AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

Communication: Approaching exact quantum chemistry by cluster analysis of full configuration interaction quantum Monte Carlo wave functions

J. Emiliano Deustua,[1] Ilias Magoulas,[1] Jun Shen,[1] and Piotr Piecuch[1,2,a)]

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

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

Abstract

We propose to accelerate convergence toward full configuration interaction (FCI) energetics by using the coupled-cluster approach, in which singly and doubly excited clusters, needed to determine the energy, are iterated in the presence of their three- and four-body counterparts extracted from FCI quantum Monte Carlo (FCIQMC) propagations. Preliminary calculations for the water molecule at the equilibrium and stretched geometries show that we can accurately extrapolate the FCI energetics based on the early stages of FCIQMC propagations.

ACCEPTED MANUSCRIPT

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

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

Approximate quantum chemistry methods reduce prohibitive costs of the exact, full configuration interaction (FCI) calculations, but they require using higher–than–two-body correlation operators in the expressions defining the relevant many-electron wave functions if one is interested in obtaining accurate results, which may still lead to large computational costs without guaranteeing answers having a desired precision. There has been great progress in developing *ab initio* electronic structure methods that can help us achieve a highly accurate description of single- and multi-reference regions of molecular potential energy surfaces and weakly as well as strongly correlated regimes with manageable computational costs, but in any quantum chemistry method development work one is always tempted to ask a fundamental question if there is a way to directly aim at the exact, FCI-level, energetics and properties in a computationally efficient manner that avoids prohibitive costs of FCI calculations. Questions of this type have prompted the development of many interesting concepts and methods over the years, such as the density-matrix renormalization group approach,[1-6] the graphically contracted function approach,[7-10] the correlation energy extrapolation by intrinsic scaling,[11-13] incremental FCI,[14,15] the generalized coupled-cluster (CC) theory using two-body correlation operators,[16-20] the adaptive CI (ACI)[21,22] and adaptive sampling CI (ASCI)[23,24] methods, the stochastic FCI quantum Monte Carlo (FCIQMC) methodology[25,26] and its semi-stochastic extension,[27] and the heat-bath sampling of many-body quantum states in the Fock space combined with CI and semi-stochastic perturbation theory,[28-30] to mention a few representative more recent examples. We have demonstrated that one can greatly benefit from incorporating the stochastic FCIQMC methodology and its CC counterpart[31-34] into the deterministic CC framework.[35] We believe that we can take the merger of the stochastic and deterministic ideas to the next level and use it to produce the FCI-level energetics based on the early stages of FCIQMC propagations.

Several ideas have been explored to improve the convergence of FCIQMC computations, including the initiator FCIQMC (*i*-FCIQMC) approach[26] and the perturbative corrections built from the information discarded when applying initiator criteria.[36] Here, we propose to accelerate convergence toward FCI by combining the stochastic wave function sampling with the cluster analysis of the FCIQMC wave functions. The proposed idea takes advantage of the observation, which has historically been used to develop a wide variety of the externally corrected CC methods,[37-46] that, as long as the Hamiltonian H does not contain higher–than–two-body

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

interactions, the amplitude equations of the single-reference CC theory[47-50] projected on the singly and doubly excited determinants,

$$\left\langle \Phi_i^a \left| [H_N(1+T_1+T_2+(1/2)T_1^2+\mathbf{T_3}+T_1T_2+(1/6)T_1^3)]_C \right| \Phi \right\rangle = 0, \tag{1}$$

$$\begin{aligned}\left\langle \Phi_{ij}^{ab} \left| [H_N(1+T_1+T_2+(1/2)T_1^2+\mathbf{T_3}+T_1T_2+(1/6)T_1^3+\mathbf{T_4}+T_1\mathbf{T_3}+(1/2)T_2^2 \right.\right. \\ \left.\left. +(1/2)T_1^2T_2+(1/24)T_1^4)]_C \right| \Phi \right\rangle = 0,\end{aligned} \tag{2}$$

yield the exact, FCI, values of the one- and two-body clusters, T_1 and T_2, respectively, and, as a consequence, the exact ground-state energy,

$$E_0 = \langle \Phi | H | \Phi \rangle + \left\langle \Phi \left| [H_N(T_1+T_2+(1/2)T_1^2)]_C \right| \Phi \right\rangle, \tag{3}$$

