



Title	Selected configuration interaction method using sampled first-order corrections to wave functions
Author(s)	Ohtsuka, Yuhki; Hasegawa, Junya
Citation	Journal of Chemical Physics 147(3)034102 https://doi.org/10.1063/1.4993214
Issue Date	2017-07-17
Doc URL	http://hdl.handle.net/2115/70978
Rights	The following article appeared in J Chem Phys 147(3)4102(2017) and may be found at http://dx.doi.org/10.1063/1.4993214
Type	article
File Information	14993214.pdf



[Instructions for use](#)

Selected configuration interaction method using sampled first-order corrections to wave functions

Yuhki Ohtsuka and Jun-ya Hasegawa

Citation: *The Journal of Chemical Physics* **147**, 034102 (2017); doi: 10.1063/1.4993214

View online: <http://dx.doi.org/10.1063/1.4993214>

View Table of Contents: <http://aip.scitation.org/toc/jcp/147/3>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Hybrid stochastic-deterministic calculation of the second-order perturbative contribution of multireference perturbation theory](#)

The Journal of Chemical Physics **147**, 034101 (2017); 10.1063/1.4992127

[Incremental full configuration interaction](#)

The Journal of Chemical Physics **146**, 104102 (2017); 10.1063/1.4977727

[Time-dependent N-electron valence perturbation theory with matrix product state reference wavefunctions for large active spaces and basis sets: Applications to the chromium dimer and all-trans polyenes](#)

The Journal of Chemical Physics **146**, 244102 (2017); 10.1063/1.4986975

[A new near-linear scaling, efficient and accurate, open-shell domain-based local pair natural orbital coupled cluster singles and doubles theory](#)

The Journal of Chemical Physics **146**, 164105 (2017); 10.1063/1.4981521

[Coupled cluster valence bond theory for open-shell systems with application to very long range strong correlation in a polycarbene dimer](#)

The Journal of Chemical Physics **147**, 024107 (2017); 10.1063/1.4991797

[Stochastic multi-reference perturbation theory with application to the linearized coupled cluster method](#)

The Journal of Chemical Physics **146**, 044107 (2017); 10.1063/1.4974177



**COMPLETELY
REDESIGNED!**

**PHYSICS
TODAY**

Physics Today Buyer's Guide
Search with a purpose.

Selected configuration interaction method using sampled first-order corrections to wave functions

Yuhki Ohtsuka^{1,a)} and Jun-ya Hasegawa^{1,2}

¹Institute for Catalysis, Hokkaido University, N21 W10 Kita-ku, Sapporo, Hokkaido 001-0021, Japan

²JST-CREST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

(Received 18 April 2017; accepted 22 June 2017; published online 17 July 2017)

A new selected configuration interaction (CI) method was proposed for the potential energy surfaces of quasi-degenerate and excited states. Slater determinants are generated by sampling the first-order corrections to the target-state wave functions using the quantum Monte Carlo method in determinant space. As in the Monte Carlo (MC) CI method, the wave function is improved at each iteration by generating new determinants and applying a pruning step. Compared to the random generation in the MCCI calculations, the number of iterations before convergence is significantly reduced. Regarding the potential energy curves of the ground and excited states of C_2 , the non-parallelity errors were sufficiently small, thus indicating the method's applicability to the calculations of potential energy surfaces. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4993214>]

I. INTRODUCTION

Wave functions become multi-configurational due to the static electron correlations in the quasi-degenerate and excited states. For these complicated electronic structures, the complete active space self-consistent field (CASSCF) method¹ has been employed as a standard theoretical model. However, the problem in the CASSCF calculation is that the size of the active space is limited because the number of electronic configurations increases factorially with the number of electrons and orbitals in the active space. Therefore, a method that can treat a large active space is greatly desired in the computational chemistry community.

One solution for the active space problem is to select important configurations from a large active space. In the restricted active space (RAS)² and generalized active space (GAS)³ methods, the selection is made by restricting the number of electrons in each subspace. The perturbation selection method⁴ uses a second-order perturbation energy correction, and configurations that exceed a predefined energy threshold are selected. This selection method was used in the multi-reference double (MRD) excitation-configuration interaction (CI)⁵ and symmetry adapted cluster (SAC)-CI⁶ methods, and was successfully applied to various systems in the ground and excited states.^{7,8} In a multi-configurational method (configuration interaction by perturbation with a multiconfigurational zeroth-order wave function selected by an iterative process, CIPSI),⁹ the configuration space is expanded iteratively by adding important configurations in the first-order corrections (FOCs) to the wave functions. The Monte Carlo (MC) method is also used for the configuration selection. In the MCCI method,^{10,11} new configurations are selected using the Monte Carlo technique from the first-order interacting space (FOIS).

In the quantum Monte Carlo (QMC) methods in the determinant space,^{12–14} configurations are selected by MC simulation of the imaginary-time propagation of the wave function. Among them, the full CI (FCI) QMC method has been actively developed and extended to quasi-degenerate electronic states and excited states.^{15–18} The algorithm of FCIQMC is also used for the estimation of the second-order energy correction in the heat bath CI method.^{19,20} Recently, the energy difference between configurations has been used as the selection criterion. In the lambda (Λ) CI method, the energy difference from the lowest configurations is evaluated.²¹ The adaptive CI (ACI) method,²² which was proposed by the same authors as the Λ CI, employs the energy and coefficient thresholds for selecting configurations and successfully controls the accuracy of the final results.

Because the configuration space is expanded iteratively, MCCI is applicable to quasi-degenerate states that require large active spaces. MCCI has been applied to the calculations of various potential energy curves (PECs)^{23–26} and dissociation energies.²⁷ In the result, MCCI reproduced the full CI (FCI) results from the equilibrium to dissociated bond length with similar accuracy by using many fewer configurations. MCCI was also applied to the excited states for small molecules.²⁸ The accuracy in the excited-state calculations was similar to that in the ground state, and the excitation energies are in good agreement with the FCI values. However, due to the random sampling of new configurations, the wave functions converge very slowly, and many iterations are necessary.

In this article, we introduce the QMC method in the determinant space to a new selected CI method, the MC correction CI (MCCCI or MC3I) method. Because the FOCs to all target states are sampled and the group sum of the sampled determinants is added to the determinant space, the iterative process is expected to converge much faster than MCCI for both the ground and excited state calculations. After the diagonalization

^{a)} Author to whom correspondence should be addressed: ohtsuka@cat.hokudai.ac.jp

of the Hamiltonian matrix in the space spanned by the sampled determinants, the less important determinants are pruned to reduce the number of determinants without losing much accuracy. By repeating the generation of new determinants and the pruning of the less important determinants, the wave function is improved iteratively. The idea of the iterative improvement of the determinant space is similar to the one adopted in CIPSI. The present MC3I could be considered as a stochastic version of CIPSI. The convergence of MC3I becomes slower than that of CIPSI due to the MC sampling of new determinants. However, the present MC3I method is applicable to systems that are larger than those which can use CIPSI because the MC method is less sensitive to the size of the active space.

