISD-TQ-MC are extremely efficient in capturing the combined effects of $T_3$ and $T_4$ correlations, even in the most challenging $R_{O-H} > 2R_e$ region, where the $T_4$ contributions, which have to overcome the massive failures of the CCSDT approach, are very large and difficult to balance with their $T_3$ counterparts. They accurately reproduce the parent CCSDTQ energetics already in the early stages of the FCIQMC and CISD-TQ-MC propagations, greatly accelerating convergence of the underlying CC($P$) calculations, in spite of the fact that in our current implementation of the semi-stochastic CC($P;Q$) routines aimed at CCSDTQ the noniterative correction $\delta(P;Q)$ corrects the energy obtained by solving the CC($P$) equations in the space of all singles and doubles and subsets of triples and quadruples captured by FCIQMC or CISD-TQ-MC for the triples outside the stochastically determined $P$ space, but not for the quadruples missed by CIQMC.

Similarly to the previously discussed CC($P;Q$) calculations aimed at CCSDT, the CC($P;Q$) approach targeting CCSDTQ that adopts the CC($P;Q$)$_{EN}$ correction is generally most effective, although the results of the CC($P;Q$)$_{MP}$ calculations, in which the Epstein–Nesbet denominator $D_K(P)$ in Eq. (4) is replaced by its Møller–Plesset form, are as accurate as their CC($P;Q$)$_{EN}$ counterparts in the quasi-degenerate $R_{O-H} > 2R_e$ region. Indeed, when we look at the results in Table IV corresponding to $R_{O-H} = 2.5R_e$ and $3R_e$, where the $T_4$ effects, estimated by forming the differences of the CCSDTQ and CCSD energies, exceed 22 and 35 millihartree, respectively.$^{63,118}$ and where the differences between the CCSDT and CCSD energies, which measure the magnitude of the $T_3$ contributions, are about $-45$ and $-51$ millihartree, respectively.$^{63,118}$ the FCIQMC- and CISD-TQ-MC-based CC($P;Q$)$_{EN}$ computations reduce the large $-20.739$ ($R_{O-H} = 2.5R_e$) and $-35.823$ ($R_{O-H} = 3R_e$) millihartree errors relative to CCSDTQ obtained in the initial CR-CC(2,3)$_D$ ($\tau = 0$ CC($P;Q$)$_{EN}$) calculations to fractions of a millihartree after only 20,000 $\delta t = 0.0001$ a.u. MC iterations, i.e., after the FCIQMC and CISD-TQ-MC propagations capture as little as 5–6% of the triples and 1% of the quadruples in the corresponding $P$ spaces. The FCIQMC- and CISD-TQ-MC-driven CC($P;Q$)$_{MP}$ calculations using the same QMC propagation time $\tau$ are similarly effective though. They reduce the large $-13.469$ and $-28.302$ millihartree errors relative to CCSDTQ resulting from the initial CCSD(2)$_P$ or CR-CC(2,3)$_A$ ($\tau = 0$ CC($P;Q$)$_{MP}$) computations to a submillihartree level too.

The situation changes in the $R_{O-H} = R_e - 2R_e$ region, where the $T_4$ effects are much smaller than those originating from the $T_3$ clusters. In this case, the con-
The convergence of the energies obtained in the semi-stochastic CC(P;Q) calculations toward CCSDTQ is slower than that obtained with the CC(P;Q)\text{EN} approach, i.e., our earlier conclusion, drawn from the calculations discussed in Sections 11A and 11B and Ref. 68, that the use of the CC(P;Q)\text{EN} corrections to the semi-stochastic CC(P) energies is generally most effective still stands. This becomes particularly clear when we consider the results of the FCIQMC- and CISDTQ-MC-driven CC(P;Q)\text{MP} and CC(P;Q)\text{EN} calculations at R_{O-H} = R_e and 1.5R_e. For example, it takes only 40,000 0.0001 a.u. MC time steps, or about 10% of the triples and 2% of the quadruples captured in the P space, for the CC(P;Q)\text{EN} approach to reach a 0.1 millihartree accuracy level relative to CCSDTQ at R_{O-H} = R_e. The CC(P;Q)\text{MP} calculations reach the same accuracy level after 100,000 MC time steps that capture about 35% of the triples and 10% of the quadruples. When the R_{O-H} = 1.5R_e geometry is considered, the CC(P;Q)\text{EN} calculations reach a 0.1 millihartree accuracy level relative to CCSDTQ after 60,000–70,000 MC iterations that capture about 30% of the triples and 6–9% of the quadruples, i.e., in the relatively early stages of the FCIQMC and CISDTQ-MC propagations. The CC(P;Q)\text{MP} calculations reach a similar accuracy level 20,000–30,000 MC iterations later, after capturing about 40% of the triples and more than 10% of the quadruples. It is certainly reassuring that the CC(P;Q)\text{EN} calculations using FCIQMC and CISDTQ-MC to identify the leading triply and quadruply excited determinants for the inclusion in the underlying P spaces are capable of reproducing the CCSDTQ energies of the water molecule over a wide range of geometries along the C2v-symmetric cut of the ground-state potential energy surface considered in Table 11 and Fig. 11 to within ~0.1 millihartree out of the early stages of the CIQMC propagations, after capturing about 10% (R_{O-H} = R_e) or 30% (R_{O-H} > R_e) of the triples and 2% (R_{O-H} = R_e) or about 10% (R_{O-H} > R_e) of the quadruples. Having said this, it is interesting to observe that both types of the CC(P;Q) corrections tested in this study, abbreviated as CC(P;Q)\text{MP} and CC(P;Q)\text{EN}, perform equally well when R_{O-H} > 2R_e, i.e., when the T_3 and T_4 effects are both very large. We observed a similar behavior in Ref. 63, when examining the relative performance of the CC(P;Q)-based CC(t,q;3)\text{A} and CC(t,q;3)\text{D} corrections to CCSDtq using the double dissociation of water as one of the examples. This should not be surprising, since the CC(t,q;3)\text{A} and CC(t,q;3)\text{D} methods investigated in Ref. 63 can be regarded as the deterministic counterparts of the semi-stochastic CC(P;Q)\text{MP} and CC(P;Q)\text{EN} approaches targeting the CCSDTQ energetics implemented in this work.