if one solves them in the presence of the exact triply (T_3) and quadruply (T_4) excited clusters (marked in Eqs. (1) and (2) in bold roman), extracted from FCI. Here, we use the notation in which $i, j, \ldots$ ($a, b, \ldots$) are the spin-orbitals occupied (unoccupied) in the reference determinant $|\Phi\rangle$, $H_N = H - \langle \Phi | H | \Phi \rangle$ is the Hamiltonian in the normal-ordered form, and subscript C represents the connected operator product. Clearly, when we neglect the T_3 and T_4 clusters in Eqs. (1) and (2), as is done in the conventional CCSD approach,[51,52] the exactness of the resulting T_1 and T_2 amplitudes for systems with more than two electrons is gone, but we can greatly improve the results of the CCSD calculations by correcting them for the T_3 and T_4 contributions, as shown in Eqs. (1) and (2), if we know how to estimate them. In the externally corrected CC methods, one tries to find a non-CC source of the T_3 and T_4 clusters, which behaves well in physical situations of interest, such as bond breaking or strong correlations, and the aforementioned Refs. 37-46 have examined various ideas in this direction, but one can also use Eqs. (1)–(3) in the previously unexplored way proposed in this communication, where T_3 and T_4 are determined during the FCIQMC wave function propagation, which is guaranteed to converge to FCI.

In order to extract information about the values of the T_3 and T_4 clusters corresponding to the FCIQMC wave function at a given propagation time τ, we represent it as

$$\left| \Psi_0^{(\mathrm{MC})}(\tau) \right\rangle = \left[1 + \sum_{n=1}^{N} C_n^{(\mathrm{MC})}(\tau) \right] | \Phi \rangle, \tag{4}$$

where we imposed the intermediate normalization condition on $\left| \Psi_0^{(\mathrm{MC})}(\tau) \right\rangle$ to carry out the subsequent cluster analysis by comparing the CI wave function expansion, Eq. (4), with the corresponding CC ansatz

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

$$\left|\Psi_0^{(\mathrm{MC})}(\tau)\right\rangle = \exp\left[\sum_{n=1}^{N} T_n^{(\mathrm{MC})}(\tau)\right]\left|\Phi\right\rangle. \tag{5}$$

Following the definitions adopted by the FCIQMC approach and because of the intermediate normalization used to represent $\left|\Psi_0^{(\mathrm{MC})}(\tau)\right\rangle$, the time-dependent $c_K(\tau)$ amplitudes entering the CI excitation operators in Eq. (4),

$$C_n^{(\mathrm{MC})}(\tau) = \sum_K c_K(\tau) E_K, \tag{6}$$

where E_K is the elementary particle-hole excitation operator generating the corresponding excited determinant $|\Phi_K\rangle$ from $|\Phi\rangle$ via the expression $|\Phi_K\rangle = E_K|\Phi\rangle$, are determined by forming the ratios of the numbers of signed walkers at a given time τ located on $|\Phi_K\rangle$ and $|\Phi\rangle$, i.e.,

$$c_K(\tau) = N_K / N_0. \tag{7}$$

We recall that $N_K = \sum_\alpha s_\alpha \delta_{K,K_\alpha}$, where K_α designates the determinant on which walker α is located and $s_\alpha = \pm 1$ is the corresponding walker sign.[25,26] Once the CI excitation operators $C_n^{(\mathrm{MC})}(\tau)$ representing the FCIQMC wave function $\left|\Psi_0^{(\mathrm{MC})}(\tau)\right\rangle$ at a given time τ are determined, we can extract the corresponding T_3 and T_4 clusters, which we designate $T_3^{(\mathrm{MC})}(\tau)$ and $T_4^{(\mathrm{MC})}(\tau)$, respectively, using the well-known relationships between the CI and CC excitation operators,

$$\begin{aligned} C_1 &= T_1, \\ C_2 &= T_2 + (1/2)T_1^2, \\ C_3 &= T_3 + T_1T_2 + (1/6)T_1^3, \\ C_4 &= T_4 + T_1T_3 + (1/2)T_2^2 + (1/2)T_1^2T_2 + (1/24)T_1^4, \text{ etc.} \end{aligned} \tag{8}$$

As one can see, all one needs to know to determine the $T_3^{(\mathrm{MC})}(\tau)$ and $T_4^{(\mathrm{MC})}(\tau)$ clusters corresponding to the FCIQMC wave function $\left|\Psi_0^{(\mathrm{MC})}(\tau)\right\rangle$ are the CI excitation amplitudes through quadruples ($C_n^{(\mathrm{MC})}(\tau)$, $n = 1-4$).

The above description suggests the following procedure, termed a cluster-analysis-driven FCIQMC and abbreviated as CAD-FCIQMC, to converge the FCI-level energetics:

(1) Start a FCIQMC propagation by placing a certain number of walkers on the reference determinant $\left|\Phi\right\rangle$.

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

(2) Extract $C_n^{(\mathrm{MC})}(\tau)$, $n = 1-4$, from a FCIQMC wave function, Eq. (4), at a given time τ using Eq. (7).