The present paper has the following structure. In Sec. II, a theoretical description is provided. After the computational details are provided in Sec. III, the results are discussed in Sec. IV. To assess the accuracy and convergence, we applied the MC3I method to the ground state of H₂O and the ground and excited states of C₂. The sampling efficiency was determined by comparing the sampled and exact FOCs to the Hartree-Fock (HF) wave function of H₂O. The energy convergence of the MC3I calculations was examined by comparing the number of iterations before convergence with the MCCI calculations for the ground and excited states in the same condition. The applicability to dissociation was investigated using the non-parallelity errors (NPEs) of the PECs of C₂ for the ground and excited states. A summary of the study is provided in Sec. V.

II. THEORY

In the MC3I method, the walker distribution on the determinants is generated from the CI wave function,

$$|\Psi_K\rangle = C_{K,0} |\Phi_0\rangle + C_{K,1} |\Phi_1\rangle + C_{K,2} |\Phi_2\rangle + \cdots + C_{K,N} |\Phi_N\rangle, \quad (1)$$

where $\{C_{K,I}\}$ denotes the CI coefficient of the determinant Φ_I for the K th state. The wave function in Eq. (1) is redefined in terms of the number of walkers,

$$|\Psi'_K\rangle = R_{K,0} |\Phi_0\rangle + R_{K,1} |\Phi_1\rangle + R_{K,2} |\Phi_2\rangle + \cdots + R_{K,N} |\Phi_N\rangle, \quad (2)$$

where $\{R_{K,I}\}$ represents the number of walkers on determinant $|\Phi_I\rangle$ and is proportional to the coefficient $\{C_{K,I}\}$,

$$R_{K,I} = C_{K,I} \times N_w, \quad (3)$$

where N_w is an integer parameter and determines the number of walkers on each determinant.

As shown in Fig. 1, the FOCs to the wave functions in the preceding iteration are generated by using MC sampling for target states that involve not only the ground state but also excited states. The quality of the sampled FOC can be improved as N_w increases because more determinants are generated by the sampling. The group sum of the newly generated determinants and the selected determinants in the preceding iteration is the selected determinants in the present iteration.

One of the crucial parts of the MC3I method is the selection of the determinants from those generated by the MC sampling. Our idea is to invoke perturbation theory for improving the sampling. As the zero-th order wave function, we assume the CI wave functions in the preceding iteration which reasonably approximate the target state. In perturbation theory, the FOC to the K th state wave function, $|\Psi_K\rangle$, is defined as

$$|\Psi_K^{(1)}\rangle = \sum_I \frac{\langle \Phi_I | \hat{H} | \Psi_K \rangle}{E_K - \langle \Phi_I | \hat{H} | \Phi_I \rangle} |\Phi_I\rangle = \sum_I C_{K,I}^{(1)} |\Phi_I\rangle. \quad (4)$$

Here, E_K is the energy of $|\Psi_K\rangle$. Φ_I runs over the determinants that interact with $|\Psi_K\rangle$. Our key idea is to use the stochastically evaluated coefficient $\{C_{K,I}^{(1)}\}$ to select the determinants. As shown in Fig. 2, the Monte Carlo simulation is performed from the walker distribution which represents the CI wave function in Eq. (2) to obtain the FOC to the CI wave function.

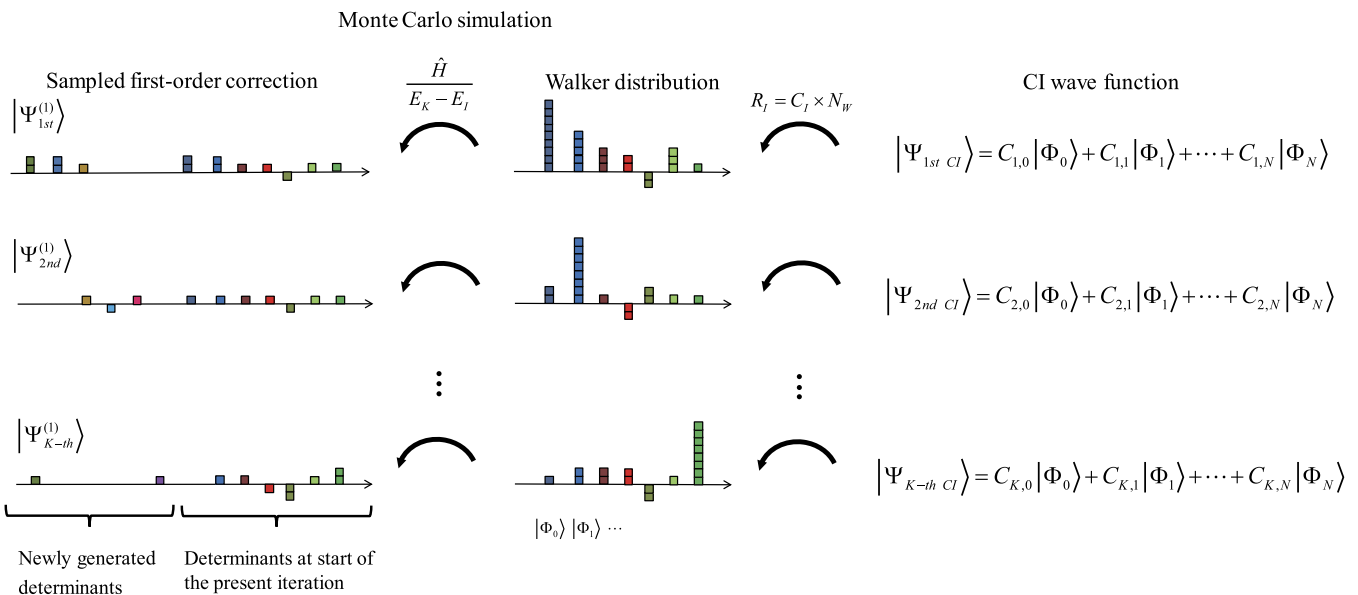


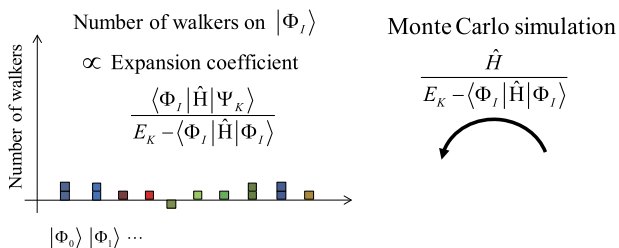
FIG. 1. Schematic representation of sampling of first-order corrections to wave functions by the QMC method in the determinant space.