As in the case of the CC(P;Q) calculations targeting CCSDT, discussed in Sections 11A and 11B the observed fast convergence of the semi-stochastic CC(P;Q) calculations aimed at recovering the CCSDTQ energetics does not seem to be affected by the type of the CIQMC approach used to identify the leading triply and quadruply excited determinants. This should facilitate future applications of the semi-stochastic CC(P;Q) methodology, including cases of stronger electronic quasi-degeneracies characterized by large T_3 and T_4 contributions, helping us to converge the CCSDTQ-level energetics at the small fraction of the deterministic CCSDTQ effort by taking advantage of the least expensive forms CIQMC capable of capturing triples and quadruples, represented in this study by CISDTQ-MC.

IV. CONCLUSIONS

We have recently started exploring a novel way of obtaining accurate electronic energies equivalent to high-level CC calculations, at the small fraction of the computational effort and preserving the black-box character of conventional single-reference computations, by merging the deterministic CC(P;Q) formalism, originally proposed in Refs. 57 and 61, along with the underlying CC(P)/EOMCC(P) framework, with the stochastic CIQMC\textsuperscript{58–61} and CCSDT\textsuperscript{60} approaches. When combined with the FCIQMC and CCSDT-MC wave function sampling, used to identify the leading triply excited determinants or cluster/excitation amplitudes, and correcting the CC(P)\textsuperscript{58} and EOMCC(P)\textsuperscript{59} energies for the remaining triples not captured by FCIQMC or CCSDT-MC, the resulting semi-stochastic CC(P;Q) methodology\textsuperscript{58} and its excited-state extension\textsuperscript{60} turned out to be very promising, allowing us to converge the CCSDT and EOMCCSDT energetics out of the early stages of the underlying QMC propagations.

This study can be regarded as the next key step in the development and exploration of the semi-stochastic CC(P;Q) approaches, in which we have extended our initial work\textsuperscript{23} focusing on recovering the CCSDT energetics and relying on FCIQMC and CCSDT-MC, to more efficient ways of identifying the leading higher-than-doubly excited determinants for the inclusion in the underlying P spaces. We have accomplished this goal by replacing FCIQMC by its less expensive CISDT-MC and CISDTQ-MC counterparts. We have also developed and tested the initial variant of the semi-stochastic CC(P;Q) method aimed at converging the CCSDTQ energetics, in which the results of CC(P) calculations in the subspaces spanned by singles, doubles, and subsets of triples and quadruples identified by FCIQMC or CISDTQ-MC are corrected for the remaining triples outside the stochastically determined P spaces. By comparing the FCIQMC-driven CC(P;Q) calculations targeting CCSDT, carried out in this work, in which the noniterative corrections δ(P;Q) to the CC(P) energies have been treated fully, as required by Eqs. 65, 69, and 71, with the analogous computations reported in Ref. 58, where the same corrections were treated in a somewhat simplified manner by neglecting the three-body component of the de-excitation operator A\textsuperscript{P} used to construct amplitudes t\textsubscript{K}(P) entering Eq. 65, we have examined the sig-
nificance of the full vs approximate treatment of these
 Corrections for the accuracy of the resulting CC(P;Q)
 energies. Other important issues, such as the benefits of
 using the Epstein–Nesbet form of the denominators
 $D_{\alpha}(P)$ that enter the definition of corrections $\delta(P;Q)$,
 resulting in the $CC(P;Q)_{\text{EN}}$ variant of $CC(P;Q)$, as com-
 pared to their Møller–Plesset counterparts defining the
 $CC(P;Q)_{\text{MP}}$ corrections, have been investigated as well.

The ability of the semi-stochastic CC(P;Q) approaches
 to converge the CCSDT and CCSDTQ energies, based
 on the truncated CISDT-MC and CISDTQ-MC propa-
gations, and their FCIQMC counterparts in which the
 noniterative corrections $\delta(P;Q)$ have been treated fully,
 has been illustrated using a few molecular examples, for
 which the deterministic CCSDT and CCSDTQ calcula-
tions that provide the reference data are feasible and
 which require a high-level CC treatment to obtain a
 reliable description. Thus, we have reported the re-
 sults of the semi-stochastic CC(P;Q) calculations using
 CISDT-MC, CISDTQ-MC, and FCIQMC aimed at con-
 verging the CCSDT energetics for the F–F bond breaking
 in $F_2$ and the automerization of cyclobutadiene, which
 require an accurate treatment of $T_3$ clusters account-
ing for the relaxation of $T_1$ and $T_2$ amplitudes in the
 presence of large $T_3$ contributions, and the CISDTQ-
 MC- and FCIQMC-driven CC(P;Q) computations for the
 $C_{2v}$-symmetric stretching of the O–H bonds in the
 water molecule targeting CCSDTQ, where the $T_3$ and $T_4$
 clusters become large and difficult to balance.