(3) Use Eq. (8) to perform cluster analysis of the FCIQMC wave function at time τ to extract the corresponding $T_3^{(\mathrm{MC})}(\tau)$ and $T_4^{(\mathrm{MC})}(\tau)$ clusters.

(4) Solve the CCSD amplitude equations corrected for terms containing $T_3^{(\mathrm{MC})}(\tau)$ and $T_4^{(\mathrm{MC})}(\tau)$, i.e., Eqs. (1) and (2), where $\mathbf{T_3} = T_3^{(\mathrm{MC})}(\tau)$ and $\mathbf{T_4} = T_4^{(\mathrm{MC})}(\tau)$, to determine the singly and doubly excited clusters relaxed in the presence of $T_3^{(\mathrm{MC})}(\tau)$ and $T_4^{(\mathrm{MC})}(\tau)$ and the corresponding ground-state energy $E_0(\tau)$, Eq. (3).

(5) Check the convergence by repeating steps 2–4 at some later FCIQMC propagation time $\tau' > \tau$ and inspecting if the resulting energy $E_0(\tau')$ and singly and doubly excited clusters are converged within a given convergence threshold. It is guaranteed that $\lim_{\tau\to\infty} E_0(\tau) = E_0^{(\mathrm{FCI})}$, where $E_0^{(\mathrm{FCI})}$ is the FCI energy.

In order to test the proposed CAD-FCIQMC algorithm and compare its convergence with the *i*-FCIQMC calculations using the HANDE code,[53] we performed preliminary computations, using our home-grown CC routines, interfaced with GAMESS,[54] and cluster analysis codes, for the symmetric stretching of the O-H bonds in the water molecule, as described by the cc-pVDZ basis set.[55] Thus, in addition to the equilibrium geometry, $R_{\mathrm{O\text{-}H}} = R_{\mathrm{e}}$, we considered two other geometries in which both O-H bonds were simultaneously stretched by factors of 1.5 and 2 without changing the ∠(H-O-H) angle. The corresponding geometries and FCI energies can be found in Ref. 56. We started the *i*-FCIQMC propagation by placing 1,000 walkers on the restricted Hartree–Fock (RHF) reference determinant, and the time step used in the *i*-FCIQMC runs was $\Delta\tau = 0.0001$ a.u. As shown in Table I and Fig. 1, the convergence toward FCI resulting from the proposed merger of the *i*-FCIQMC propagation and the deterministic CCSD-like calculations using the triply and quadruply excited cluster amplitudes extracted from the FCIQMC wave functions at selected time steps τ, which were subsequently read into Eqs. (1) and (2) to obtain the corresponding singly and doubly excited clusters and energies $E_0(\tau)$ (we performed the $T_3^{(\mathrm{MC})}(\tau)$- and $T_4^{(\mathrm{MC})}(\tau)$-corrected CCSD calculations every 1,000 MC time steps, called in Table I and Fig. 1 iterations), is smoother and considerably faster than in the case of the

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

purely stochastic *i*-FCIQMC runs. As anticipated, some of the noise resulting from the *i*-FCIQMC calculations propagates, via the stochastically determined $T_3^{(MC)}(\tau)$ and $T_4^{(MC)}(\tau)$ clusters, to the corresponding CAD-FCIQMC runs, but the convergence of the CAD-FCIQMC algorithm toward FCI, allowing us to obtain millihartree or submillihartree errors without going to longer propagation times, is significantly faster than in the case of the *i*-FCIQMC simulations. As one can see in Table I and Fig. 1, unlike in the case of the *i*-FCIQMC runs, the convergence of the CAD-FCIQMC energies toward FCI is similar at the equilibrium and stretched geometries.