First-order correction to Ψ_K

$$\Psi_K^{(1)} = \sum_I \frac{\langle \Phi_I | \hat{H} | \Psi_K \rangle}{E_K - \langle \Phi_I | \hat{H} | \Phi_I \rangle} |\Phi_I\rangle \quad \xleftarrow{\text{Perturbation}}$$

K-th state CI wave function

$$\Psi_K = \sum_I C_I |\Phi_I\rangle$$

Sampled first-order correction to Ψ_K 

The determinant whose expansion coefficient in the FOC is larger is sampled more frequently, and the resultant number of walkers on the determinant Φ_I corresponds to the expansion coefficient, $C_I^{(1)}$, in Eq. (4)

In the below discussion, we explain a part of the selection procedure in the case of a single walker on the determinant $|\Phi_I\rangle$ of the K th state. First, a determinant $|\Phi_I\rangle$ is selected randomly from the FOIS of $|\Phi_I\rangle$. The probability of the candidate $|\Phi_I\rangle$ being selected is $(N_I)^{-1}$, where N_I is the number of determinants that interact with $|\Phi_I\rangle$ ($|\Phi_I\rangle$ itself is excluded).

Next, the transition probability of a walker moving from $|\Phi_I\rangle$ to $|\Phi_J\rangle$, $P_{Transition}$, is calculated as follows:

$$Q_{Transition} = \frac{S \times r_I \times N_I \times \langle \Phi_J | \hat{H} | \Phi_I \rangle}{(E_K - \langle \Phi_J | \hat{H} | \Phi_J \rangle)}, \quad (5)$$

$$P_{Transition} = |Q_{Transition}|. \quad (6)$$

Here, S is a scale parameter that controls the acceptance ratio for the transition [usually of a similar magnitude to $(N_I)^{-1}$] and r_I is a weight parameter that is introduced to account for the decimal part of $R_{K,I}$, the number of walkers on $|\Phi_I\rangle$. Because $R_{K,I}$ is a real number, as defined in Eq. (3), the magnitude of the decimal part is also considered in the sampling procedure. The trial of the transition is repeated $\text{Int}(|R_{K,I}|) + 1$ times²⁹ with $r_I = 1.0$ and with $r_I = |R_{K,I} - \text{Int}(R_{K,I})|$ for the last trial.

A random number (between 0.0 and 1.0) is provided. If the random number is smaller than the transition probability $P_{Transition}$ in Eq. (6), the transition is accepted, and the walker on $|\Phi_I\rangle$ is moved to $|\Phi_J\rangle$. The number of walkers on $|\Phi_J\rangle$ is thus increased by $1.0 \times \text{sgn}(R_I \times Q_{Transition})$, where $\text{sgn}(R_I \times Q_{Transition})$ represents the sign of $R_I \times Q_{Transition}$. If $|\Phi_J\rangle$ does not exist, it is added to the determinant space. When $P_{Transition}$ exceeds 1.0, the transition is unconditionally accepted, and the number of walkers on $|\Phi_J\rangle$ is increased by $\text{sgn}(R_I) \times Q_{Transition}$. These procedures are applied to all of the walkers in all states in one iteration. As shown in Fig. 1, the determinant space of the CI wave function is augmented by the sampled determinants in the FOIS and spanned by the group

FIG. 2. Relation between the sampled CI wave function and sampled first-order corrections to CI wave function.

sum of the determinants for all of the target states. The CI wave functions are updated by solving the eigenvalue equations with the Davidson diagonalization.³⁰

Each iteration ends with the pruning step. Checking the CI vector, the less important determinants whose absolute coefficients are smaller than the coefficient threshold, c_{\min} , in all target states are pruned. The remaining CI vectors after the pruning step are orthonormalized and used to construct the walker distributions in the next iteration. The generation of new determinants, Davidson diagonalization, and removal of less important determinants by the coefficient threshold are repeated until stationary wave functions are obtained as in the MCCI calculations.

We summarize the present algorithm of the MC3I calculation.

- Step (i): The number of solutions, Abelian space symmetry, active space, and S_z value of the Slater determinants are specified for the MC3I wave functions. The initial reference wave functions for the target states are obtained by the HF or complete active space (CAS) CI calculations.
- Step (ii): Construction of the walker distributions from the CI wave functions. The number of walkers on $|\Phi_I\rangle$ is defined.
- Step (iii): From the walker distributions given by the previous step, the MC simulation is performed to sample the FOCs to the target states. The determinant space is spanned by the group sum of the determinants that are augmented using the MC sampling.
- Step (iv): The CI wave functions of the target states are updated by solving the eigenvalue problem with the Davidson diagonalization. The initial vectors to be used for the Davidson diagonalization are taken from the solutions in the preceding iteration or the CI vectors in the first iteration.
- Step (v): The pruning step in the MCCI method is carried out. The less important determinants whose absolute coefficients are smaller than the

threshold c_{\min} in all solutions are discarded, and the resultant CI vectors are used to construct the walker distribution in the next iteration. The vectors are orthonormalized to each other and used for the initial vectors of the Davidson diagonalization for the next iteration.

In the present algorithm, the procedures in Steps (ii)–(v) are repeated until the iteration number reaches a pre-defined number. The number of iterations should be large enough to obtain a stationary energy and stationary number of determinants after the pruning step.

III. COMPUTATIONAL DETAILS

We are developing the MC3I and MCCI programs as a part of SMASH open source software.³¹ The present development code is interfaced with the GAMESS program package,^{32,33} from which the molecular integrals are imported. The present algorithm treats all of the molecular integrals in the core. The Hamiltonian matrix elements, which are necessary for the calculation of the transition probabilities of the walkers and for the Davidson diagonalization, are calculated on the fly. The formation of the sigma vectors, which is the most time-consuming part of the MC3I calculations, has been parallelized using OpenMP. The computational costs of the sigma vector calculation are proportional to the square of the number of determinants in our algorithm. Because we cannot predict which determinants are selected, the sigma vectors are calculated from the matrix elements between all pairs of the determinants. The MC3I and MCCI energies in a calculation are those at the last iteration. The final MC3I and MCCI energies are defined as the average energies over ten calculations with different seeds for the random number generator. The standard deviations were also estimated from the results of the ten calculations. For the random number generator, the Mersenne twister method was used.³⁴ The number of iterations is 1000 and 500 for H₂O and C₂, respectively. The scale parameter S in Eq. (5) is fixed as 0.01 and 0.001 for H₂O and C₂, respectively.

In Sec. IV C, the number of determinants in the MCCI calculations is controlled by the branching factor f , which is the number of determinants generated from one determinant in one iteration. We first performed the MC3I calculations and recorded the number of newly generated determinants in all iterations. In the succeeding MCCI calculations, the number of generated new determinants was very similar to that for the MC3I calculation in each iteration by varying $f = (\text{Number of newly generated determinants in the present iteration of MC3I calculation}) / (\text{Number of determinants at the start of the present iteration in MCCI})$. The number of determinants in the MCCI calculation was not exactly the same as that obtained using the corresponding MC3I calculation because the determinants were generated stochastically using random numbers.