The numerical results reported in this article clearly
 show that the semi-stochastic CC(P;Q) calculations are
 capable of accurately reproducing the parent CCSDT
 and CCSDTQ energetics, even when electronic quasi-
degeneracies and higher–than–two-body components of
 the cluster operator become large, out of the early stages
 of the corresponding CIQMC propagations, accelerating
 convergence of the underlying $CC(P)$ computations at
 the same time. The convergence of the $CC(P;Q)$ ener-
gies toward their CCSDT and CCSDTQ parents does not
 seem to be affected by the type of the CIQMC approach
 used to identify the leading triply or triply and quadruply
 excited determinants. In the case of the CC(P;Q) cal-
culations targeting the CCSDT energetics, one can use
 FCIQMC or one of its less expensive truncated forms,
such as CISDTQ-MC, or even the crude CISDT-MC ap-
proach, with virtually no impact on the systematic con-
vergence pattern toward CCSDT as the propagation time
 $\tau$ approaches $\infty$. Similarly, one can replace FCIQMC
 by CISDTQ-MC without any significant effect on the
 convergence of the semi-stochastic CC(P;Q) calculations
 toward CCSDTQ. Our calculations also suggest that a
 complete treatment of the $CC(P;Q)$ corrections $\delta(P;Q)$,
as defined by Eqs. [3], [9], and [11], including the use of
 the $CC(P;Q)_{\text{EN}}$ variant, as opposed to its more ap-
proximate $CC(P;Q)_{\text{MP}}$ version, is more important than
 the actual type of the CIQMC approach used to deter-
mine the relevant $P$ spaces, especially when one is inter-
ested in accelerating convergence of the semi-stochastic
 $CC(P;Q)$ calculations in the early stages of the QMC
 propagations. We have demonstrated that independent
 of the type of the CIQMC approach used to identify the
 leading triply or triply and quadruply excited determin-
ants for the inclusion in the relevant $P$ spaces and inde-
pendent of the magnitude of $T_3$ and $T_4$ effects, the semi-
stochastic $CC(P;Q)$ calculations allow us to reach sub-
millihartree accuracy levels relative to the parent CCSDT
 and CCSDTQ energetics with small fractions of higher–
than–double excited determinants captured in the early
 stages of the corresponding CIQMC runs.

By relaxing $T_1$ and $T_2$ clusters in the presence of
 their $T_3$ or $T_3$ and $T_4$ counterparts defined using the
 excitation lists provided by full or truncated CIQMC,
 the semi-stochastic $CC(P;Q)$ computations are capable
 of considerably improving accuracy of the more estab-
lished noniterative corrections to CCSD without mak-
ing the calculations a lot more expensive. In this
 sense, the semi-stochastic $CC(P;Q)$ methodology using
 CIQMC is very similar to the deterministic $CC(t;3)$,
 $CC(t;q;3)$, and $CC(t;q;3;4)$ hierarchy developed and
tested in Refs. [57, 61–63, 82, and 100], which uses the
 $CC(P;Q)$ corrections to correct the results of the active-
 space CCSDt or CCSDtq calculations for the remain-
ing $T_3$ or $T_3$ and $T_4$ correlations that were not cap-
tured via active orbitals. There is, however, one ma-
jor advantage of the semi-stochastic $CC(P;Q)$ frame-
work over the $CC(t;3)$, $CC(t;q;3)$, and $CC(t;q;3;4)$ ap-
proaches, namely, the use of FCIQMC or truncated
 CIQMC propagations, which can efficiently identify the
 leading higher–than–double excited determinants for the
 inclusion in the relevant $P$ spaces, combined with the
 $\delta(P;Q)$ corrections to capture the remaining correla-
tions of interest, offers an automated way of performing
 accurate $CC(P;Q)$ computations without any refer-
ence to the user- and system-dependent active orbitals.

The analogies between the active-space CCSDt (for ex-
 acted states, EOMCCSD[26, 27, 145] and semi-stochastic
 $CC(P)/EOMCC(P)$ approaches, on which the deter-
mistic $CC(t;3)$ (in the case of CCSDt/EOMCCSDt)
 and CIQMC-driven (in the case of semi-stochastic
 $CC(P)/EOMCC(P)$) $CC(P;Q)$ approaches are based,
 have been investigated in Ref. [60].

The findings presented in this article are encourag-
ing from the point of view of future applications of the
 semi-stochastic $CC(P;Q)$ methodology using CIQMC, in-
cluding challenging cases of stronger electronic quasi-
degeneracies characterized by large $T_3$ or $T_3$ and $T_4$
 contributions that other approximations to CCSDT or CCS-
 DTQ may struggle with, but the story is not over yet. We
 certainly need to improve the efficiency of our $CC(P;Q)$
codes, especially the underlying $CC(P)$ routines, to ob-
tain full benefits offered by the semi-stochastic $CC(P;Q)$
approaches, discussed in Section [13]. This is especially
true in the case of our current $CC(P;Q)$ codes aimed at
converging the CCSDTQ energetics, which have a
largely pilot character. In this case, we also need to ex-
amine if one can further improve the convergence of
the FCIQMC- or CISDTQ-MC-driven CC(P,Q) calculations aimed at CCSDTQ by correcting the underlying CC(P) energies for both the missing triples and quadruples not captured by CIQMC at a given time \( \tau \), not just for the missing triples, as has been done in this work. It would also be useful to examine if one can extend the semi-stochastic CC(P) and CC(P,Q) approaches to the higher CC theory levels, beyond CCSDTQ examined in this work and beyond EOMCCSDT explored in Refs. 58 and 61, and investigate if our observations regarding the utility of the truncated CIQMC methods, such as CISDT-MC and CISDTQ-MC, remain true in the excited-state and open-shell CC(P,Q) calculations. In this study, we have adopted the original form of the \( \iota \)-CIQMC algorithm proposed in Ref. 63, but it would be interesting to examine if one could obtain additional benefits by interfacing our semi-stochastic CC(P,Q) methods with the improved ways of converging CIQMC, such as the adaptive-shift approach developed Refs. 63 and 114. All of the above ideas are presently pursued in our group, and the results will be reported as soon as they become available. Last, but not least, we have recently interfaced our CC(P) and CC(P,Q) routines with some of the modern versions of the selected CI approaches, which date back to the late 1960s and early 1970s\(^{146, 149}\) and which have recently regained significant attention\(^{150, 160}\). Our initial numerical results, which we hope to report in a separate publication\(^{58, 60}\), indicate that selected CI methods can be as effective in generating meaningful \( P \) spaces for the CC(P) calculations, which precede the determination of the \( \delta(P;Q) \) moment corrections, as the stochastic CIQMC propagations advocated in this and our earlier\(^{6, 16} \) studies.