Further insights into the performance of our CAD-FCIQMC approach can be gained if, in analogy to our recent study,[35] we correlate the accuracy of the CAD-FCIQMC energies with the percentages of triples and quadruples (meaning triply and quadruply excited determinants having at least one positive or negative walker on them) captured during the *i*-FCIQMC runs. We focus on triples and quadruples, since higher–than–quadruply excited determinants do not enter the cluster analysis of FCIQMC wave functions needed to determine the CAD-FCIQMC energies. As shown in Tables I and II and Fig. 1, the CAD-FCIQMC calculations are capable of reaching reasonably stable submillihartree accuracy levels relative to FCI in the early stages of the *i*-FCIQMC propagations, after capturing about 10 % of triples and 1 % of quadruples at $R_{O\text{-}H} = R_e$ (i.e., after about 10,000 $\Delta\tau$ = 0.0001 a.u. MC iterations), ~25 % of triples and ~5 % of quadruples at $R_{O\text{-}H} = 1.5\ R_e$ (after ~60,000 MC iterations), and ~40 % of triples and ~10 % of quadruples at $R_{O\text{-}H} = 2.0\ R_e$ (after ~100,000 MC iterations). They reach these high accuracies long before converging walker populations or achieving walker equilibration in the corresponding *i*-FCIQMC runs. Indeed, as shown in Table II, the walker populations corresponding to the above fractions of triples and quadruples captured during the *i*-FCIQMC propagations are very small, even when compared to the numbers of walkers at τ = 16.0 a.u., when we stopped our *i*-FCIQMC runs. Because of the relatively fast convergence of the CAD-FCIQMC energies toward FCI, which is one of the key advantages of the proposed CAD-FCIQMC algorithm, we did not run our *i*-FCIQMC calculations long enough to stabilize and equilibrate walker populations. As shown in Table II, our *i*-FCIQMC propagations are still in the equilibration stage, with walker numbers increasing with the iteration count. As a consequence, we could not perform a rigorous analysis of the *i*-FCIQMC data used in conventional FCIQMC calculations, where one resorts to the appropriate blocking analysis[57] of the resulting energies after longer *i*-FCIQMC runs, averaging a series of independent calculations, and applying energy

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

shifts once critical numbers of walkers are reached.[25,26] Thus, in examining and further comparing our CAD-FCIQMC and *i*-FCIQMC calculations we resorted to straightforward mean value and standard deviation analysis shown in Table II. This elementary statistical analysis confirms that the CAD-FCIQMC energies stabilize early on, achieving submillihartree accuracies long before the corresponding *i*-FCIQMC calculations can do it. The primary role of the *i*-FCIQMC propagations within the CAD-FCIQMC computations is to identify the dominant triply and quadruply excited wave function components and, through cluster analysis of the FCIQMC wave functions, the leading T_3 and T_4 contributions, needed to determine accurate T_1 and T_2 values by solving Eqs. (1) and (2), and this does not require long *i*-FCIQMC runs or equilibration of walkers on all higher–than–doubly excited determinants. As a result, we obtain accurate estimates of the FCI energies out of the short *i*-FCIQMC propagations and relatively inexpensive CCSD-like computations.

Clearly, the above test is only a preliminary raw calculation, where we did not make any attempt to further accelerate our CAD-FCIQMC runs or to make them even smoother than they already are by feeding the T_1 and T_2 cluster amplitudes obtained deterministically in step 4 of the CAD-FCIQMC algorithm back to the *i*-FCIQMC propagations. One could envision doing this by using numerical values of the T_1 and T_2 amplitudes resulting from solving Eqs. (1) and (2), in which $\mathbf{T_3} = T_3^{(\mathrm{MC})}(\tau)$ and $\mathbf{T_4} = T_4^{(\mathrm{MC})}(\tau)$, to determine the corresponding C_1 and C_2 excitation coefficients that would subsequently be translated back into the numbers of walkers located on the singly and doubly excited determinants using Eq. (7), potentially accelerating the *i*-FCIQMC propagations themselves. By extracting the $T_3^{(\mathrm{MC})}(\tau)$ and $T_4^{(\mathrm{MC})}(\tau)$ clusters from the *i*-FCIQMC calculations and using Eqs. (1) and (2) to determine improved T_1 and T_2 clusters, and then feeding the information about these clusters back into the *i*-FCIQMC algorithm, one should be able to converge the FCI energetics even faster than what is seen in Table I and Fig. 1. This is one of the aspects of the CAD-FCIQMC procedure suggested in this communication that we would like to thoroughly explore in the future. Also, as already alluded to above, we do not have to use the deterministic step 4 of the CAD-FCIQMC procedure as frequently as we did it in our test calculations for the water molecule shown in Table I and Fig. 1, where we performed the $T_3^{(\mathrm{MC})}(\tau)$- and $T_4^{(\mathrm{MC})}(\tau)$-corrected CCSD calculations once every 1,000 MC iterations for the purpose of learning about the CAD-FCIQMC convergence pattern. The fast and relatively

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

smooth convergence of the CAD-FCIQMC energies, observed in Table I and Fig. 1, and further examined in Table II, indicates that we should be able to rely on a few such CCSD-like calculations to accurately extrapolate the desired FCI energetics, especially if we incorporate the above feedback mechanism in our codes.

In summary, in this communication we have proposed the idea of accelerating convergence of the FCI-level energetics by cluster analysis of the FCIQMC wave functions, followed by deterministic CCSD-like calculations, in which the singly and doubly excited clusters, needed to determine the energy, are iterated in the presence of their three- and four-body counterparts extracted from FCIQMC. Based on our encouraging preliminary results, shown in Table I and Fig. 1, and possible future enhancements of the proposed CAD-FCIQMC algorithm, such as the feedback mechanism discussed above, we anticipate that we will be able to accurately extrapolate the exact, FCI, energetics based on the early stages of the deterministically driven *i*-FCIQMC propagations. Although the initial results reported in this communication need more study, one can think of extending them to accelerate convergence of other methodologies aimed at FCI, such as the ACI and ASCI approaches of Refs. 21-24.