To evaluate the sampling efficiencies of the MCCI and MC3I algorithms, we performed the so-called full pruning at each iteration. Full pruning means that all determinants are subjected to pruning in that iteration. We note that in standard MCCI calculations, only newly generated determinants in the iteration are pruned if their absolute coefficients are smaller

than a threshold c_{\min} to prevent the duplicate sampling of the same determinants. In a certain iteration (usually every ten iterations), full pruning is performed. The combination of the pruning of only the newly generated determinants and full pruning will improve the sampling efficiency in the MCCI and MC3I calculations. However, to illustrate the difference in the sampling efficiencies between MCCI and MC3I, we performed full pruning in all iterations.

IV. RESULTS AND DISCUSSION

A. Sampling the FOCs to the wave functions of the target states

One of the important points relevant to the efficiency of the MC3I method is the quality of the FOC after the MC sampling. The convergence of the MC3I wave functions will be accelerated if the important determinants are selected efficiently. To evaluate the accuracy, the sampled FOCs are compared with the exact ones given by $C_{K,I}^{(1)}$ in Eq. (4). If N_w reaches infinity, the relative magnitude of the sampled FOCs approach the exact result. A benchmark calculation was carried out for H₂O at the equilibrium geometry that was employed in Ref. 35, an O–H distance of 1.843 45 Bohrs and a HOH angle of 110.6°. For the basis sets, the cc-pVDZ sets³⁶ were used. In Fig. 3, the sampled FOCs to the HF wave function using $N_w = 2000, 5000,$ and $10\,000$ are compared with the exact FOC. The vertical axis shows the number of walkers on $|\Phi_I\rangle$ for the sampled FOC and the $C_{HF,I}^{(1)}$ coefficient for the exact FOC. The horizontal axis is the labels of the determinants, and its maximum value is 3415, which corresponds to the number of singly and doubly excited determinants from the closed-shell HF determinant in

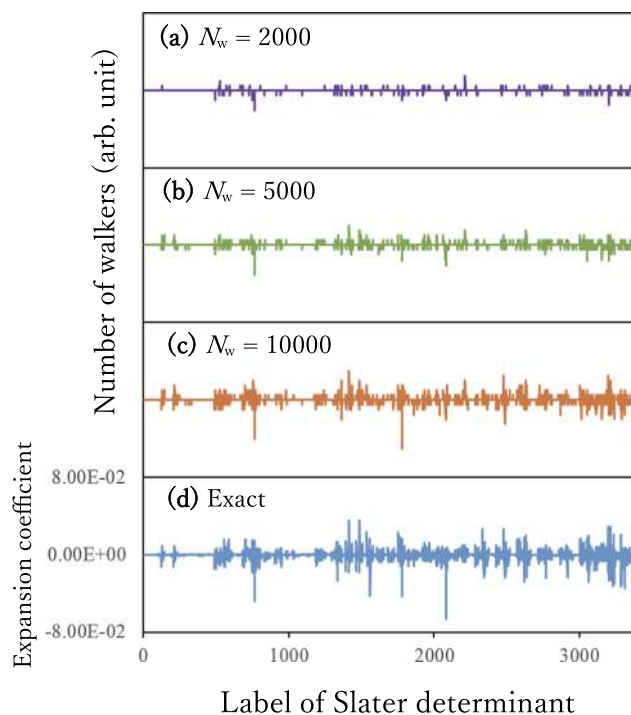


FIG. 3. Comparison between exact and sampled first-order corrections to the HF determinant of H₂O. ((a), (b), and (c)) Sampled first-order corrections using $N_w = 2000, 5000,$ and $10\,000$, respectively. (d) Exact first-order correction.

A_1 symmetry. As shown in Fig. 3, the shape of the resultant walker distribution became closer to the exact shape as N_w increased. The average number of sampled determinants over ten MC simulations was 150, 344, and 582 with $N_w = 2000$, 5000, and 10 000, respectively. The quality of the sampled determinants was promising; in the case of $N_w = 2000$, 46% of the sampled determinants had the largest 10% of the coefficients in the exact FOC. In contrast, with random sampling, only 10% of the sampled determinants should have the largest 10% of the coefficients.

With the selected sets of determinants, the CI energy was evaluated. This energy corresponds to that at the first iteration of the MC3I and MCCI calculations. The average CI energies with $N_w = 2000$, 5000, and 10 000 are $-76.080(6)$, $-76.125(4)$, and $-76.163(3)$ E_h , respectively. These energies were lower than those obtained with a very similar number of determinants selected by random sampling, $-76.036(3)$, $-76.050(3)$, and $-76.066(5)$ E_h , respectively. The variational principle tells that the FOIS selected by the sampling of the FOCs is much more appropriate for a CI wave function than one selected by random sampling. As shown in Sec. IV C, the convergence behavior of the MC3I and MCCI calculations clearly reflects this difference.

B. Dependence of energy convergence on N_w and c_{\min}

Next, we discuss how the energy convergence depends on the walker parameter, N_w , and the coefficient parameter, c_{\min} . The examination was carried out for the ground state of H_2O at the equilibrium structure. Figure 4 shows the results of the MC3I calculations with three different walker parameters, $N_w = 2 \times 10^3$, 5×10^3 , and 1×10^4 , and three different coefficient thresholds, $c_{\min} = 1 \times 10^{-3}$, 5×10^{-4} , and 2×10^{-4} . The HF wave function was used for the initial wave function. For each c_{\min} , the energy convergence became faster as N_w increased because the number of sampled determinants in each iteration increases with N_w . Consequently, important determinants are selected at earlier iterations.

As shown in Fig. 4 and Table I, the converged energy is sensitive to the coefficient threshold c_{\min} but is less dependent on N_w . For example, the energies using $c_{\min} = 5 \times 10^{-4}$ were $-76.232\ 327(9)$, $-76.232\ 42(1)$, and $-76.232\ 59(2)$ E_h for $N_w = 2000$, 5000, and 10 000, respectively. Although the number of determinants before the pruning step increased with

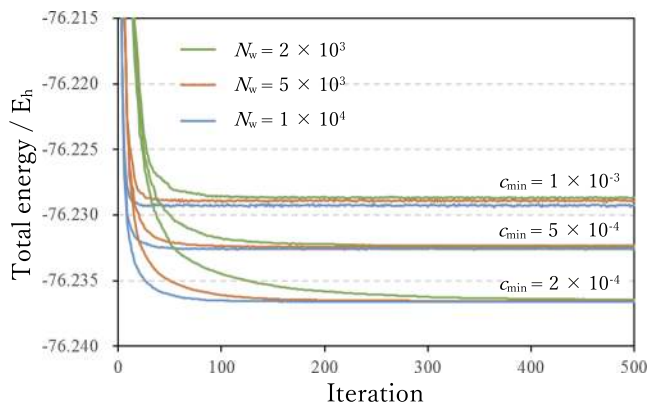


FIG. 4. MC3I energy convergence using various c_{\min} and N_w of H_2O for the ground state at the equilibrium bond length.