ACKNOWLEDGMENTS

This work has been supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy (Grant No. DE-FG02-01ER15228 to P.P.), the National Science Foundation (Grant No. CHE-1763371 to P.P.), and Phase I and II Software Fellowships awarded to J.E.D. by the Molecular Sciences Software Institute funded by the National Science Foundation grant ACI-1547580. P.P. thanks Professors Ali Alavi, George H. Booth, and Alex J. W. Thom for useful discussions.

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

\(^1\)J. Hubbard, Proc. R. Soc. Lond., Ser. A 240, 539 (1957).
\(^2\)N. M. Hugenholtz, Physica 23, 481 (1957).
\(^3\)F. Coester, Nucl. Phys. 7, 421 (1958).
\(^4\)F. Coester and H. Kümml, Nucl. Phys. 17, 477 (1960).
\(^5\)J. Čiežek, J. Chem. Phys. 45, 4256 (1966).
\(^6\)J. Čiežek, Adv. Chem. Phys. 14, 35 (1969).
\(^7\)J. Paldus, J. Čiežek, and I. Shavitt, Phys. Rev. A 5, 50 (1972).
\(^8\)J. Paldus and X. Li, Adv. Chem. Phys. 110, 1 (1999).
\(^9\)P. Piecuch and K. Kowalski, Int. J. Mol. Sci. 3, 676 (2002).
\(^10\)R. J. Bartlett and M. Musial, Rev. Mod. Phys. 79, 291 (2007).
\(^11\)D. I. Lyakh, M. Musial, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. 112, 182 (2012).
\(^12\)F. A. Evangelista, J. Chem. Phys. 149, 030901 (2018).
\(^13\)G. D. Purvis, III and R. J. Bartlett, J. Chem. Phys. 76, 1910 (1982).
\(^14\)J. M. Cullen and M. C. Zerner, J. Chem. Phys. 77, 4088 (1982).
\(^15\)G. E. Scuseria, A. C. Scheiner, T. J. Lee, J. E. Rice, and H. F. Schaefer, III, J. Chem. Phys. 86, 2881 (1987).
\(^16\)P. Piecuch and J. Paldus, Int. J. Quantum Chem. 36, 429 (1989).
\(^17\)J. Noga and R. J. Bartlett, J. Chem. Phys. 86, 7041 (1987), 89, 3401 (1988) [Erratum].
\(^18\)G. E. Scuseria and H. F. Schaefer, III, Chem. Phys. Lett. 152, 382 (1988).
\(^19\)J. D. Watts and R. J. Bartlett, J. Chem. Phys. 93, 6104 (1990).
\(^20\)N. Oliphant and L. Adamowicz, J. Chem. Phys. 95, 6645 (1991).
\(^21\)S. A. Kucharski and R. J. Bartlett, Theor. Chim. Acta 80, 387 (1991).
\(^22\)S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. 97, 4282 (1992).
\(^23\)K. Emrich, Nucl. Phys. A 351, 379 (1981).
\(^24\)J. Geertsen, M. Rittby, and R. J. Bartlett, J. Chem. Phys. Lett. 164, 57 (1989).
\(^25\)J. F. Stanton and R. J. Bartlett, J. Chem. Phys. 98, 7029 (1993).
\(^26\)K. Kowalski and P. Piecuch, J. Chem. Phys. 115, 643 (2001).
\(^27\)K. Kowalski and P. Piecuch, Chem. Phys. Lett. 347, 237 (2001).
\(^28\)S. A. Kucharski, M. Wloch, M. Musial, and R. J. Bartlett, J. Chem. Phys. 115, 8263 (2001).
\(^29\)M. Källay and J. Gauss, J. Chem. Phys. 121, 9257 (2004).
\(^30\)S. Hirata, 121, 51 (2004).
\(^31\)H. Monkhorst, Int. J. Quantum Chem. Symp. 11, 421 (1977).
\(^32\)E. Dalgaard and H. Monkhorst, Phys. Rev. A 28, 1217 (1983).
\(^33\)D. Mukherjee and P. K. Mukherjee, Chem. Phys. 39, 325 (1979).
\(^34\)H. Sekino and R. J. Bartlett, Int. J. Quantum Chem. Symp. 18, 255 (1984).
\(^35\)M. Takahashi and J. Paldus, J. Chem. Phys. 85, 1486 (1986).
\(^36\)H. Koch and P. Jørgensen, J. Chem. Phys. 93, 3333 (1990).
\(^37\)H. Koch, H. J. A. Jensen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. 93, 3345 (1990).
\(^38\)A. E. Kondo, P. Piecuch, and J. Paldus, J. Chem. Phys. 102, 6511 (1995).
\(^39\)A. E. Kondo, P. Piecuch, and J. Paldus, J. Chem. Phys. 104, 8566 (1995).
\(^40\)M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, J. Chem. Phys. 83, 4041 (1985).
\(^41\)K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head Gordon, Chem. Phys. Lett. 157, 479 (1989).
\(^42\)Y. S. Lee and R. J. Bartlett, J. Chem. Phys. 80, 4371 (1984).
\(^43\)Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. 81, 5906 (1984), 82, 5761 (1985) [Erratum].
\(^44\)H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. 244, 75 (1995).
\(^45\)O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. 103, 7429 (1995).
\(^46\)P. Piecuch, K. Kowalski, I. S. O. Pimenta, and M. J. McGuire, Int. Rev. Phys. Chem. 21, 527 (2002).
\(^47\)P. Piecuch, K. Kowalski, I. S. O. Pimenta, P.-D. Fan, M. Lodriguito, M. J. McGuire, S. A. Kucharski, T. Kuś, and M. Musial, Theor. Chem. Acc. 112, 349 (2004).
\(^48\)P. Piecuch, Mol. Phys. 108, 2987 (2010).
\(^49\)K. Kowalski, D. J. Dean, M. Hjorth-Jensen, T. Papenbrock, and P. Piecuch, Phys. Rev. Lett. 92, 132501 (2004).
\(^50\)M. Wloch, D. J. Dean, J. R. Gour, M. Hjorth-Jensen, K. Kowalski, T. Papenbrock, and P. Piecuch, Phys. Rev. Lett. 94, 212501
TABLE I. Convergence of the CC(P), CC(P,Q)\textsubscript{MF}, and CC(P,Q)\textsubscript{EN} energies toward CCSDT, where the P spaces consisted of all singles and doubles and subsets of triples identified during the i-FCIQMC, i-CISDTQ-MC, or i-CISDT-MC propagations with $\sigma = 0.0001$ a.u. and where the corresponding Q spaces consisted of the triples not captured by the corresponding QMC simulations, for the F$_2$/cc-pVDZ molecule in which the F-F distance was set at 3.0 A, 3.5 A, and 4.0 A, with $R = 2.66816$ bohr representing the equilibrium geometry. The i-FCIQMC, i-CISDTQ-MC, and i-CISDT-MC calculations preceding the CC(P) and CC(P,Q) steps were initiated by placing 100 walkers on the RHF determinant and the $n_0$ parameter of the initiator algorithm was set at 3. In all post-RHF calculations, the lowest two core orbitals were kept frozen and the Cartesian components of d orbitals were employed throughout.