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).

[1] S.R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).

[2] S.R. White and R.L. Martin, *J. Chem. Phys*. **110**, 4127 (1999).

[3] A.O. Mitrushenkov, G. Fano, F. Ortolani, R. Linguerri, and P. Palmieri, *J. Chem. Phys*. **115**, 6815 (2001).

[4] G.K.-L. Chan and M. Head-Gordon, *J. Chem. Phys*. **116**, 4462 (2002).

[5] S. Keller, M. Dolfi, M. Troyer, and M. Reiher, *J. Chem. Phys*. **143**, 244118 (2015), and references therein.

[6] G.K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, and S.R. White, *J. Chem. Phys*. **145**, 014102 (2016), and references therein.

[7] R. Shepard, *J. Phys. Chem. A* **109**, 11629 (2005).

[8] G. Gidofalvi and R. Shepard, *Mol. Phys*. **108**, 2717 (2010).

[9] R. Shepard, G. Gidofalvi, and S.R. Brozell, *J. Chem. Phys*. **141**, 064105 (2014).

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

[10] R. Shepard, G. Gidofalvi, and S.R. Brozell, *J. Chem. Phys*. **141**, 064106 (2014).

[11] L. Bytautas and K. Ruedenberg, *J. Chem. Phys*. **121**, 10905 (2004).

[12] L. Bytautas and K. Ruedenberg, *J. Chem. Phys*. **121**, 10919 (2004).

[13] L. Bytautas and K. Ruedenberg, *J. Chem. Phys*. **121**, 10852 (2004).

[14] P.M. Zimmerman, *J. Chem. Phys*. **146**, 104102 (2017).

[15] P.M. Zimmerman, *J. Chem. Phys*. **146**, 224104 (2017).

[16] M. Nooijen, *Phys. Rev. Lett.* **84**, 2108 (2000).

[17] H. Nakatsuji, *J. Chem. Phys*. **113**, 2949 (2000).

[18] T. Van Voorhis and M. Head-Gordon, *J. Chem. Phys*. **115**, 5033 (2001).

[19] P. Piecuch, K. Kowalski, P.-D. Fan, and K. Jedziniak, *Phys. Rev. Lett.* **90**, 113001 (2003).

[20] P.-D. Fan and P. Piecuch, *J. Mol. Struct.: THEOCHEM* **768**, 3 (2006).

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

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

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

[24] N.M. Tubman, C.D. Freeman, D.S. Levine, D. Hait, M. Head-Gordon, and K.B. Whaley, arXiv:1807.00821v1 [physics.comp-ph].

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

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

[27] N.S. Blunt, S.D. Smart, J.A.F. Kersten, J.S. Spencer, G.H. Booth, and A. Alavi, *J. Chem. Phys.* **142**, 184107 (2015).

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

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

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

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

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

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

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

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

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

[36] N.S. Blunt, *J. Chem. Phys.* **148**, 221101 (2018).

[37] J. Paldus, J. Čížek, and M. Takahashi, *Phys. Rev. A* **30**, 2193 (1984).

[38] P. Piecuch and J. Paldus, *Theor. Chim. Acta* **78**, 65 (1990).

[39] P. Piecuch, R. Toboła, and J. Paldus, *Phys. Rev. A* **54**, 1210 (1996).

[40] J. Paldus and J. Planelles, *Theor. Chim. Acta* **89**, 13 (1994).

[41] L.Z. Stolarczyk, *Chem. Phys. Lett.* **217**, 1 (1994).

[42] G. Peris, J. Planelles, and J. Paldus, *Int. J. Quantum Chem*. **62**, 137 (1997).

[43] G. Peris, J. Planelles, J.-P. Malrieu, and J. Paldus, *J. Chem. Phys*. **110**, 11708 (1999).

[44] X. Li and J. Paldus, *J. Chem. Phys.* **107**, 6257 (1997).

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

[46] J. Paldus, *J. Math. Chem*. **55**, 477 (2017), and references therein.

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

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

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

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

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

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

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

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

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

[56] J. Olsen, P. Jørgensen, H. Koch, A. Balkova, and R.J. Bartlett, *J. Chem. Phys*. **104**, 8007 (1996).

[57] H. Flyvbjerg and H.G. Petersen, *J. Chem. Phys*. **91**, 461 (1989).