TABLE I. The MC3I result for H_2O for the ground state at the equilibrium bond length (E_h) with walker parameters (N_w) and coefficient thresholds (c_{\min}). The standard deviations over ten calculations are shown in parentheses. FCI energy = $-76.241\ 860$ (E_h).³⁵

c_{\min}	N_w		
	2 000	5 000	10 000
1×10^{-3}	$-76.228\ 66(2)$	$-76.228\ 91(4)$	$-76.229\ 26(5)$
5×10^{-4}	$-76.232\ 327(9)$	$-76.232\ 42(1)$	$-76.232\ 59(2)$
2×10^{-4}	$-76.236\ 508(4)$	$-76.236\ 539(3)$	$-76.236\ 587(5)$
1×10^{-4}	$-76.238\ 831(2)$	$-76.238\ 880(2)$	$-76.238\ 897(2)$

N_w (see Table II), the number of selected determinants at the convergence was very similar when a common c_{\min} value was used. The number of determinants after the pruning step at the last iteration was approximately 1660, 2870, 9200, and 26 000 for $c_{\min} = 1 \times 10^{-3}$, 5×10^{-4} , 2×10^{-4} , and 1×10^{-4} , respectively.

The result indicates that the total energy and the number of iterations necessary for the convergence are controlled using c_{\min} and N_w , respectively. We note that a smaller c_{\min} increases the computational time. With the smaller c_{\min} , more determinants should be sampled, and more determinants must be included in the CI diagonalization. To determine the proper c_{\min} , we first perform a MC3I calculation with $c_{\min} = 5 \times 10^{-4}$, which is a typical value in MCCI calculations, and decrease c_{\min} gradually till the desired accuracy is achieved. Subsequently, N_w is chosen so that the number of newly generated determinants in each iteration is on the same order as the number of determinants at the start of the iteration.

C. Comparison of energy convergence between MC3I and MCCI calculations

Here, the result of the MC3I calculations is compared with that of the MCCI calculations. Figure 5 shows the convergence curves of ten MCCI and ten MC3I calculations of H_2O at the equilibrium bond length (R_e) and $3.0 R_e$ (the two O–H bonds are simultaneously stretched three times longer than the equilibrium distance) with $c_{\min} = 5.0 \times 10^{-4}$ and $N_w = 10\ 000$ using different seeds for the random number generator. For each iteration, the number of newly generated determinants in the MCCI calculations was kept very close to that of the corresponding MC3I calculation.

The most important finding is that all of the MC3I calculations converged much more quickly than did the MCCI calculations. For O–H bond lengths = R_e , it took only 16.2

TABLE II. Number of determinants at the last iteration in the MC3I calculation of H_2O for the ground state at the equilibrium bond length before the pruning step. Number of determinants in FCI = 451 681 246.

c_{\min}	N_w		
	2 000	5 000	10 000
1×10^{-3}	2 219	3 043	4 369
5×10^{-4}	3 456	4 323	5 743
2×10^{-4}	9 825	10 767	12 301
1×10^{-4}	26 226	27 505	29 110

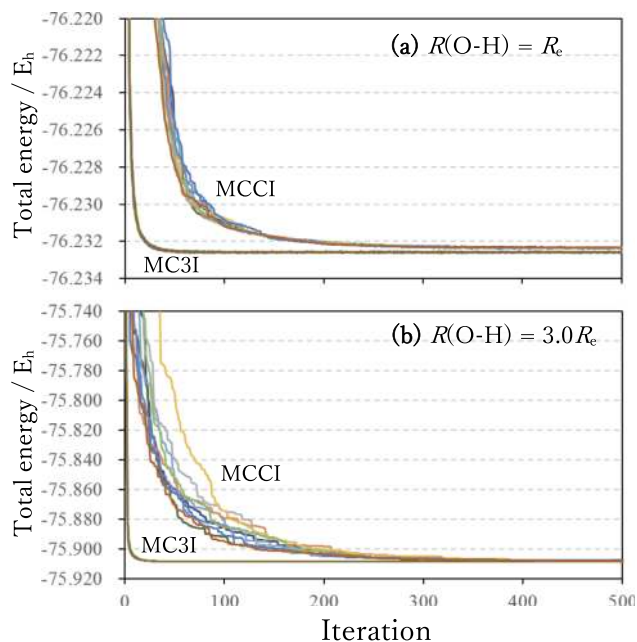


FIG. 5. Energy convergences in ten MC3I and ten MCCI calculations of H_2O for the ground state at (a) the equilibrium bond length (R_e) and (b) O–H bond length = $3.0R_e$ with different seeds for the random number generator.

iterations on average until the energy difference from the converged energy became less than $1 mE_h$ in the MC3I calculations, whereas 116.9 iterations were required in the MCCI calculations. In addition, the convergence behavior is much less dependent on random numbers. As shown in Fig. 5(a), the ten MC3I convergence curves overlap with each other very well. The energy at convergence should be mentioned; although a common c_{min} was used, the converged MC3I energy, $-76.23259(2) E_h$, was slightly lower than the MCCI energy, $-76.232362(9) E_h$, because more of the important determinants were sampled in one iteration in the MC3I calculations.

The reason for the fast convergence behavior of the MC3I calculation is that the important determinants are selected at earlier iterations. Figure 6 shows the number of determinants

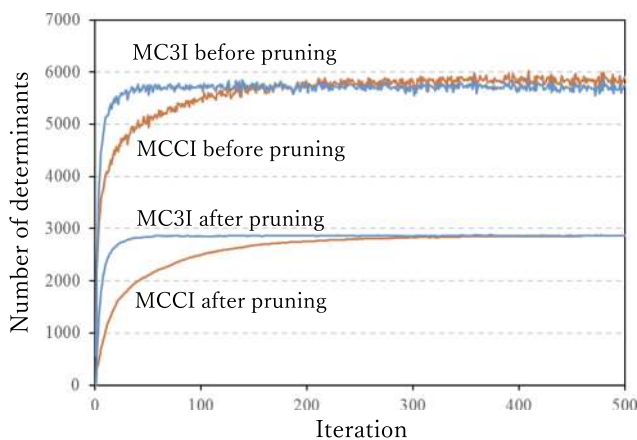


FIG. 6. Number of determinants before and after the pruning step in MC3I and MCCI calculations of H_2O for the ground state at the equilibrium bond length. The fluctuation behavior in the number of determinants before the pruning step originates from the stochastic nature of the MC sampling using random numbers.

before and after the pruning step in each iteration. The MC3I case is compared with the MCCI case for H_2O , as shown in Fig. 5(a). The number of newly generated determinants in each iteration was controlled to be similar to each other in the MC3I and MCCI calculations. However, the number of determinants before and after the pruning step in the MC3I calculations was larger than those in MCCI at the early stage of the calculations. Because more important determinants were generated and approved in the MC3I algorithm, the number of determinants after pruning is larger than that in the MCCI algorithm. At the converging limit, the determinants after the pruning step in the MCCI and MC3I calculations became almost the same because all of the important determinants had been selected using the common coefficient threshold c_{min} .