| $R/R_e$ | MC iterations | $\%$ of triples | CC(P)$^{a}$ | CC(P,Q)$^{a,b,c}$ | CC(P,Q)$^{a,b,c,d}$ |
|---------|---------------|-----------------|---------------|------------------|-----------------|
| 1.0     | 10000         | 3 3 4           | 9.2$^{g}$     | 1.39$^{g}$       | -1.0$^{g}$     |
|         | 20000         | 9 8 8           | 3.58$^{g}$    | 3.96$^{g}$       | 3.68$^{g}$     |
|         | 30000         | 15 16 14        | 2.29$^{g}$    | 2.76$^{g}$       | 2.84$^{g}$     |
|         | 40000         | 26 22 22        | 1.79$^{g}$    | 1.76$^{g}$       | 2.02$^{g}$     |
|         | 50000         | 37 38 34        | 0.93$^{g}$    | 1.15$^{g}$       | 0.93$^{g}$     |
|         | 60000         | 51 52 46        | 0.53$^{g}$    | 0.98$^{g}$       | 0.96$^{g}$     |
|         | 70000         | 64 64 58        | 0.38$^{g}$    | 0.41$^{g}$       | 0.44$^{g}$     |
|         | 80000         | 72 72 61        | 0.12$^{g}$    | 0.23$^{g}$       | 0.13$^{g}$     |
|         | 90000         | 85 85 99        | 0.01$^{g}$    | 0.02$^{g}$       | 0.01$^{g}$     |
|         | 120000        | 99 99 99        | 0.00$^{g}$    | 0.00$^{g}$       | 0.00$^{g}$     |

- $^{a}$ Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.
- $^{b}$ FCI stands for i-FCIQMC.
- $^{c}$ CIQ stands for i-CISDTQ-MC.
- $^{d}$ CIT stands for i-CISDT-MC.
- $^{e}$ Equivalent to CCSD.
- $^{f}$ Equivalent to the CCSD energy corrected for the effects of T3 clusters using the CCSDQ(2) approach of Ref. 88, which is equivalent to the approximate form of the completely renormalized CR-CC(2,3) approach of Refs. 23 and 24, abbreviated sometimes as CR-CC(2,3),A or CR-CC(2,3)\textsubscript{A}.
- $^{g}$ Equivalent to the CCSD energy corrected for the effects of T3 clusters using the most complete variant of the completely renormalized CR-CC(2,3) approach of Refs. 24 and 25, abbreviated sometimes as CR-CC(2,3)\textsubscript{D} or CR-CC(2,3)\textsubscript{D}.
- $^{h}$ Total CCSDT energy in hartree.
and the spherical components of \(d\) and \(P\) calculations preceding the CC(P) step.

| Basis set | MC iterations | \% of triples | \(CC(P)\) \(a\) | \(CC(P;Q)_{MP}\) \(b\) | \(CC(P;Q)_{EN}\) \(c\) |
|-----------|---------------|---------------|----------------|----------------|----------------|
| VTZ       | 0             | 0             | 62.816 \(d\) | 9.241 \(d\)  | 4.254 \(d\)   |
| 10000     | 1 1 1 1       | 29.714 1.31937 31.571 2.738 3.163 2.636 0.728 0.896 0.539 |
| 20000     | 2 2 2 2       | 11.179 14.687 20.194 0.824 1.097 1.487 0.071 0.151 0.217 |
| 30000     | 6 6 4 4       | 5.787 6.031 9.294 0.400 0.425 0.617 0.028 0.030 0.025 |
| 40000     | 14 14 10      | 2.406 2.574 4.203 0.160 0.171 0.284 0.002 0.001 0.014 |
| 50000     | 27 26 19      | 1.193 1.237 2.177 0.076 0.078 0.138 0.003 0.002 0.002 |
| 60000     | 42 42 30      | 0.490 0.489 1.144 0.029 0.029 0.071 0.002 0.002 0.005 |
| 70000     | 59 57 44      | 0.178 0.171 0.576 0.011 0.010 0.037 0.001 0.001 0.002 |
| 80000     | 72 71 56      | 0.045 0.054 0.309 0.003 0.003 0.020 0.000 0.000 0.001 |
| 100000    | 90 89 78      | 0.002 0.003 0.130 0.000 0.000 0.009 0.000 0.000 0.000 |
| \(\infty\) | 100           | -199.238344 \(e\) | - | - | - |