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

TABLE I. Convergence of the energies resulting from the all-electron *i*-FCIQMC and CAD-FCIQMC calculations with $\Delta\tau$ = 0.0001 a.u. toward FCI at the equilibrium and two displaced geometries corresponding to a simultaneous stretching of both O-H bonds in the H_2O/cc-pVDZ molecule by factors of 1.5 and 2.[a] The *i*-FCIQMC calculations were initiated by placing 1,000 walkers on the RHF determinant and the n_a parameter of the initiator algorithm[26] was set at 3.

Iterations	$R_{O\text{-}H} = R_e$		$R_{O\text{-}H} = 1.5\ R_e$		$R_{O\text{-}H} = 2.0\ R_e$	
	CAD-FCIQMC	*i*-FCIQMC	CAD-FCIQMC	*i*-FCIQMC	CAD-FCIQMC	*i*-FCIQMC
0	3.744[b]	217.821[c]	10.043[b]	269.961[c]	22.032[b]	363.954[c]
10,000	0.611	8.381	3.335	45.802	12.351	119.896
20,000	-0.073	1.596	2.597	18.345	8.485	72.650
30,000	-0.175	0.586	0.473	15.937	4.794	49.203
40,000	-0.211	-2.217	0.873	4.855	0.138	44.627
50,000	-0.440	1.456	1.310	3.247	-1.693	31.448
60,000	-0.046	1.911	0.501	0.588	-0.225	19.660
70,000	-0.235	0.302	0.685	0.811	-0.377	11.333
80,000	0.189	-0.686	0.063	-0.128	-0.425	12.611
90,000	-0.177	-0.981	-0.171	0.014	-1.657	10.089
100,000	-0.036	0.139	-0.302	1.956	-0.816	5.680
110,000	0.129	-0.710	-0.189	1.088	-0.580	4.797
120,000	-0.035	0.597	0.020	-0.241	-0.555	4.041
130,000	0.086	-0.503	-0.020	-0.720	-1.166	3.107
140,000	0.098	0.080	-0.084	0.497	-0.666	1.981
150,000	-0.055	0.308	-0.156	0.990	-0.620	1.630
160,000	0.078	-0.400	-0.059	-1.002	-0.434	1.328
∞	-76.241860[d]		-76.072348[d]		-75.951665[d]	

[a] The equilibrium geometry, $R_{O\text{-}H} = R_e$, and the geometries that represent a simultaneous stretching of both O-H bonds by factors of 1.5 and 2.0 without changing the ∠(H-O-H) angle were taken from Ref. 56. Unless otherwise stated, all energies are errors relative to FCI in millihartree.
[b] Equivalent to CCSD.
[c] Equivalent to RHF.
[d] Total FCI energy taken from Ref. 56.

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

TABLE II. Statistical analysis of the *i*-FCIQMC and CAD-FCIQMC calculations for the water molecule shown in Table I and Fig. 1.

Iterations (I)	%T [a]	%Q [b]	f_I [c]	m_I [d]		σ_I [d]	
				CAD-FCIQMC	*i*-FCIQMC	CAD-FCIQMC	*i*-FCIQMC
				$R_{O\text{-}H} = R_e$			
0	0	0	0.09[e]	0.087	4.035	0.593	18.944
20,000	11.8	1.38	3.95	-0.036	0.109	0.128	1.471
40,000	16.4	2.28	6.52	-0.020	0.043	0.122	1.298
60,000	21.3	3.48	10.3	0.003	-0.004	0.101	1.171
80,000	26.5	5.12	16.1	0.013	-0.165	0.095	1.043
100,000	32.1	7.47	25.7	0.017	0.055	0.080	0.823
120,000	38.0	10.3	40.0	0.049	-0.099	0.053	0.716
140,000	44.2	14.0	63.5	0.029	-0.075	0.055	0.354
				$R_{O\text{-}H} = 1.5\ R_e$			
0	0	0	0.03[e]	0.853	10.496	1.816	29.168
20,000	12.4	1.58	1.72	0.280	2.528	0.476	4.511
40,000	17.9	2.85	3.33	0.202	0.914	0.385	1.976
60,000	24.7	4.77	6.15	0.091	0.360	0.272	1.355
80,000	31.5	7.46	11.1	0.019	0.067	0.215	0.996
100,000	38.6	11.1	19.6	-0.019	-0.081	0.155	0.983
120,000	46.1	15.7	34.2	0.021	-0.527	0.139	0.744
140,000	53.6	21.1	58.7	0.062	-0.346	0.168	0.799
				$R_{O\text{-}H} = 2.0\ R_e$			
0	0	0	0.01[e]	1.645	32.718	5.095	50.838
20,000	10.6	1.34	0.55	-0.063	17.098	2.035	18.864
40,000	16.6	2.72	1.29	-0.721	10.707	0.730	10.739
60,000	24.5	5.03	2.85	-0.776	6.492	0.468	4.981
80,000	33.2	8.46	6.02	-0.709	4.574	0.356	3.268
100,000	42.5	13.2	12.4	-0.619	2.951	0.298	1.285
120,000	51.7	19.2	25.1	-0.711	2.241	0.253	0.814
140,000	60.9	26.6	50.3	-0.529	1.651	0.154	0.229

[a] Percentages of triply excited determinants captured during the *i*-FCIQMC propagations.