The new sampling method also improves the convergence of the multi-reference wave function at elongated bond lengths. Figure 5(b) shows the results of ten MC3I and ten MCCI calculations for O–H bond lengths = $3.0R_e$. The calculations also started from the HF determinant. In the MCCI calculations, 375.5 iterations were required until the energy difference from the converged energy became less than $1 mE_h$, while only 18.1 iterations were required to achieve the same accuracy in the MC3I calculations.

As for the computational costs, the sampling in the MC3I algorithm takes more time than the random sampling in MCCI because the FOC [Eq. (4)] is evaluated. However, the computing time for the sampling of the FOC is negligibly smaller than that for the Davidson diagonalization of the Hamiltonian matrix. In the MC3I calculations of H_2O with $c_{min} = 5.0 \times 10^{-4}$ and $N_w = 10\,000$, only 0.3% of the total wall-clock time was used for sampling the FOC. Therefore, the computational costs of MCCI and MC3I calculations are proportional to $\sum_I N_I^2$, where N_I is the number of the determinants before the pruning step at iteration I , because the formation of sigma vectors in the diagonalization step dominates most of the computational time. For the calculations of Fig. 6, the computational cost of MC3I/MCCI ratio was 0.105.

D. PECs of the excited states of C_2

The MC3I method was also applied to the PECs of the excited states. The results for the lowest three singlet ($X^1\Sigma_g^+$, $B^1\Delta_g$, and $B'^1\Sigma_g^+$) states of the C_2 molecule were compared with the FCI values.³⁷ The 6-31G* basis set^{38,39} was used, and 1s orbitals were excluded from the active space. All of the three states have A_g symmetry in the D_{2h} point group and were calculated simultaneously in one MC3I calculation. The MC3I and MCCI calculations started from the CASCI(6e, 6o) wave functions. The lowest four states were calculated in all geometries because a quintet state is nearly degenerated with the $B'^1\Sigma_g^+$ state at the C–C bond length of 3.0 Å. We used four different c_{min} values to check the convergence in energy. With $c_{min} = 1.0 \times 10^{-3}$, 5.0×10^{-4} , 2.0×10^{-4} , and 2.0×10^{-5} , the number of the selected determinants before the pruning step ranged from 8127 to 14 643, from 15 340 to 27 383, from 41 604 to 65 902, and from 408 037 to 576 802, respectively. These selected determinants amount to only 0.028%, 0.052%, 0.126%, and 1.101% of the total number of determinants, 52 407 353, in the FCI wave function.

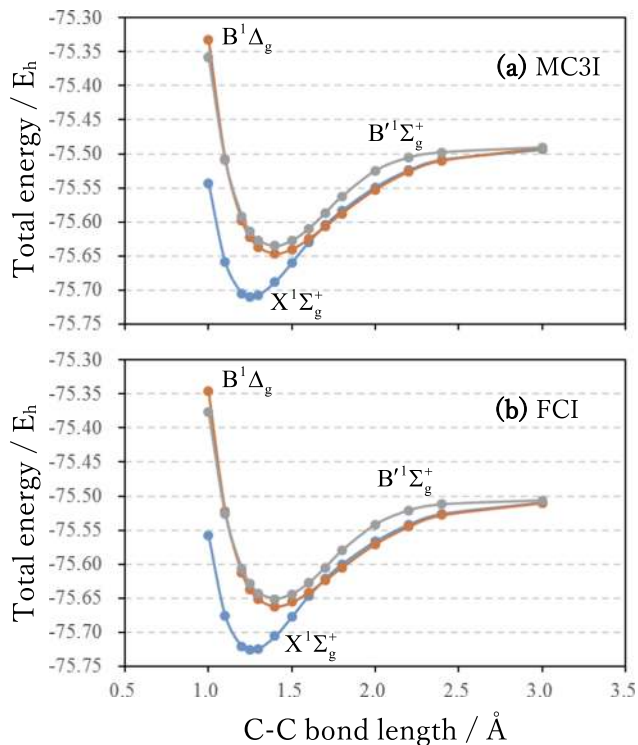


FIG. 7. Comparison between (a) MC3I with $c_{\min} = 1 \times 10^{-3}$ and $N_w = 10\,000$ and (b) FCI PECs of C_2 for the lowest three singlet states.

Figure 7 shows the comparison between the MC3I ($c_{\min} = 1.0 \times 10^{-3}$) and FCI PECs for the lowest three states. Regardless of the crude threshold for c_{\min} , the features of the curves were reproduced qualitatively. For instance, the $B^1\Delta_g$ state crosses with the $X^1\Sigma_g^+$ state between 1.6 and 1.7 Å and becomes the ground state at longer bond distances. In the FCI result, the energy of the $B^1\Delta_g$ state is lower than that of the $X^1\Sigma_g^+$ state by 0.062 mE_h at 3.0 Å, whereas the $X^1\Sigma_g^+$ state is slightly below the $B^1\Delta_g$ state by 2.00, 0.524, and 0.075 mE_h with $c_{\min} = 1.0 \times 10^{-3}$, 5.0×10^{-4} , and 2.0×10^{-4} , respectively, in the MC3I results.

The applicability of MC3I to the PECs should be evaluated using the non-parallelity error (NPE). Approximate methods are useful if the error from the exact energy is similar all through the potential energy surface. The NPE is defined as the difference between the maximum and minimum errors. In Fig. 8, the deviations of the MC3I energies from the FCI energies are plotted along the C–C bond length. The results with four different coefficient thresholds c_{\min} are compared. In Table III, the maximum error, minimum error, and NPE are summarized. As clearly shown in Fig. 8, the deviations were significantly reduced as c_{\min} became small. The maximum error of 18.9 mE_h with $c_{\min} = 1.0 \times 10^{-3}$ became 0.7 mE_h when $c_{\min} = 2.0 \times 10^{-5}$ was used. However, as Table III shows, the NPEs with $c_{\min} = 1.0 \times 10^{-3}$, 5.0×10^{-4} , 2.0×10^{-4} , and 2.0×10^{-5} are, at most, 4.7, 2.3, 1.0, and 0.16 mE_h, respectively, which indicates that the NPEs are less dependent within the given range of the threshold. The result indicates that the MC3I method with the proper parameter settings provides reliable PECs for the ground and excited states with quasi-degeneracy in the electronic structure.