| AVTZ      | 0             | 0             | 65.035 \(e\) | 9.809 \(e\)  | 5.595 \(e\)   |
| 10000     | 0 0 0 0       | 36.316 38.874 42.801 3.641 4.144 4.851 1.594 1.796 2.304 |
| 20000     | 1 1 1 1       | 17.190 20.799 26.557 1.276 1.656 2.288 0.382 0.512 0.791 |
| 30000     | 4 4 3 3       | 8.065 9.272 13.279 0.549 0.623 0.928 0.138 0.138 0.246 |
| 40000     | 10 10 7       | 4.408 4.677 7.477 0.291 0.307 0.499 0.057 0.062 0.106 |
| 50000     | 23 22 15      | 2.208 2.425 3.951 0.136 0.150 0.244 0.016 0.019 0.038 |
| 60000     | 41 39 27      | 1.021 1.137 2.052 0.058 0.070 0.124 0.002 0.005 0.013 |
| 70000     | 61 58 61      | 0.385 0.455 0.385 0.021 0.025 0.059 0.000 0.000 0.001 |
| 80000     | 78 76 78      | 0.125 0.154 0.125 0.007 0.008 0.026 0.000 0.000 0.000 |
| 100000    | 97 96 97      | 0.007 0.009 0.007 0.000 0.001 0.004 0.000 0.000 0.000 |
| \(\infty\) | 100           | -199.253022 \(h\) | - | - | - |

\(a\) Unless otherwise stated, all energies are reported as errors relative to CCSDT in millihartree.

\(b\) FCI stands for \(i\)-FCIQMC.

\(c\) CIQ stands for \(i\)-CISDTQ-MC.

\(d\) CFT stands for \(i\)-CISDT-MC.

\(e\) Equivalent to CCSD.

\(f\) Equivalent to the CCSD energy corrected for the effects of \(T_3\) clusters using the CCSD(2) approach of Ref. \[78\], which is equivalent to the approximate form of the completely renormalized CR-CC(2,3) approach of Refs. \[75\] and \[76\], abbreviated sometimes as CR-CC(2,3),A or CR-CC(2,3),B\[75,76\].

\(g\) Equivalent to the CCSD energy corrected for the effects of \(T_3\) clusters using the most complete variant of the completely renormalized CR-CC(2,3) approach of Refs. \[75\] and \[76\], abbreviated sometimes as CR-CC(2,3),B or CR-CC(2,3),D\[75,76\].

\(h\) Total CCSDT energy in hartree.


TABLE III. Convergence of the CC(P), CC(P;Q)\textit{MP}, and CC(P;Q)\textit{EN} energies toward CCSDT, where the \textit{P} spaces consisted of all singles and doubles and subsets of triples identified during the \textit{i}-FCIQMC, \textit{i}-CISDTQ-MC, or \textit{i}-CISDT-MC propagations with \( \sigma = 0.0001 \) a.u. and where the corresponding \textit{Q} spaces consisted of the triples not captured by the corresponding QMC simulations, for the reactant (R) and transition state (TS) structures defining the automerization of cyclobutadiene, as described by the \textit{cc}-pVDZ basis set, optimized in the MR-AQCC calculations reported in Ref. 134, and for the corresponding activation barrier. The \textit{i}-FCIQMC, \textit{i}-CISDTQ-MC, and \textit{i}-CISDT-MC calculations preceding the CC(P) and CC(P;Q) steps were initiated by placing 100 walkers on the RHF determinant and the \( n_a \) parameter of the initiator algorithm was set at 3. In all post-RHF calculations, the lowest four core orbitals were kept frozen and the spherical components of \( d \) orbitals were employed throughout.

| Species | MC iterations | FCI\textit{a} | CIQ\textit{a} | CIT\textit{a} | FCI\textit{b} | CIQ\textit{b} | CIT\textit{b} | FCI\textit{c} | CIQ\textit{c} | CIT\textit{c} | FCI\textit{d} | CIQ\textit{d} | CIT\textit{d} | FCI\textit{e} | CIQ\textit{e} | CIT\textit{e} |
|---------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| R       | 10000         | 0             | 0             | 0             | 25.758        | 25.484        | 4.437         | 4.762         | 4.324         | 0.696         | 0.763         | 0.625         | 0.696         | 0.763         | 0.625         |
| TS      | 10000         | 0             | 0             | 0             | 45.875        | 45.777        | 18.899        | 20.080        | 14.636        | 13.680        | 13.842        | 12.665        | 13.680        | 13.842        | 12.665        |
| Barrier | 10000         | 0             | 0             | 0             | 12.624        | 12.825        | 12.734        | 13.274        | 9.611         | 8.653         | 8.208         | 7.553         | 8.653         | 8.208         | 7.553         |
| ∞       | 100           | 0             | 0             | 0             | 13.274        | 9.611         | —             | —             | —             | —             | —             | —             | —             | —             | —             |

\textit{a} Unless otherwise stated, all energies are reported as errors relative to CCSDT, in millihartree for the reactant and transition state and in kcal/mol for the activation barrier.