[b] Percentages of quadruply excited determinants captured during the *i*-FCIQMC propagations.

[c] Walker populations characterizing the *i*-FCIQMC propagations reported as percentages of the total walker numbers at $I = I_{max} = 160{,}000$, which in the specific *i*-FCIQMC runs shown in Table I and Fig. 1 were 1,169,396 at $R_{O\text{-}H} = R_e$, 3,327,438 at $R_{O\text{-}H} = 1.5\ R_e$, and 10,146,724 at $R_{O\text{-}H} = 2.0\ R_e$.

[d] The mean signed error (m_I) and standard deviation (σ_I) values relative to FCI, in millihartree, calculated using subsets of either the *i*-FCIQMC or CAD-FCIQMC energies that were obtained at the MC iterations $I, I + \Delta I, I + 2\Delta I, \ldots, I_{max}$, where $I_{max} = 160{,}000$ is the final iteration at which the *i*-FCIQMC propagations were stopped and ΔI is an iteration increment used to select the energy points for statistical analysis. In the case of the calculations reported in this work, $\Delta I = 20$ for *i*-FCIQMC and 1,000 for CAD-FCIQMC (although the *i*-FCIQMC energies were reported once every 20 MC iterations, the FCIQMC-corrected CCSD calculations were performed only once every 1,000 iterations). The $m_{I=0}$ and $\sigma_{I=0}$ values correspond to averaging the entire energy dataset characterizing a given method and geometry. The m_I and σ_I values for $I > 0$ correspond to averaging a subset of energy points that starts at the MC iteration I and ends up at the last iteration I_{max}.

[e] The initial walker population, meaning 1,000 walkers placed on the RHF reference determinant.

ACCEPTED MANUSCRIPT

AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

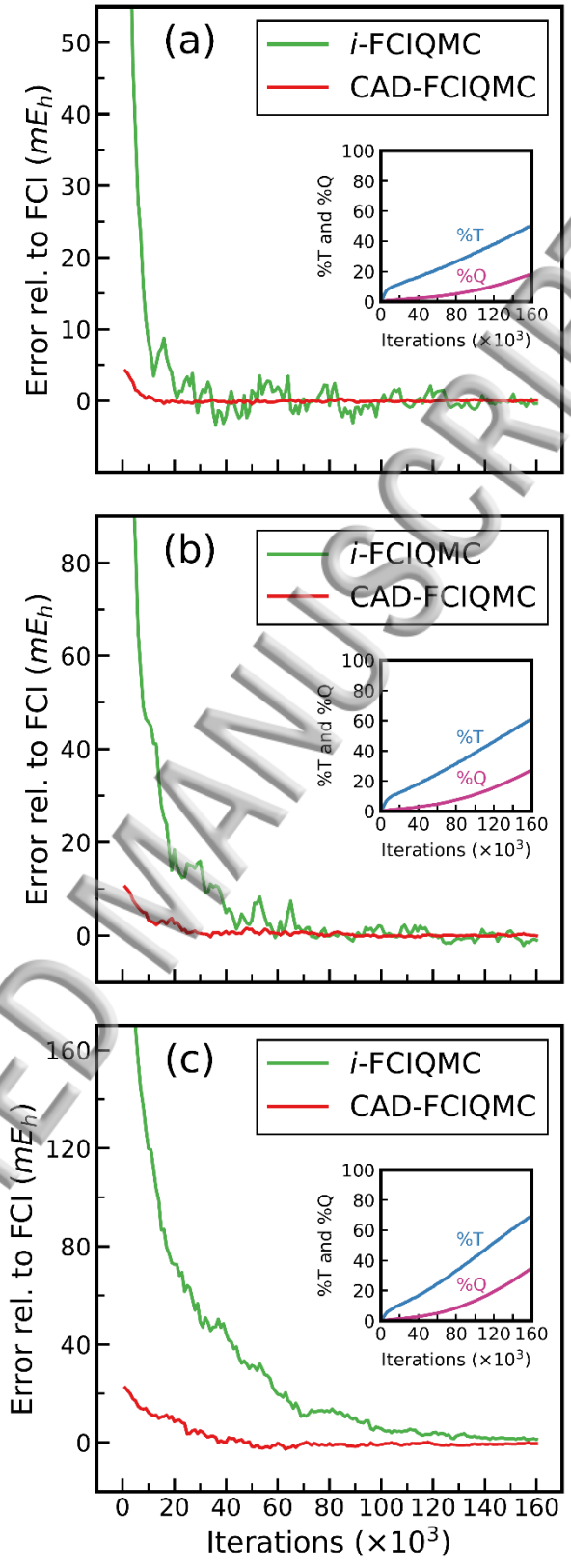


FIG. 1.