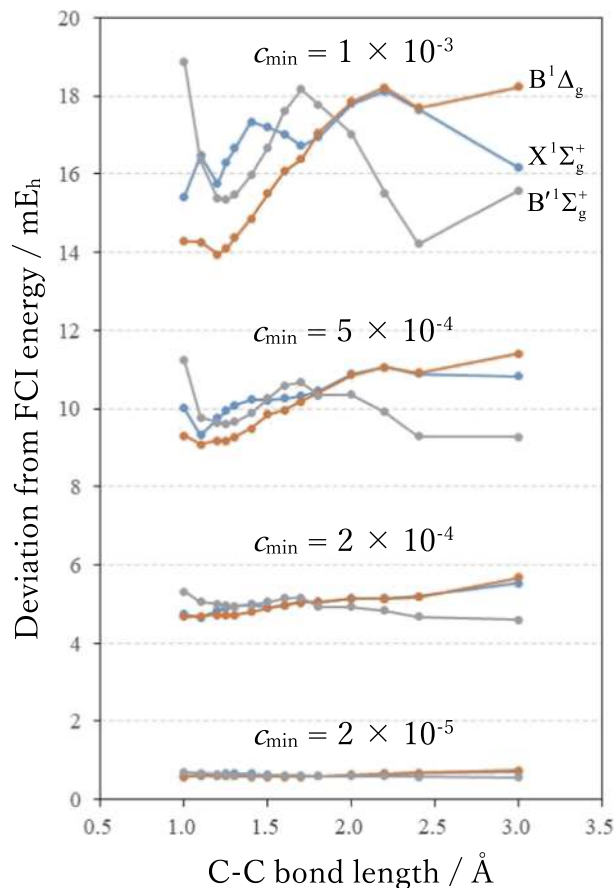


FIG. 8. Deviations of MC3I energies using $c_{\min} = 1 \times 10^{-3}$, 5×10^{-4} , 2×10^{-4} , and 2×10^{-5} for the lowest three singlet states from FCI values.

The energy convergences of the lowest three singlet states at C–C bond distances 1.25 and 3.0 Å are shown in Figs. 9(a) and 9(b), respectively. One of the ten MC3I and ten MCCI calculations is shown for clarity because the other nine MC3I

TABLE III. The MC3I calculations of the PECs of the ground and excited states of C_2 . Maximum, minimum, and non-parallelity errors in mE_h for the $X^1\Sigma_g^+$, $B^1\Delta_g$, and $B'^1\Sigma_g^+$ states. The results of $c_{\min} = 1 \times 10^{-3}$, 5×10^{-4} , 2×10^{-4} , and 2×10^{-5} are compared. The standard deviations over ten calculations are shown in parentheses.

State	Maximum errors	Minimum errors	NPE
(1) $c_{\min} = 1 \times 10^{-3}$			
$X^1\Sigma_g^+$	18.12 (3)	15.42 (2)	2.70
$B^1\Delta_g$	18.23 (3)	13.95 (2)	4.28
$B'^1\Sigma_g^+$	18.88 (2)	14.22 (4)	4.67
(2) $c_{\min} = 5 \times 10^{-4}$			
$X^1\Sigma_g^+$	11.060 (9)	9.315 (9)	1.745
$B^1\Delta_g$	11.41 (2)	9.085 (7)	2.33
$B'^1\Sigma_g^+$	11.23 (1)	9.27 (1)	1.96
(3) $c_{\min} = 2 \times 10^{-4}$			
$X^1\Sigma_g^+$	5.544 (7)	4.656 (4)	0.888
$B^1\Delta_g$	5.681 (9)	4.673 (4)	1.008
$B'^1\Sigma_g^+$	5.322 (2)	4.592 (2)	0.729
(4) $c_{\min} = 2 \times 10^{-5}$			
$X^1\Sigma_g^+$	0.715 (6)	0.5895 (9)	0.126
$B^1\Delta_g$	0.732 (6)	0.577 (2)	0.156
$B'^1\Sigma_g^+$	0.690 (2)	0.563 (3)	0.127

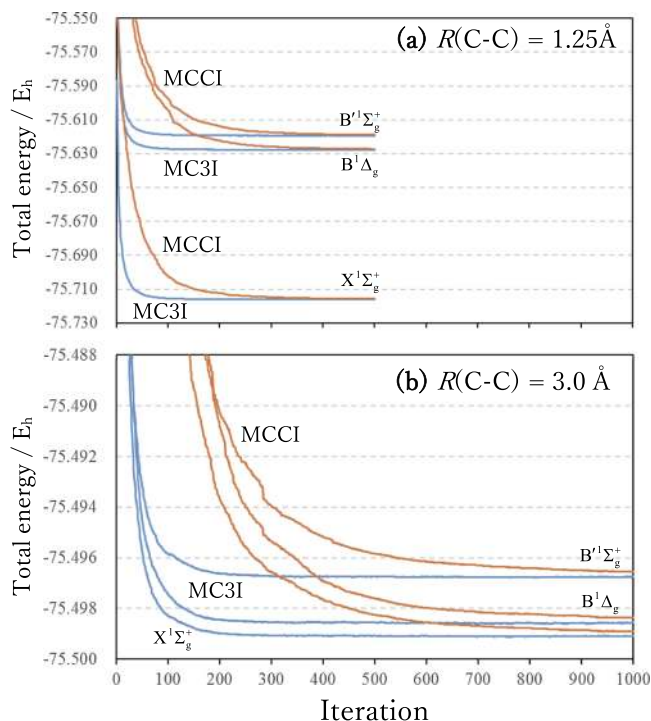


FIG. 9. MC3I and MCCI energy convergences of C_2 for the lowest three singlet states at C–C bond length = (a) 1.25 \AA and (b) 3.0 \AA .

and MCCI calculations show similar trends. The MC3I calculations were performed with $c_{\min} = 5.0 \times 10^{-4}$ and $N_w = 10\,000$, and a very similar number of determinants were generated in each step of the MCCI calculations. The result resembles that obtained for H_2O , as shown in Fig. 5. For the ground state at a C–C bond distance of 1.25 \AA , the MC3I calculations took 84.7 iterations on average until the energy difference from the converged energy became less than 1 mE_h , but 312.9 iterations were required in the MCCI calculations to achieve the same accuracy.

The convergence of excited states was also improved. As shown in Fig. 9(a), the convergence of the MCCI calculation for the excited states was worse than that for the ground state. For the $B^1\Delta_g$ and $B^1\Sigma_g^+$ states, 329.4 and 330.0 iterations were required until the energy difference from the converged energy became less than 1 mE_h . In the MC3I calculation, in which the FOCs to the excited states were sampled, the number of iterations was 72.1 and 74.7 for $B^1\Delta_g$ and $B^1\Sigma_g^+$ states, respectively.