\textit{b} FCI stands for \textit{i}-FCIQMC.

\textit{c} CIQ stands for \textit{i}-CISDTQ-MC.

\textit{d} CIT stands for \textit{i}-CISDT-MC.

\textit{e} Equivalent to CCSD.

\textit{f} Equivalent to the CCSD energy corrected for the effects of \textit{T}_3 clusters using the CCSD(2)\textit{FT} approach of Ref. 88, which is equivalent to the approximate form of the completely renormalized CR-CC(2,3) approach of Refs. 17 and 24, abbreviated sometimes as CR-CC(2,3)\textit{A} or CR-CC(2,3)\textit{D}, or CR-CC(2,3)\textit{D}, equivalent to the CCSD energy corrected for the effects of \textit{T}_3 clusters using the most complete variant of the completely renormalized CR-CC(2,3)\textit{A} or CR-CC(2,3)\textit{D} D = 63.78.80.83.

\textit{g} The total CCSDT energy in hartree.

\textit{i} The CCSDT activation barrier in kcal/mol.
| R/R_{\text{eq}} | MC iterations | % of triples/quadruples | CC(P)_{\text{CIQ}} | CC(P;Q)_{\text{MP}} | CC(P;Q)_{\text{EN}} |
|-----------------|----------------|-------------------------|--------------------|----------------------|----------------------|
| 1.0             | 0              | 2/0                     | 3.291              | 0.718                | 0.220                |
|                 | 20000          | 4/1                     | 2.874              | 0.633                | 0.205                |
|                 | 30000          | 5/1                     | 2.637              | 0.544                | 0.143                |
|                 | 40000          | 9/3                     | 2.052              | 0.441                | 0.142                |
|                 | 50000          | 14/3                    | 1.910              | 0.390                | 0.105                |
|                 | 60000          | 18/4                    | 1.481              | 0.304                | 0.087                |
|                 | 70000          | 22/5                    | 1.238              | 0.245                | 0.065                |
|                 | 80000          | 27/6                    | 0.956              | 0.207                | 0.073                |
|                 | 100000         | 36/10                   | 0.586              | 0.127                | 0.048                |
|                 | ∞              | 100                     | −76.241841         | −2.704               | 1.021                |
| 1.5             | 0              | 3/1                     | 6.612              | 1.393                | 0.290                |
|                 | 20000          | 7/1                     | 4.908              | 0.613                | 0.144                |
|                 | 30000          | 11/2                    | 3.000              | 0.481                | 0.231                |
|                 | 40000          | 17/3                    | 1.878              | 0.377                | 0.185                |
|                 | 50000          | 22/4                    | 1.465              | 0.254                | 0.133                |
|                 | 60000          | 26/4                    | 0.930              | 0.220                | 0.079                |
|                 | 70000          | 31/8                    | 0.786              | 0.229                | 0.061                |
|                 | 80000          | 36/10                   | 0.552              | 0.106                | 0.060                |
|                 | 100000         | 46/17                   | 0.259              | 0.036                | 0.029                |
|                 | ∞              | 100                     | −76.072229         | −5.816               | −20.396              |
| 2.0             | 0              | 2/0                     | 11.766             | 1.966                | 0.044                |
|                 | 20000          | 6/1                     | 4.172              | 1.129                | 0.036                |
|                 | 30000          | 10/2                    | 3.132              | 0.708                | 0.036                |
|                 | 40000          | 13/2                    | 1.728              | 0.603                | 0.036                |
|                 | 50000          | 19/4                    | 1.123              | 0.421                | 0.036                |
|                 | 60000          | 25/6                    | 0.794              | 0.305                | 0.036                |
|                 | 70000          | 30/8                    | 0.429              | 0.129                | 0.036                |
|                 | 80000          | 36/11                   | 0.327              | 0.106                | 0.036                |
|                 | 100000         | 47/18                   | 0.107              | 0.030                | 0.036                |
|                 | ∞              | 100                     | −75.951635         | −0.038               | 0.036                |
| 2.5             | 0              | 3/0                     | 18.305             | −1.136               | −18.549              |
|                 | 20000          | 6/1                     | 5.254              | 0.010                | 0.036                |
|                 | 30000          | 10/2                    | 2.578              | 0.513                | 0.036                |
|                 | 40000          | 13/3                    | 1.021              | 0.304                | 0.036                |
|                 | 50000          | 22/5                    | 0.459              | 0.264                | 0.036                |
|                 | 60000          | 27/6                    | 0.340              | 0.122                | 0.036                |
|                 | 70000          | 31/9                    | 0.131              | 0.059                | 0.036                |
|                 | 80000          | 36/11                   | 0.088              | 0.052                | 0.036                |
|                 | 100000         | 53/28                   | 0.020              | 0.020                | 0.036                |
|                 | ∞              | 100                     | −75.920352         | −1.290               | −35.8236             |
| 3.0             | 0              | 3/0                     | 15.583             | −2.390               | −3.945               |
|                 | 20000          | 5/1                     | 10.165             | −2.199               | −3.945               |
|                 | 30000          | 8/2                     | 4.282              | −0.084               | −0.403               |
|                 | 40000          | 11/3                    | 1.636              | 0.544                | 0.141                |
|                 | 50000          | 17/5                    | 0.969              | 0.267                | 0.199                |
|                 | 60000          | 24/7                    | 0.523              | 0.251                | 0.231                |
|                 | 70000          | 30/12                   | 0.185              | 0.093                | 0.090                |
|                 | 80000          | 36/16                   | 0.030              | 0.022                | 0.021                |
|                 | 100000         | 51/28                   | 0.005              | 0.008                | 0.005                |
|                 | ∞              | 100                     | −75.916679         | −1.697               | −35.8236             |