ACCEPTED MANUSCRIPT

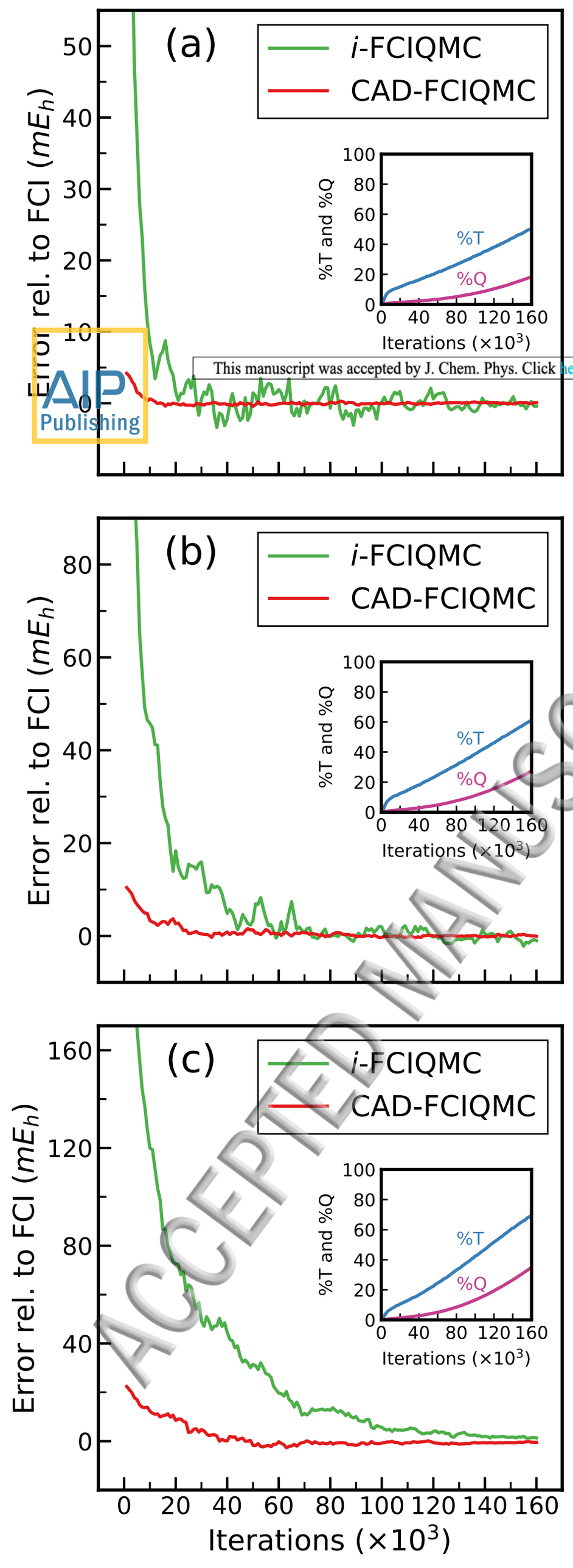
AIP Publishing

This manuscript was accepted by J. Chem. Phys. Click here to see the version of record.

Figure caption

FIG. 1. Convergence of the energies resulting from the all-electron *i*-FCIQMC and CAD-FCIQMC calculations with $\Delta\tau$ = 0.0001 a.u. toward FCI at the equilibrium and two displaced geometries corresponding to a simultaneous stretching of both O-H bonds in the H_2O/cc-pVDZ molecule by factors of 1.5 and 2 taken from Ref. 56. Panel (a) corresponds to the equilibrium geometry. Panels (b) and (c) correspond to the geometries that represent a simultaneous stretching of both O-H bonds by factors of 1.5 and 2.0, respectively, without changing the ∠(H-O-H) angle. The *i*-FCIQMC calculations were initiated by placing 1,000 walkers on the RHF determinant and the n_a parameter of the initiator algorithm[26] was set at 3. All energies are errors relative to FCI in millihartree, and the insets show the percentages of triply (%T) and quadruply (%Q) excited determinants captured during the *i*-FCIQMC propagations.

ACCEPTED MANUSCRIPT



AIP Publishing
This manuscript was accepted by J. Chem. Phys. Click he
ACCEPTED MANUSCRIPT