It should be noted that at longer C–C separations, the convergence becomes slow, as shown in Fig. 9(b). At the C–C bond length of 3.0 \AA , the number of iterations until the energy difference converges to within 1 mE_h was 91.5, 105.4, and 97.4 for the $X^1\Sigma_g^+$, $B^1\Delta_g$, and $B^1\Sigma_g^+$ states in the MC3I calculation, respectively, whereas MCCI took 453.1, 507.2, and 491.7 iterations, respectively. The results clearly illustrate that a key feature for the acceleration of the MC convergence is the sampling of FOCs to construct the excited-state wave function.

V. CONCLUSION

In this paper, a newly selected CI method, MC3I, has been proposed. This method is one of the variants of the QMC

methods in the determinant space and is viewed as a stochastic version of the CIPSI method.⁹ To extend the applicability of the MCCI-type methods to larger molecules with quasi-degenerated electronic structures, an efficient algorithm has to be considered for the generation of the determinants in the MC sampling. We introduced sampling of the FOCs to all target states including ground and excited states. Similar to the branching factor f in MCCI, the convergence is controlled by setting the N_w parameter to a proper value. The accuracy on the energy is controlled by the coefficient threshold c_{\min} .

The efficiency and accuracy of the MC3I method were assessed by benchmark applications to H_2O in the ground state at the equilibrium and elongated bond distances and to the PECs of C_2 in the ground and excited states. Compared to the MCCI calculations, the MC3I calculations significantly reduce the number of iterations until the convergence. This successful trend was commonly observed for both the ground and excited states at any bond distance.

As for the PECs of the C_2 , the accuracy was well-balanced at different bond distances. The relative energies between states were in good agreement with the FCI values. The calculated NPEs were within 0.2 mE_h for the accurate c_{\min} (2×10^{-5}) and still within 5 mE_h (3.1 kcal/mol) for a loose c_{\min} (1×10^{-3}). The MC3I energies were slightly lower than the MCCI values in all test calculations because the important determinants can be sampled more in the same calculation condition. We note, however, that the MC3I energy is not size-extensive.

By accelerating the convergence of the ground and excited states, the application target of the MCCI framework has been expanded. The MCCI energies can be improved by considering dynamical correlations as the second-order perturbation.^{26,27} We are also developing a program to calculate the perturbation to the reference MC3I wave functions.

ACKNOWLEDGMENTS

This study is supported by a Grant-in-Aid for Challenging Exploratory Research (JSPS KAKENHI Grant No. JP16K13935). This study is also supported by JSPS KAKENHI Grant Nos. JP15H05805 and the FLAGSHIP2020 (priority study 5) program from MEXT. Part of the computations was carried out at ACCMS (Kyoto University).

¹B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, *Chem. Phys.* **48**, 157 (1980).

²J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, *J. Chem. Phys.* **89**, 2185 (1988).

³D. Ma, G. Li Manni, and L. Gagliardi, *J. Chem. Phys.* **135**, 044128 (2011).

⁴C. F. Bender and E. R. Davidson, *Phys. Rev.* **183**, 23 (1969).

⁵R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta* **35**, 33 (1974).

⁶H. Nakatsuji, *Chem. Phys. Lett.* **59**, 362 (1978); **67**, 329 (1979); **67**, 334 (1979).

⁷R. J. Buenker, S. D. Peyerimhoff, and W. Butscher, *Mol. Phys.* **35**, 771 (1978).

⁸M. Ehara, J. Hasegawa, and H. Nakatsuji, "SAC-CI method applied to molecular spectroscopy," in *Theory and Applications of Computational Chemistry: The First 40 Years*, A Volume of Technical and Historical Perspectives (Elsevier, Oxford, 2005).

⁹B. Huron, J. P. Malrieu, and P. Rancurel, *J. Chem. Phys.* **58**, 5745 (1973).

¹⁰J. C. Greer, *J. Chem. Phys.* **103**, 1821 (1995).

¹¹J. C. Greer, *J. Comput. Phys.* **146**, 181 (1998).

¹²Y. Ohtsuka and S. Nagase, *Chem. Phys. Lett.* **463**, 431 (2008).

- ¹³Y. Ohtsuka and S. Nagase, *Theor. Chem. Acc.* **130**, 501 (2011).
- ¹⁴G. H. Booth, A. J. W. Thom, and A. Alavi, *J. Chem. Phys.* **131**, 054106 (2009).
- ¹⁵F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, and C. J. Umrigar, *Phys. Rev. Lett.* **109**, 230201 (2012).
- ¹⁶S. Ten-no, *J. Chem. Phys.* **138**, 164126 (2013).
- ¹⁷N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, *J. Chem. Phys.* **145**, 044112 (2016).
- ¹⁸T. Zhang and F. A. Evangelista, *J. Chem. Theory Comput.* **12**, 4326 (2016).
- ¹⁹A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 3674 (2016).
- ²⁰S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, *J. Chem. Theory Comput.* **13**, 1595 (2017).
- ²¹F. A. Evangelista, *J. Chem. Phys.* **140**, 124114 (2014).
- ²²J. B. Schriber and F. A. Evangelista, *J. Chem. Phys.* **144**, 161106 (2016).
- ²³J. P. Coe, D. J. Taylor, and M. J. Paterson, *J. Chem. Phys.* **137**, 194111 (2012).
- ²⁴J. P. Coe and M. J. Paterson, *J. Chem. Phys.* **139**, 154103 (2013).
- ²⁵J. P. Coe, P. Murphy, and M. J. Paterson, *Chem. Phys. Lett.* **604**, 46 (2014).
- ²⁶J. P. Coe and M. J. Paterson, *J. Chem. Phys.* **137**, 204108 (2012).
- ²⁷T. P. Kelly, A. Perera, R. J. Bartlett, and J. C. Greer, *J. Chem. Phys.* **140**, 084114 (2014).
- ²⁸W. Gyrfy, R. J. Bartlett, and J. C. Greer, *J. Chem. Phys.* **129**, 064103 (2008).
- ²⁹ $\text{Int}(R)$ is the integer part of a real number R .
- ³⁰E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975).
- ³¹K. Ishimura, SMASH Code, <http://smash-qc.sourceforge.net/>.
- ³²M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993).
- ³³M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005).
- ³⁴M. Matsumoto and T. Nishimura, "Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Trans. Model. Comput. Simul.* **8**(1), 3–30 (1998).
- ³⁵J. Olsen, P. Jørgensen, H. Koch, A. Balkova, and R. J. Bartlett, *J. Chem. Phys.* **104**, 8007 (1996).
- ³⁶T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- ³⁷M. L. Abrams and C. D. Sherrill, *J. Chem. Phys.* **121**, 9211 (2004).
- ³⁸W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- ³⁹P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* **28**, 213 (1973).