a The equilibrium geometry, R_{\text{O-H}} = R_{\text{eq}}, and the geometries that represent a simultaneous stretching of both O-H bonds by factors of 1.5, 2.0, 2.5, and 3.0 without changing the \angle(H-O-H) angle were taken from Ref. 113.
b Unless otherwise stated, all energies are reported as errors relative to CCSDTQ in millihartree.
c FCI stands for i-FCIQMCC.
d CIQ stands for i-CISD-TQ-MC.
e Equivalent to CCSD.
f Equivalent to the CCSD energy corrected for the effects of T_{3} clusters using the CCSD(2,3) approach of Ref. 85, which is equivalent to the approximate form of the completely renormalized CR-CC(2,3) approach of Refs. 73 and 74, abbreviated sometimes as CR-CC(2,3)_{m} or CR-CC(2,3)_{D}.
g Equivalent to the CCSD energy corrected for the effects of T_{3} clusters using the most complete variant of the completely renormalized CR-CC(2,3) approach of Refs. 73 and 74, abbreviated sometimes as CR-CC(2,3)_{m} or CR-CC(2,3)_{D}.
h Total CCSDTQ energy in hartree.
FIG. 1. Convergence of the CC(P) (red filled circles and dashed lines) and CC(P:Q)_{EN} (black open squares and solid lines) energies toward CCSDT for the F\textsubscript{2}/cc-pVDZ molecule in which the F–F distance R was set at (a) $R_e$, (b) 1.5$R_e$, (c) 2$R_e$, and (d) 5$R_e$, where $R_e = 2.66816$ bohr is the equilibrium geometry. 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. (depicted by the green lines representing the corresponding projected energies). The $Q$ spaces consisted of the triples not captured by $i$-FCIQMC. All energies are errors relative to CCSDT in millihartree and the insets show the percentages of triples captured during the $i$-FCIQMC propagations.

FIG. 2. Same as Fig. 1 except that the subsets of triples included in the CC(P) calculations are now identified by the $i$-CISDTQ-MC simulations and the corresponding $Q$ spaces consist of the triples not captured by $i$-CISDTQ-MC.
FIG. 3. Same as Fig. 1 except that the subsets of triples included in the CC($P$) calculations are now identified by the $i$-CISDT-MC simulations and the corresponding $Q$ spaces consist of the triples not captured by $i$-CISDT-MC.

FIG. 4. The key molecular structures defining the automerization of cyclobutadiene. The leftmost and rightmost structures represent the degenerate reactant/product minima, whereas the structure in the center corresponds the transition state.
FIG. 5. Convergence of the CC($P$) (red filled circles and dashed lines) and CC($P;Q_{EN}$) (black open squares and solid lines) energies toward CCSDT for the reactant [panels (a)–(c)] and transition state [panels (d)–(f)] structures defining the automerization of cyclobutadiene, as described by the cc-pVDZ basis set. The relevant $i$-CIQMC runs (all using $\delta \tau = 0.0001$ a.u.) are depicted by the green lines representing the corresponding projected energies. Panels (a) and (d) correspond to the calculations in which the $P$ spaces employed in the CC($P$) steps consisted of all singles and doubles and subsets of triples identified during the $i$-FCIQMC propagations; the $Q$ spaces needed to define the corresponding $\delta(P;Q)$ corrections consisted of the triples that were not captured by $i$-FCIQMC. Panels (b) and (e) correspond to the calculations in which the $P$ spaces employed in the CC($P$) steps consisted of all singles and doubles and subsets of triples identified during the $i$-CISDTQ-MC propagations; in this case, the $Q$ spaces needed to define the $\delta(P;Q)$ corrections consisted of the triples that were not captured by $i$-CISDTQ-MC. Panels (c) and (f) correspond to the calculations in which the $P$ spaces employed in the CC($P$) steps consisted of all singles and doubles and subsets of triples identified during the $i$-CISDT-MC propagations; in this case, the $Q$ spaces needed to define the $\delta(P;Q)$ corrections consisted of the triples that were not captured by $i$-CISDT-MC. All reported energies are errors relative to CCSDT in millihartree. The insets show the percentages of triples captured during the relevant $i$-CIQMC propagations.
FIG. 6. Convergence of the CC($P$) (red filled circles and dashed lines) and CC($P;Q$)$_{\text{EN}}$ (black open squares and solid lines) energies toward CCSDTQ for the water molecule, as described by the cc-pVDZ basis set. The relevant $i$-CIQMC runs (all using $\delta \tau = 0.0001$ a.u.) are depicted by the green lines representing the corresponding projected energies. Panels (a) and (b) correspond to the calculations in which the $P$ spaces employed in the CC($P$) steps consisted of all singles and doubles and subsets of triples and quadruples identified during the $i$-FCIQMC propagations; the $Q$ spaces needed to define the corresponding $\delta(P;Q)$ corrections consisted of the triples that were not captured by $i$-FCIQMC. Panels (c) and (d) correspond to the calculations in which the $P$ spaces employed in the CC($P$) steps consisted of all singles and doubles and subsets of triples and quadruples identified during the $i$-CISDTQ-MC propagations; in this case, the $Q$ spaces needed to define the corresponding $\delta(P;Q)$ corrections consisted of the triples that were not captured by $i$-CISDTQ-MC. Panels (a) and (c) correspond to the equilibrium geometry. Panels (b) and (d) correspond to the geometry in which both O–H bonds in water are simultaneously stretched by a factor of 3 without changing the $\angle$(H–O–H) angle. All reported energies are errors relative to CCSDTQ in millihartree. The insets show the percentages of triples (blue line) and quadruples (purple line) captured during the relevant $i$-CIQMC propagations.